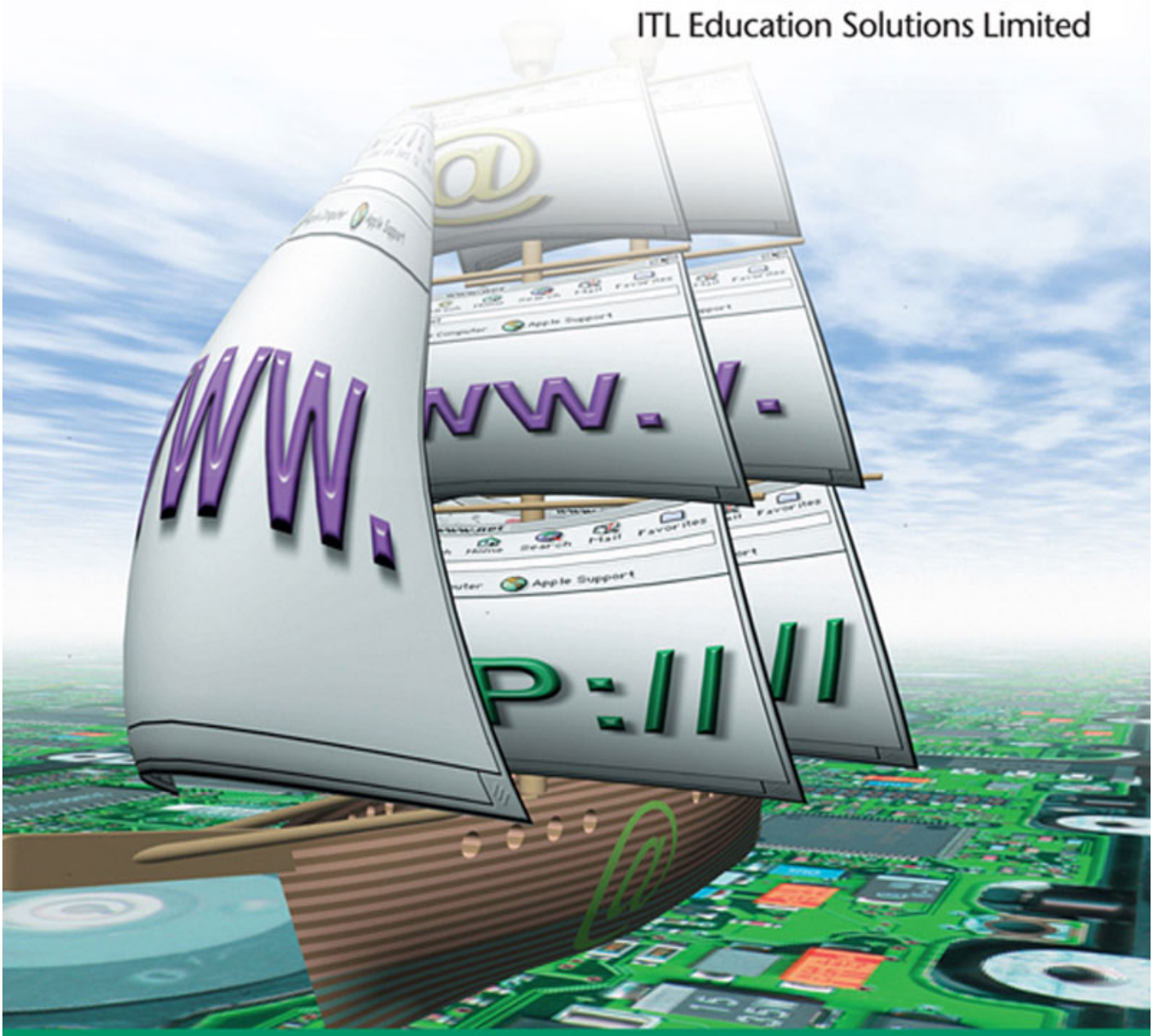
SECOND EDITION

Introduction to INFORMATION TECHNOLOGY

ITL Education Solutions Limited



ALWAYS LEARNING

PEARSON

INTRODUCTION TO INFORMATION TECHNOLOGY

SECOND EDITION

ITL Education Solutions Limited
Research and Development Wing
New Delhi

PEARSON

Delhi • Chennai • Chandigarh

Brief Contents

Preface

1. Computer Basics

2. Computer Organization and Architecture

3. Computer Memory and Storage

4. Input Output Media

5. Operating System

6. Microsoft Windows XP

7. Overview of Microsoft Windows 7

8. Information Technology Basics

9. Multimedia Essentials

10. Computer Programming and Languages

11. Computer Software

12. Introduction to Microsoft Office 2007

13. Microsoft Office Word 2003

14. Microsoft Office Excel 2003

15. Microsoft Office PowerPoint 2003

16. Data Communication and Computer Networks

17. The Internet

18. Internet Tools

19. Computer Security

20. Database Fundamentals

21. Structured Query Language (SQL)

22. Microsoft Office Access 2003

23. Current and Future Trends in IT

24. Artificial Intelligence (AI)

Contents

Preface

1. Computer Basics

1.1 Introduction

1.1.1 Characteristics of Computers

1.2 Evolution of Computers

1.3 Generations of Computers

1.3.1 First Generation (1940 to 1956): Vacuum Tubes

1.3.2 Second Generation (1956 to 1963): Transistors

1.3.3 Third Generation (1964 to Early 1970s): Integrated Circuits

1.3.4 Fourth Generation (Early 1970s to Till Date): Microprocessors

1.3.5 Fifth Generation (Present and Beyond): Artificial Intelligence

1.4 Classification of Computers

1.4.1 Microcomputers

1.4.2 Minicomputers

1.4.3 Mainframes

1.4.4 Supercomputers

1.5 The Computer System

1.5.1 Components of a Computer System

1.5.2 How Does a Computer Work?

1.6 Applications of Computers

Let Us Summarize

Exercises

Answers

2. Computer Organization and Architecture

2.1 Introduction

2.2 Central Processing Unit

2.2.1 Arithmetic/Logic Unit

2.2.2 Registers

2.2.3 Control Unit

2.2.4 System Bus

2.2.5 Main Memory Unit

2.2.6 Cache Memory

2.3 Communication among Various Units

2.3.1 Processor to Memory Communication

2.3.2 Processor—I/O Devices Communication

2.4 Instruction Format

2.5 Instruction Cycle

2.6 Instruction Set

2.6.1 CISC Architecture

2.6.2 RISC Architecture

2.6.3 Comparing the CISC with RISC

2.7 Inside a Computer

2.7.1 Power Supply (SMPS)

2.7.2 Motherboard

2.7.3 Ports and Interfaces

2.7.4 Expansion Cards

2.7.5 Ribbon Cables

2.7.6 Memory Chips

2.7.7 Storage Devices

2.7.8 Processors

2.8 Data Representation in Computers

2.8.1 Types of Number System

2.8.2 Conversion between Number Bases

2.9 Coding Schemes

Let Us Summarize

Exercises

Answers

3. Computer Memory and Storage

3.1 Introduction

3.1.1 Memory Representation

3.2 Memory Hierarchy

3.3 Random Access Memory

3.3.1 Types of RAM

3.4 Read Only Memory

3.4.1 Types of ROM

3.5 RAM, ROM and CPU Interaction

3.6 Types of Secondary Storage Devices

3.7 Magnetic Tape

3.7.1 Magnetic Tape Organization

3.7.2 Advantages of Magnetic Tapes

3.7.3 Disadvantages of Magnetic Tapes

3.8 Magnetic Disk

3.8.1 Storage Organization of a Magnetic Disk

3.8.2 Accessing Data from Magnetic Disk

3.9 Types of Magnetic Disks

3.9.1 Floppy Disk

3.9.2 Hard Disk

3.9.3 Zip Disk

3.9.4 Advantages of Magnetic Disks

3.9.5 Disadvantages of Magnetic Disks

3.10 Optical Disk

3.10.1 Storage Organization of Optical Disk

3.10.2 Access Mechanism of Optical Disk

3.11 Types of Optical Disks

3.11.1 Compact Disk

3.11.2 Digital Versatile Disk

3.11.3 Blu-ray Disk

3.11.4 Advantages of Optical Disks

3.11.5 Disadvantages of Optical Disks

3.12 Magneto-optical Storage Devices

3.12.1 Basics of MO Reading

3.12.2 Basics of MO Recording

3.13 Universal Serial Bus

3.13.1 Pen Drive

3.13.2 External Hard Disk Drive

3.14 Memory Stick

3.15 Mass Storage Devices

Let Us Summarize

Exercises

Answers

4. Input Output Media

4.1 Introduction

4.1.1 Importance of Input/Output Devices

4.2 Types of Input Devices

4.2.1 Keyboard

4.2.2 Pointing Devices

4.2.3 Speech Recognition

4.2.4 Digital Camera

4.2.5 Webcam

4.2.6 Scanners

4.2.7 Optical Character Recognition

4.2.8 Optical Mark Recognition

4.2.9 Magnetic-ink Character Recognition

4.2.10 Bar Code Reader

4.3 Types of Output Devices

4.3.1 Printers

4.3.2 Plotters

4.3.3 Computer Output Microfilm

4.3.4 Monitor

4.3.5 Voice Response System

4.3.6 Projector

4.3.7 Electronic Whiteboard

4.3.8 Headphone and Headset

4.4 Computer Terminals

Let Us Summarize

Exercises

Answers

5. Operating System

5.1 Introduction

5.2 Operating System: Definition

5.3 Evolution of Operating System

5.4 Types of Operating Systems

5.5 Functions of an Operating System

5.5.1 Process Management

5.5.2 Memory Management

5.5.3 File Management

5.5.4 Device Management

5.5.5 Security Management

5.5.6 User Interface

Let Us Summarize

Exercises

Answers

6. Microsoft Windows XP

6.1 Introduction

6.2 Features of Windows XP

6.3 Starting Windows XP Professional

6.4 The Desktop

6.4.1 Taskbar

6.4.2 Customizing Taskbar

6.4.3 Desktop Icons

6.4.4 Customizing the Desktop

6.5 My Computer

6.6 Working with Files and Folders

6.6.1 Recycle Bin

6.7 Windows Explorer

6.8 Windows Help and Support Centre

6.9 Searching in Windows XP

6.10 System Utilities in Windows XP

6.11 Control Panel

6.12 Windows Accessories

6.12.1 Calculator

6.12.2 Notepad

6.12.3 Paint

Let Us Summarize

Exercises

Answers

7. Overview of Microsoft Windows 7

7.1 Introduction

7.2 Different Editions of Windows 7

7.3 New Style of Windows 7

7.3.1 Aero Themes

7.3.2 Gadgets

7.3.3 Taskbar

7.3.4 Start Menu

7.3.5 Aero Effects

7.3.6 Jump Lists

7.4 Windows 7 Accessories

Let Us Summarize

Exercises

Answers

8. Information Technology Basics

8.1 Introduction

8.2 Information

8.2.1 Prerequisites of Information

8.2.2 Need for Information

8.3 Technology

8.4 Information Technology

8.4.1 Components of Information Technology

8.5 Present Scenario

8.6 Role of Information Technology

8.6.1 Information Technology in Business

8.6.2 Information Technology in Manufacturing

8.6.3 Information Technology in Mobile Computing

8.6.4 Information Technology in Public Sector

8.6.5 Information Technology in Defence Services

8.6.6 Information Technology in Media

8.6.7 Information Technology in Education

8.6.8 Information Technology in Publication

8.7 Information Technology and the Internet

8.8 Careers in the IT Industry

Let Us Summarize

Exercises

Answers

9. Multimedia Essentials

9.1 Introduction

9.2 Multimedia: Definition

9.3 Building Blocks of Multimedia

9.3.1 Text

9.3.2 Graphics

9.3.3 Audio

9.3.4 Video

9.3.5 Animation

9.4 Multimedia System

9.4.1 Desirable Features of Multimedia System

9.4.2 Data Compression

9.5 Multimedia Applications

9.5.1 Multimedia and Education

9.5.2 Multimedia and Entertainment

9.5.3 Multimedia and Training

9.5.4 Multimedia and Business

9.5.5 Multimedia Interactive Kiosks

9.6 Virtual Reality

9.6.1 Characteristics of Virtual Reality

Let Us Summarize

Exercises

Answers

10. Computer Programming and Languages

10.1 Introduction

10.1.1 Developing a Program

10.1.2 Program Development Cycle

10.2 Algorithm

10.3 Flowchart

10.3.1 Flowchart Symbols

10.3.2 Guidelines for Preparing Flowcharts

10.3.3 Benefits of Flowcharts

10.3.4 Limitations of Flowcharts

10.4 Decision Tables

10.4.1 Steps to Develop a Decision Table

10.4.2 Benefits of Decision Tables

10.4.3 Limitations of Decision Tables

10.5 Pseudocode

10.5.1 Pseudocode Guidelines

10.5.2 Benefits of Pseudocode

10.5.3 Limitations of Pseudocode

10.6 Program Control Structures

10.6.1 Sequence Control Structure

10.6.2 Selection Control Structure

10.6.3 Repetition Control Structure

10.7 Programming Paradigms

10.7.1 Unstructured Programming

10.7.2 Structured Programming

10.7.3 Object-oriented Programming

10.8 Characteristics of a Good Program

10.9 Programming Languages

10.9.1 Types of Programming Languages

10.10 Generations of Programming Languages

10.10.1 First Generation: Machine Language

10.10.2 Second Generation: Assembly Language

10.10.3 Third Generation: High-level Language

10.10.4 Fourth Generation

10.10.5 Fifth Generation: Very High-level Languages

10.11 Features of a Good Programming Language

Let Us Summarize

Exercises

Answers

11. Computer Software

11.1 Introduction

11.2 Software: Definition

11.2.1 Relationship between Software and Hardware

11.3 Categories of Software

11.3.1 System Software

11.3.2 Application Software

11.4 Installing and Uninstalling Software

11.5 Software Piracy

11.5.1 Types of Software Piracy

11.5.2 Preventing Software Piracy

11.6 Software Terminologies

Let Us Summarize

Exercises

Answers

12. Introduction to Microsoft Office 2007

12.1 Introduction

12.2 Introduction to Office 2007 User Interface

12.3 Microsoft Office Applications

12.3.1 Microsoft Office Word 2007

12.3.2 Microsoft Office Excel 2007

12.3.3 Microsoft Office Access 2007

12.3.4 Microsoft Office PowerPoint 2007

Let Us Summarize

Exercises

Answers

13. Microsoft Office Word 2003

13.1 Introduction

13.2 Starting Microsoft Office Word 2003

13.2.1 Microsoft Word Environment

13.2.2 Getting Help

13.3 Working with Word 2003 Documents

13.3.1 Creating a New Document

13.3.2 Opening an Existing Document

13.3.3 Saving a Document

13.3.4 Closing a Document and Exiting Microsoft Word

13.4 Working with Text

13.4.1 Inserting Text

13.4.2 Selecting Text

13.4.3 Formatting the Text

13.4.4 Aligning and Indenting Text

13.4.5 Formatting Style

13.4.6 Format Painter

13.4.7 Using Cut, Copy, and Paste in Word

13.4.8 Finding and Replacing Text

13.4.9 Undo and Redo

13.4.10 Bullets and Numbering

13.4.11 Header and Footer

13.5 Working with Tables

13.5.1 Table Properties

13.5.2 Moving and Resizing Table

13.5.3 Formatting Text in Table

13.5.4 Adding and Deleting Rows and Columns

13.5.5 Converting Text to Table

13.5.6 Converting Table to Text

13.6 Checking Spelling and Grammar

13.6.1 Using the Auto Spell and Auto Grammar Features

13.6.2 Using Spelling and Grammar Dialog Box

13.6.3 Using AutoCorrect Options

13.6.4 Using Synonyms and Thesaurus

13.7 Adding Graphics to Document

13.7.1 Pictures

13.7.2 Editing a Picture

13.7.3 Drawing Objects

13.8 Mail Merge

13.8.1 Components of a Mail Merge

13.8.2 Mail Merge Steps

13.9 Printing a Document

13.9.1 Setting Page Margins

13.9.2 Print Preview

13.9.3 Printing the Document

Let Us Summarize

Exercises

Answers

14. Microsoft Office Excel 2003

14.1 Introduction

14.2 Starting Microsoft Office Excel 2003

14.2.1 Microsoft Excel Environment

14.3 Working with Excel Workbook

14.3.1 Creating an Excel Workbook

14.3.2 Opening an Existing Workbook

14.3.3 Saving Workbook

14.3.4 Closing a Workbook

14.4 Working with Worksheet

14.4.1 Entering Data in Cells

14.4.2 Navigating through Cells

14.4.3 Naming of a Range of Cells

14.4.4 Editing a Worksheet

14.4.5 Using Cut, Copy, and Paste in Excel

14.4.6 Finding and Replacing Cells

14.4.7 Undo and Redo

14.5 Formulas and Functions

14.5.1 Functions

14.5.2 Using AutoSum

14.5.3 Formulas

14.6 Inserting Charts

14.6.1 Creating Charts

14.6.2 Using Chart Toolbar

14.6.3 Saving a Chart

14.7 Sorting

14.8 Printing in Excel

14.8.1 Setting Page Layout

14.8.2 Print Preview

14.8.3 Printing Worksheets

Let Us Summarize

Exercises

Answers

15. Microsoft Office PowerPoint 2003

15.1 Introduction

15.2 Starting Microsoft Office PowerPoint 2003

15.2.1 Microsoft PowerPoint Environment

15.3 Working with PowerPoint 2003

15.3.1 Creating a New Presentation

15.3.2 Opening a Presentation

15.3.3 Saving a Presentation

15.3.4 Closing a Presentation

15.4 Working with Different Views

15.5 Designing Presentation

15.5.1 Choosing Design Template

15.5.2 Using Masters

15.5.3 Working with Slides

15.5.4 Working with Text

15.5.5 Formatting the Text

15.5.6 Working with Graphics

15.5.7 Adding Multimedia

15.5.8 Running Slides

15.5.9 Package for CD

15.6 Printing in PowerPoint

15.6.1 Setting Page Layout

15.6.2 Printing Presentation

Let Us Summarize

Exercises

Answers

16. Data Communication and Computer Networks

16.1 Introduction

16.2 Data Communication

16.2.1 Components of Data Communication

16.2.2 Data Transmission Mode

16.2.3 Analog and Digital Data Transmission

16.2.4 Data Communication Measurement

16.3 Transmission Media

16.3.1 Guided/Wired Media

16.3.2 Unguided/Wireless Media

16.4 Modulation

16.4.1 Amplitude Modulation (AM)

16.4.2 Frequency Modulation (FM)

16.4.3 Phase Modulation (PM)

16.5 Multiplexing

16.5.1 Multiplexers

16.5.2 Asynchronous and Synchronous Transmission

16.6 Switching

16.6.1 Circuit Switching

16.6.2 Packet Switching

16.6.3 Message Switching

16.7 Computer Network

16.7.1 Types of Computer Networks

16.8 Network Topologies

16.8.1 Bus Topology

16.8.2 Ring Topology

16.8.3 Star Topology

16.8.4 Tree Topology

16.8.5 Mesh Topology

16.9 Communication Protocol

16.9.1 The OSI Model

16.10 Network Devices

Let Us Summarize

Exercises

Answers

17. The Internet

17.1 Introduction

17.2 Evolution of Internet

17.3 Basic Internet Terms

17.4 Getting Connected to the Internet

17.5 Internet Applications

17.5.1 Electronic Mail (E-mail)

17.5.2 File Transfer Protocol (FTP)

17.5.3 Telnet

17.5.4 Internet Relay Chat (IRC)

17.5.5 Chatting and Instant Messaging

17.5.6 Internet Telephony

17.5.7 Video Conferencing

17.5.8 Commerce through the Internet

17.5.9 Newsgroups (Usenet)

17.5.10 Mailing Lists (Listserver)

17.6 Data Over the Internet

17.6.1 Static Component

17.6.2 Dynamic Component

Let Us Summarize

Exercises

Answers

18. Internet Tools

18.1 Introduction

18.2 Web Browser

18.2.1 Web Browser Features

18.2.2 Opening Internet Explorer

18.2.3 Internet Explorer Environment

18.3 Browsing Internet Using Internet Explorer

18.4 Electronic Mail (E-mail)

18.4.1 E-mail Address Structure

18.4.2 Checking E-mails

18.4.3 Sending E-mails

18.4.4 E-mail Attachments

18.4.5 How E-mail Works

18.4.6 Netiquette

18.4.7 Advantages and Disadvantages of E-mail

18.5 E-mail Client

18.5.1 Features of E-mail Client

18.5.2 Opening Outlook Express

18.5.3 Creating an Account

18.5.4 Sending E-mails Using Outlook Express

18.5.5 Checking E-mails Using Outlook Express

18.5.6 Creating Newsgroup Account

18.6 Search Engines

18.6.1 Searching the Internet

18.6.2 Refining the Search

18.7 Instant Messaging

18.7.1 Features of Messenger

18.7.2 Getting Started with Yahoo Messenger

18.7.3 Sending Instant Messages

18.7.4 Sending Files

Let Us Summarize

Exercises

Answers

19. Computer Security

19.1 Introduction

19.2 Computer Security: Definition

19.2.1 Security Threats

19.3 Malicious Programs

19.3.1 Virus

19.3.2 Other Destructive Programs

19.3.3 Affecting Computer System

19.3.4 Protecting Computer System

19.4 Cryptography

19.4.1 Secret Key Cryptography

19.4.2 Public Key Cryptography

19.4.3 Hash Function

19.5 Digital Signature

19.5.1 How Is Digital Signature Used?

19.6 Firewall

19.6.1 How Firewall Works?

19.6.2 Types of Firewall

19.7 User Identification and Authentication

19.7.1 User Requirements

19.7.2 User Possessions

19.7.3 Biometric Technique

19.8 Data Backup and Recovery

19.8.1 Backup Methods

19.8.2 Online Backups

19.8.3 Recovery Tools

19.9 Security Awareness and Policies

19.9.1 Awareness

19.9.2 Security Policy

Let Us Summarize

Exercises

Answers

20 Database Fundamentals

20.1 Introduction

20.1.1 Data versus Information

20.2 Database: Definition

20.2.1 Basic Database Terms

20.3 Logical Data Concepts

20.3.1 Entity

20.3.2 Attribute

20.3.3 Relationship

20.4 Physical Data Concepts

20.4.1 Sequential Files

20.4.2 Direct Files

20.4.3 Indexed Sequential Files

20.5 Database Management System

20.5.1 Need for DBMS

20.5.2 Benefits of DBMS

20.5.3 Components of a DBMS

20.5.4 Database Administrator

20.5.5 Database Languages

20.5.6 Data Dictionary

20.6 DBMS Architecture

20.6.1 Data Independence

20.7 Database Models

20.7.1 Hierarchical Database Model

20.7.2 Network Database Model

20.7.3 Relational Database Model

20.7.4 Object-oriented Database Model

20.8 Normalization

20.8.1 First Normal Form (1NF)

20.8.2 Second Normal Form (2NF)

20.8.3 Third Normal Form (3NF)

20.9 Types of Databases

20.9.1 Centralized Database

20.9.2 Distributed Database

20.10 Data Warehousing and Data Mining

Let Us Summarize

Exercises

Answers

21. Structured Query Language (SQL)

21.1 Introduction

21.1.1 Communicating with Database Using SQL

21.1.2 Characteristics of SQL

21.1.3 Brief History of SQL

21.2 Getting Started with SQL

21.2.1 Data Types

21.2.2 SQL Syntax

21.3 DDL Commands

21.3.1 CREATE TABLE Command

21.3.2 ALTER TABLE Command

21.3.3 DROP TABLE Command

21.4 DML Commands

21.4.1 INSERT Command

21.4.2 UPDATE Command

21.4.3 DELETE Command

21.4.4 SELECT Command

21.5 Querying Multiple Tables (SQL Joins)

21.5.1 Combining Queries Using UNION

21.6 Nesting Select Statements (Subquery)

Let Us Summarize

Exercises

Answers

22. Microsoft Office Access 2003

22.1 Introduction

22.2 Starting Microsoft Office Access 2003

22.3 Working with Access Database

22.3.1 Creating and Saving a Database

22.3.2 Opening a Database

22.3.3 Closing a Database and Exiting Microsoft Access

22.4 Microsoft Access Environment

22.5 Working with Tables

22.5.1 Creating a Table in Design View

22.5.2 Creating a Table Using Wizard

22.5.3 Adding, Modifying, and Deleting Records

22.5.4 Navigating a Table

22.6 Queries

22.6.1 Creating a Simple Query Using Wizard

22.6.2 Running a Query

22.6.3 Sorting Query Results

22.6.4 Using SQL in Microsoft Access

22.7 Forms

22.7.1 Creating a Form Using Wizard

22.8 Reports

22.8.1 Creating a Report Using Wizard

Let Us Summarize

Exercises

Answers

23 Current and Future Trends in IT

23.1 Introduction

23.2 Electronic Commerce (E-Commerce)

23.2.1 Types of E-Commerce

23.2.2 Processes in E-Commerce

23.2.3 Types of Electronic Payment

23.2.4 Advantages of E-Commerce

23.2.5 Disadvantages of E-Commerce

23.3 Electronic Data Interchange (EDI)

23.3.1 How EDI Works?

23.3.2 EDI Benefits

23.3.3 EDI Limitations

23.4 Wireless Application Protocol (WAP)

23.5 Smart Card

23.6 Internet Prootcol Television (IPTV)

23.7 Blogging

23.8 Radio Frequency Identification (RFID)

23.9 Brain Computer Interface (BCI)

23.10 Imminent Technologies

23.10.1 Pen Size Computer

23.10.2 Napkin PC

23.10.3 Roll Top Computer

23.10.4 B-membrane PC

23.10.5 USB Memory Stick

Let Us Summarize

Exercises

Answers

24. Artificial Intelligence (AI)

24.1 Introduction

24.1.1 What Is Intelligence?

24.1.2 Turing Test

24.2 Artificial Intelligence: Definition

24.2.1 Domain Areas of AI

24.2.2 Brief History of AI

24.3 Knowledge-based Systems

24.3.1 Knowledge Acquisition

24.3.2 Knowledge Organization

24.3.3 Knowledge Manipulation

24.3.4 Knowledge Representation

24.4 Searching Techniques

24.4.1 Uninformed Search

24.4.2 Informed Search

24.5 Natural Language Processing (NLP)

24.5.1 Uses of NLP

24.6 Expert Systems

24.6.1 Characteristics of Expert System

24.6.2 Applications Areas of Expert System

24.6.3 Users in Expert Systems

24.6.4 Benefits of Expert Systems

24.6.5 Limitations of Expert Systems

24.7 Neural Networks

24.7.1 Benefits of Neural Networks

24.7.2 Applications of Neural Networks

24.8 AI Languages

Let Us Summarize

Exercises

Answers

Preface

The world today is witnessing a new kind of revolution—the *Information Revolution*—ushered in by tchnology. This revolution is far more sweeping than any other revolution in history in its reach and influence, bringing fundamental changes in all aspects of our life.

Information technology (IT) is the engine used to drive useful information systems. This includes computers, software, Internet/intranet and telecommunication systems. IT provides the means for collecting, storing, encoding, processing, analyzing, transmitting, receiving and printing text, audio or video information.

In the present context, IT is not merely the matter of machines and microchips. In fact, now it regulates all aspects of our lives including how we live, how we work, how we communicate, how we relax, how we travel and how we talk. This universal applicability of IT has made it the most important part of all technological curricula throughout the world. The future of IT promises to be more dynamic. Researches in the field of processor performance, chip miniaturization, artificial intelligence (AI) and other areas are setting up new standards every day. Keeping pace with these changing trends and to meet the requirements of modern curricula, the book has been revised.

The revised edition of *Introduction to Information Technology* has become more comprehensive with several enhancements. Similar to its previous edition, it also dwells on fundamentals that one must learn in order to pursue virtually any endeavor in the IT field. In addition, it covers some latest trends of IT such as Internet protocol television (IPTV), radio frequency

identification (RFID), brain computer interface (BCI), blogging and many more. The main objective is to teach IT concepts to people who will be the end users and to those who need some insight into IT. Our approach is to strike a balance between the why and the how of IT use.

This book trains students with the concepts needed to lay a solid foundation for joining this exciting field. It is important to learn the fundamentals so that one has the tools necessary to learn emerging technologies as they become more common. In simple words, reading this book will benefit anyone who wishes to arm themselves with substantive computer skills.

Our long experience of training computer professionals has enabled us to come out with a book that is reader friendly and teaches important concepts in an easy-to-understand format. The treatment of the subject matter is kept quite simple and concepts are explained with the help of examples and illustrations. The end result promises to be an outstanding learning material for all those who are interested in the body of knowledge we call *Information Technology*.

FOCUS

The experts and educators are faced with the challenging task of developing courses to teach emerging computer science concepts in college and university curricula. Such an undertaking requires adequate literature and explanation of key computer concepts at an introductory level. The computer textbooks used in colleges are either too advanced that they have less instructional value or are too focused on describing particular tools such as word processing, multimedia and spreadsheet applications. Some start at a very high level of complexity because of which students get lost in details before they grasp the fundamental concepts, while others cover material without cohesion between topics. Keeping all

this in mind, this book has been developed to present the fundamentals of IT in an easily understandable manner. The book has been designed to assist the reader to grasp the fundamentals of IT.

UNIQUE FEATURES

- The book is written in a clear, concise and lucid manner and is student friendly.
- Independent thinking and research on the subject imparts a practitioner's perspective to it and makes the text thoroughly enjoyable.
- Emphasis is laid on practical application of theory. All ideas and concepts are presented with clear examples.
- Historical perspective describes components, systems or events from the past.
- The text is well structured and well supported with suitable diagrams.
- Fact files and things-to-remember boxes provide a snapshot of the subject covered.
- Major applications of IT and their strategic implications for business and social activities are described.
- Inter-chapter dependencies are kept to a minimum.
- Chapter objectives, a summary at the end of each chapter and review questions are also present.
- A comprehensive index is presented at the end of the book for quick access to the topics.

KEY ADDITIONS TO THE SECOND EDITION

Although several enhancements have been made to the text, we are highlighting the key additions to this new edition here:

- Some new secondary storage devices such as pen drive, memory stick, blue-ray disk and external hard disk have been added to *Chapter 3*.
- Some more input/output devices such as trackpad, webcam, projector, electronic whiteboard and headphone and headset have been included in *Chapter 4*.
- The basic Windows accessories including Calculator, Notepad and Paint have been discussed in *Chapter 6*.
- Two new chapters providing an overview of Microsoft Windows 7 and Microsoft Office 2007 have been added.
- The concepts of word processing, spreadsheets, PowerPoint presentations and relational database are now discussed using Microsoft Office 2003.
- Chapter 23, *Current and Future Trends in IT,* has been thoroughly revised to include the technologies currently in use such as RFID, IPTV and BCI as well as some imminent technologies such as the concept of

Pen PC, Napkin PC, Roll top computer, B-membrane PC and USB memory stick.

TARGET AUDIENCE

1. Primary usage as a textbook for students of PGDCA, BCA, BE, BIT, BSc (IT), BTech, MSc (IT) and MCA courses offered by computer science, computer engineering and IT departments in colleges and universities.
2. Also useful as a textual resource in vocational training courses offered by many colleges.
3. As a self-study guide for non-computer science professionals.

This book serves the requirements of both students and professionals. In fact, it is aimed at all those people who wish to know all about IT fundamentals.

CHAPTER ORGANIZATION

- *Chapter 1* provides an overview of the history of computers and basic digital computer concepts.
- *Chapter 2* describes how computations are carried out using arithmetic and logic units, registers and control unit.
- *Chapter 3* introduces the reader to computer memory including an in-depth understanding of RAM and ROM. The chapter also describes the secondary storage and its types such as magnetic tapes, magnetic disks, optical disks and memory sticks.
- *Chapter 4* describes how a computer interacts with input devices such as keyboard and mouse and with the output devices such as printer, monitor and projector.
- *Chapter 5* provides an overview of the operating system along with its types and functions.
- *Chapter 6* discusses Microsoft's Windows XP along with its in-built tools and features.
- *Chapter 7* provides an overview of the latest version of Windows, that is, Microsoft Windows 7.
- *Chapter 8* introduces the reader to IT. The chapter discusses the impact of IT in various areas such as business, defense, public sector and media.
- *Chapter 9* provides a broad grounding in multimedia.
- *Chapter 10* deals essentially with programming and various programming tools such as algorithm, flowchart and pseudocode. The chapter also discusses commonly used programming languages.
- *Chapter 11* provides the fundamentals of computer software and its relationship with the hardware in building up a fully functional computer system.
- *Chapter 12* discusses the user interface of Microsoft Office 2007. It also discusses enhanced and new features of its core applications including Microsoft Office Word, Excel, Access and PowerPoint.

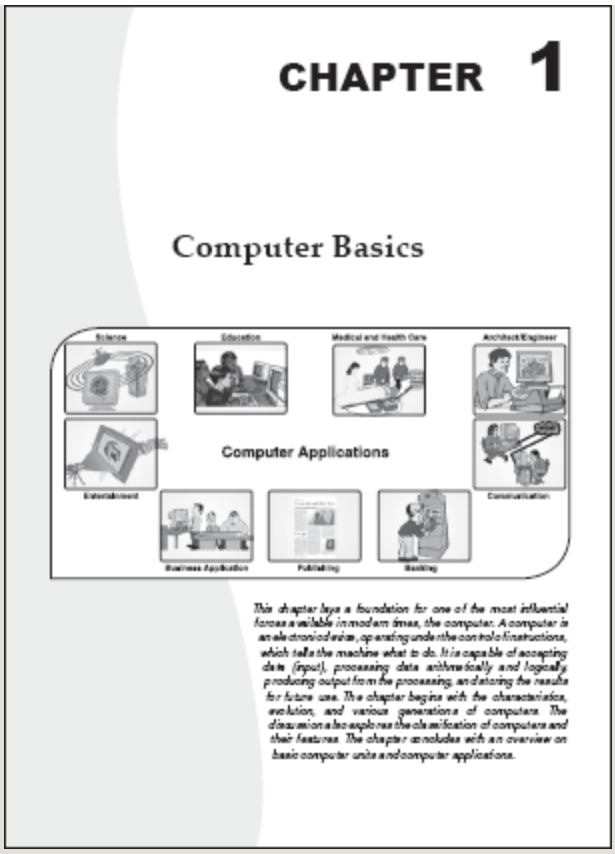
- *Chapter 13* provides a foundation to harness the power of word processing using Microsoft Office Word 2003.
- *Chapter 14* provides a base for the reader to harness the power of a spreadsheet program using Microsoft Office Excel 2003.
- *Chapter 15* provides a foundation for the reader to harness the power of creating presentations using Microsoft Office PowerPoint 2003.
- *Chapter 16* deals with data communication and computer network along with important concepts like transmission modes, network topologies and the OSI reference model.
- *Chapter 17* helps in building an understanding of the basic Internet structure and processes.
- *Chapter 18* introduces the reader to the different Internet tools such as browsing, searching and sending and receiving e-mails.
- *Chapter 19* provides a foundation in the field of computer security. Computer security is applied to the computer systems to ensure the availability, integrity and confidentiality of information against unwanted access, damage and modification.
- *Chapter 20* provides the basic understanding of database management system along with its components and architecture.
- *Chapter 21* introduces the reader to structured query language (SQL), which is a standard computer language for accessing and manipulating databases.
- *Chapter 22* provides a foundation to harness the power of databases using Microsoft Office Access 2003.
- *Chapter 23* provides a base for the reader to harness the power of IT in areas such as e-commerce, wireless applications and communication. It also discusses some new technologies that have been proposed for the future.
- *Chapter 24* introduces the reader to the fundamentals of AI.

OUTSTANDING PEDAGOGICAL FEATURES

The outstanding pedagogical features in the book have been designed to plan a clear and well-targeted route map for students' understanding and knowledge enhancement. These features are:

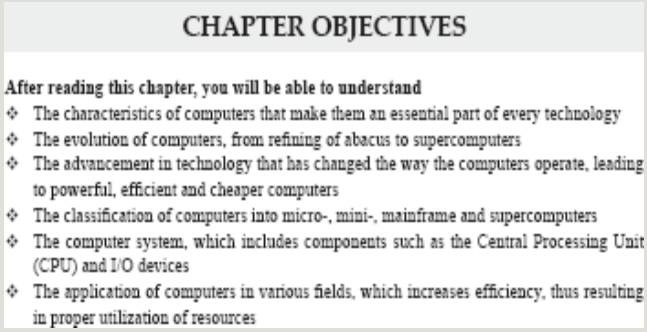
1. **Chapter Openers**

   Every chapter begins with a pictorial/ diagrammatic presentation and a nutshell description of what lies ahead in the chapter.



2. **Chapter Objectives**

An exhaustive list of chapter objectives clearly specifies learning outcomes.



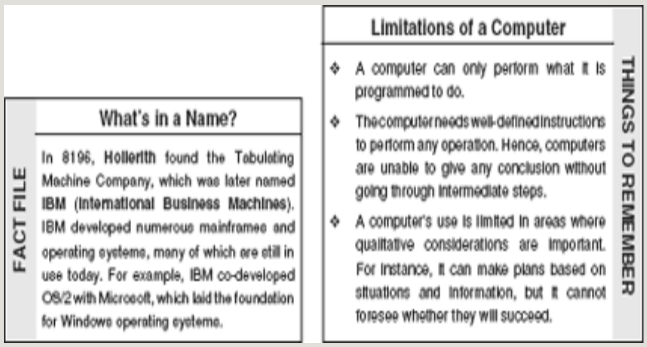
CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand

- The characteristics of computers that make them an essential part of every technology
- The evolution of computers, from refining of abacus to supercomputers
- The advancement in technology that has changed the way the computers operate, leading to powerful, efficient and cheaper computers
- The classification of computers into micro-, mini-, mainframe and supercomputers
- The computer system, which includes components such as the Central Processing Unit (CPU) and I/O devices
- The application of computers in various fields, which increases efficiency, thus resulting in proper utilization of resources

3. **Fact File and Things-to-Remember Box**

Regular fact files and things-to-remember boxes, which provide a snapshot of the subject covered, enable students to know facts about a particular section and help them to quickly revise the concepts.



FACT FILE

What's in a Name?

In 8196, **Hollerith** found the Tabulating Machine Company, which was later named **IBM (International Business Machines)**. IBM developed numerous mainframes and operating systems, many of which are still in use today. For example, IBM co-developed OS/2 with Microsoft, which laid the foundation for Windows operating systems.

THINGS TO REMEMBER

Limitations of a Computer

- A computer can only perform what it is programmed to do.
- The computer needs well-defined instructions to perform any operation. Hence, computers are unable to give any conclusion without going through intermediate steps.
- A computer's use is limited in areas where qualitative considerations are important. For instance, it can make plans based on situations and information, but it cannot foresee whether they will succeed.

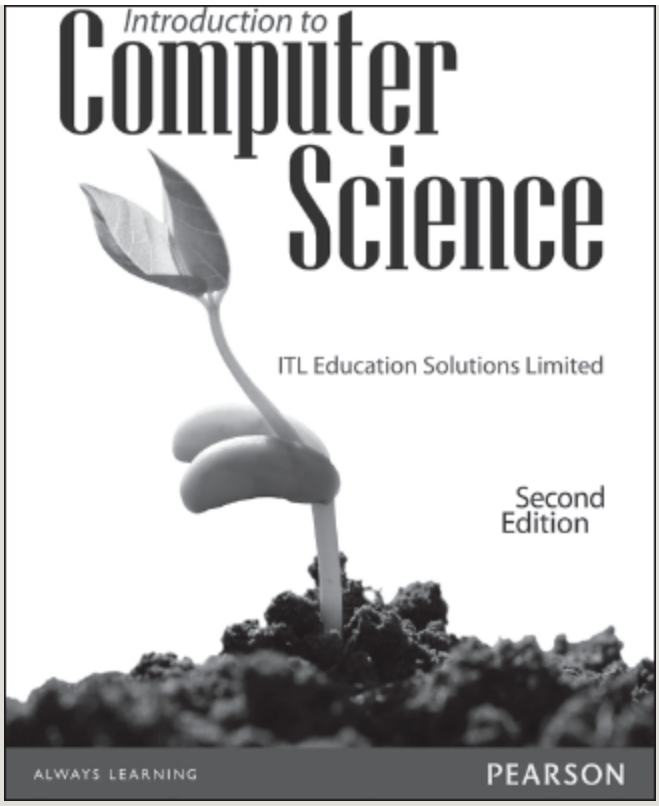
4. **Further Reading**

   For more details on the following chapters, refer to the book *Introduction to Computer Science*, second edition, published by Pearson Education.

   1. Chapter 1 Computer Basics
   2. Chapter 2 Computer Organization and Architecture
   3. Chapter 3 Computer Memory and Storage
   4. Chapter 4 Input Output Media
   5. Chapter 5 Operating System
   6. Chapter 9 Multimedia Essentials
   7. Chapter 10 Computer Programming and Languages
   8. Chapter 11 Computer Software
   9. Chapter 16 Data Communications and Computer Networks
   10. Chapter 17 The Internet
   11. Chapter 20 Database Fundamentals

5. **Exercises**

   End-of-chapter exercises with a set of objective and descriptive questions aim to develop wide and varied skill sets in the students.



INSTRUCTOR'S RESOURCE CENTER

A PowerPoint presentation is available to aid instructors in teaching the subject. The presentation is accessible from www.pearsoned.co.in/ITLEducationSolutionsLimited.

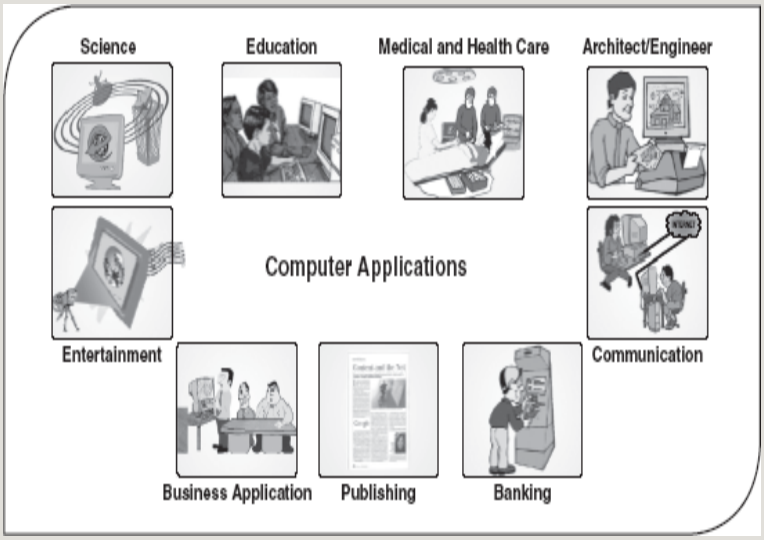
FEEDBACK

For any suggestions and comments about this book, please feel free to send an e-mail to **itlesl@rediffmail.com.**

Hope you enjoy reading this book as much as we have enjoyed writing it.

Chapter 1

Computer Basics



This chapter lays a foundation for one of the most influential forces available in modern times, the computer. A computer is an electronic device, operating under the control of instructions, which tells the machine what to do. It is capable of accepting data (input), processing data arithmetically and logically, producing output from the processing, and storing the results for future use. The chapter begins with the characteristics, evolution, and various generations of computers. The discussion also explores the classification of computers and their features. The chapter concludes with an overview on basic computer units and computer applications.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The characteristics of computers that make them an essential part of every technology
- The evolution of computers, from refining of abacus to supercomputers
- The advancement in technology that has changed the way the computers operate, leading to powerful, efficient and cheaper computers
- The classification of computers into micro, mini, mainframe and supercomputers
- The computer system, which includes components such as the Central Processing Unit (CPU) and I/O devices
- The application of computers in various fields, which increases efficiency, thus, resulting in proper utilization of resources

1.1 INTRODUCTION

In the beginning of the civilization, people used fingers and pebbles for computing purposes. In fact, the word *digitus* in Latin actually means *finger* and *calculus* means *pebble*. This gives a clue into the origin of early computing concepts. With the development of civilization, the computing needs also grew. The need for a mechanism to perform lengthy calculations led to the invention of, first, calculator and then computers.

The term *computer* is derived from the word *compute*, which means to *calculate*. A *computer* is an electronic machine devised for performing calculations and controlling operations that can be expressed either in logical or in numerical terms. In simple words, a computer is an electronic device that performs diverse operations with the help of instructions to process the data in order to achieve desired results. Although the

application domain of a computer depends totally on human creativity and imagination, it covers a huge area of applications including education, industries, government, medicine, scientific research, law, and even music and arts.

Computers are one of the most influential forces available in modern times. Harnessing the power of computers enables relatively limited and fallible human capacities for memory, logical decision making, reaction and perfection to be extended to almost infinite levels. Millions of complex calculations can be done in a mere fraction of time; difficult decisions can be made with unerring accuracy for comparatively little cost. Computers are widely seen as instruments for future progress and as tools to achieve sustainability by way of improved access to information with the help of video-conferencing and e-mail. Indeed, computers have left such an impression on modern civilization that we call this era as the “information age”.

1.1.1 Characteristics of Computers

The human race developed computers so that it could perform intricate operations, such as calculation and data processing, or simply for entertainment. Today, much of the world's infrastructure runs on computers and it has profoundly changed our lives, mostly for the better. Let us discuss some of the characteristics of computers, which make them an essential part of every emerging technology and such a desirable tool in human development.

- **Speed:** The computers process data at an extremely fast rate, at millions or billions of instructions per second. A computer can perform a huge task in a few seconds that otherwise a normal human being may take days or even years to complete. The speed of a computer is calculated in MHz (megahertz), that is, one million instructions per second. At present, a powerful computer can perform billions of operations in just one second.
- **Accuracy:** Besides the efficiency, the computers are also very accurate. The level of accuracy depends on the instructions and the

type of machines being used. Since the computer is capable of doing only what it is instructed to do, faulty instructions for data processing may lead to faulty results. This is known as *Garbage In Garbage Out* (*GIGO*).

- **Diligence:** Computer, being a machine, does not suffer from the human traits of tiredness and lack of concentration. If four million calculations have to be performed, then the computer will perform the last four-millionth calculation with the same accuracy and speed as the first calculation.
- **Reliability:** Generally, reliability is the measurement of the performance of a computer, which is measured against some predetermined standard for operation without any failure. The major reason behind the reliability of the computers is that, at hardware level, it does not require any human intervention between its processing operations. Moreover, computers have built-in diagnostic capabilities, which help in the continuous monitoring of the system.
- **Storage Capability:** Computers can store large amounts of data and can recall the required information almost instantaneously. The main memory of the computer is relatively small and it can hold only a certain amount of data; therefore, the data are stored on secondary storage devices such as magnetic tape or disks. Small sections of data can be accessed very quickly from these storage devices and brought into the main memory, as and when required, for processing.
- **Versatility:** Computers are quite versatile in nature. It can perform multiple tasks simultaneously with equal ease. For example, at one moment it can be used to draft a letter, another moment it can be used to play music and in between, one can print a document as well. All this work is possible by changing the program (computer instructions).
- **Resource Sharing:** In the initial stages of development, computers used to be isolated machines. With the tremendous growth in computer technologies, computers today have the capability to connect with each other. This has made the sharing of costly resources like printers possible. Apart from device sharing, data and information can also be shared among groups of computers, thus creating a large information and knowledge base.

Although processing has become less tedious with the development of computers, it is still a time-consuming and expensive job. Sometimes, a program works properly for some period and then suddenly produces an error. This happens because of a rare combination of events or due to an error in the instruction provided by the user. Therefore, computer parts require regular checking and maintenance in order to give correct results. Furthermore, computers need to be installed in a dust-free place. Generally, some parts of computers get heated up due to heavy processing. Therefore, the ambient temperature of the computer system should be maintained.

THINGS TO REMEMBER

Limitations of a Computer

- A computer can only perform what it is programmed to do.
- The computer needs well-defined instructions to perform any operation. Hence, computers are unable to give any conclusion without going through intermediate steps.
- A computer's use is limited in areas where qualitative considerations are important. For instance, it can make plans based on situations and information, but it cannot foresee whether they will succeed.

1.2 EVOLUTION OF COMPUTERS

The need for a device to do calculations along with the growth in commerce and other human activities explains the evolution of computers. Having the right tool to perform these tasks has always been important for human beings. In their quest to develop efficient computing devices, humankind developed many apparatuses. However, many centuries elapsed before technology was adequately advanced to develop computers.

In the beginning, when the task was simply counting or adding, people used either their fingers or pebbles along lines in the sand. In order to conveniently have the sand and pebbles all the time, people in Asia Minor built a counting device called *abacus*. This device allowed users to do calculations using a system of sliding beads arranged on a rack. The abacus was simple to operate and was used worldwide for centuries. In fact, it is still used in many countries (Figure 1.1).

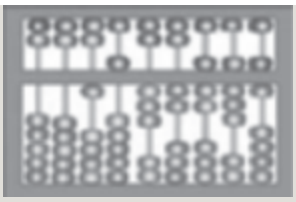


Figure 1.1 Abacus

With the passage of time, humankind invented many computing devices, such as Napier bones and slide rule. It took many centuries, however, for the next significant advancement in computing devices. In 1642, a French mathematician, Blaise Pascal, invented the first functional automatic calculator. This brass rectangular box, also called a *Pascaline*, used eight movable dials to add numbers up to eight figures long (Figure 1.2).

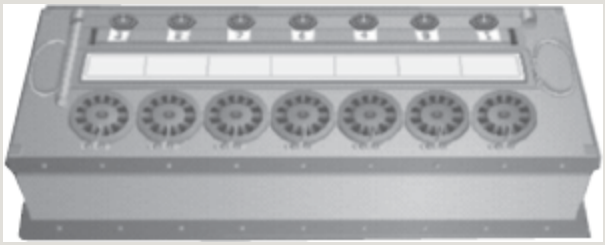


Figure 1.2 Pascaline

In 1694, a German mathematician, Gottfried Wilhem von Leibniz, extended Pascal's design to perform multiplication, division and to find square root. This machine is known as the *Stepped Reckoner*. It was the first mass-produced calculating device, which was

designed to perform multiplication by repeated addition. Like its predecessor, Leibniz's mechanical multiplier worked by a system of gears and dials. The only problem with this device was that it lacked mechanical precision in its construction and was not very reliable.

FACT FILE

What's in a Name?

In 8196, **Hollerith** found the Tabulating Machine Company, which was later named **IBM** (**International Business Machines**). IBM developed numerous mainframes and operating systems, many of which are still in use today. For example, IBM co-developed OS/2 with Microsoft, which laid the foundation for Windows operating systems.

The real beginning of computers as we know them today, however, lay with an English mathematics professor, Charles Babbage. In 1822, he proposed a machine to perform differential equations, called a *Difference Engine*. Powered by steam and as large as a locomotive, the machine would have a stored program and could perform calculations and print the results automatically. However, Babbage never quite made a fully functional difference engine and in 1833 he quitted working on this machine to concentrate on the *Analytical Engine*. The basic design of this engine included input devices in the form of perforated cards containing operating instructions and a “store” for memory of 1,000 numbers of up to 50 decimal digits long. It also contained a control unit to allow processing instructions in any sequence and output devices to produce printed results. Babbage borrowed the idea of punch cards to encode the

machine's instructions from Joseph-Marie Jacquard's loom. Although the analytical engine was never constructed, it outlined the basic elements of a modern computer.

In 1889, Herman Hollerith, who worked for the US Census Bureau, also applied Jacquard's loom concept to computing. Unlike Babbage's idea of using perforated cards to instruct the machine, Hollerith's method used cards to store the data, which he fed into a machine that compiled the results mechanically (Figure 1.3).

Figure 1.3 Hollerith's Tabulator

The start of World War II produced a substantial need for computer capacity, especially for military purposes. One early success was the *Mark I*, which was built as a partnership between Harvard Aiken and IBM in 1944. This electronic calculating machine used relays and electromagnetic components to replace mechanical components. In 1946, John Eckert and John Mauchly of the Moore School of Engineering at the University of Pennsylvania developed the *Electronic Numerical Integrator and Calculator* (*ENIAC*). This computer used

electronic vacuum tubes to make the internal parts of the computer. It embodied almost all the components and concepts of today's high-speed, electronic computers. Later on, Eckert and Mauchly also proposed the development of the *Electronic Discrete Variable Automatic Computer* (*EDVAC*). It was the first electronic computer to use the stored program concept introduced by John Von Neumann. It also had the capability of conditional transfer of control, that is, the computer could stop any time and then resume operations. In 1949, at the Cambridge University, a team headed by Maurice Wilkes developed the *Electronic Delay Storage Automatic Calculator* (*EDSAC*), which was also based on John Von Neumann's stored program concept. This machine used mercury delay lines for memory and vacuum tubes for logic. The Eckert–Mauchly Corporation manufactured the *Universal Automatic Computer* (*UNIVAC*) in 1951 and its implementation marked the real beginning of the computer era.

In the 1960s, efforts to design and develop the fastest possible computer with the greatest capacity reached a turning point with the *Livermore Advanced Research Computer* (*LARC*), which had access time of less than 1 μs (pronounced as *microsecond*) and the total capacity of 100,000,000 words. During this period, the major computer manufacturers began to offer a range of capabilities and prices, as well as accessories such as card feeders, page printers and cathode ray tube displays. During the 1970s, the trend shifted towards a larger range of applications for cheaper computer systems. During this period, many business organizations adopted computers for their offices. The vacuum deposition of transistors became the norm and entire computer assemblies became available on tiny “chips”.

In the 1980s, Very Large Scale Integration (VLSI) design, in which hundreds of thousands of transistors were

placed on a single chip, became increasingly common. The “shrinking” trend continued with the introduction of personal computers (PCs), which are programmable machines small enough and inexpensive enough to be purchased and used by individuals. Microprocessors equipped with the read-only memory (ROM), which stores constantly used and unchanging programs, performed an increased number of functions. By the late 1980s, some PCs were run by microprocessors that were capable of handling 32 bits of data at a time and processing about 4,000,000 instructions per second. By the 1990s, PCs became part of everyday life. This transformation was the result of the invention of the *microprocessor*, a processor on a single *integrated circuit* (*IC*) chip. The trend continued leading to the development of smaller and smaller microprocessors with a proportionate increase in processing powers. The computer technology continues to experience huge growth. Computer networking, electronic mail and electronic publishing are just a few applications that have grown in recent years. Advances in technologies continue to produce cheaper and more powerful computers, offering the promise that in the near future, computers or terminals will reside in most, if not all, homes, offices and schools.

1.3 GENERATIONS OF COMPUTERS

The history of computer development is often discussed with reference to the different generations of computing devices. In computer terminology, the word *generation* is described as a stage of technological development or innovation. A major technological development that fundamentally changed the way computers operate, resulting in increasingly smaller, cheaper, more powerful, and more efficient and reliable devices, characterizes each generation of computers. According to the type of “processor” installed in a machine, there are five generations of computers.

1.3.1 First Generation (1940 to 1956): Vacuum Tubes

First-generation computers were vacuum tubes/thermionic valve-based machines. These computers used vacuum tubes for circuitry and magnetic drums for memory. *A magnetic drum* is a metal cylinder coated with magnetic iron oxide material on which data and programs can be stored. The input was based on punched cards and paper tape, and the output was in the form of printouts (Figure 1.4).

Figure 1.4 Vacuum Tube

First-generation computers relied on *binary-coded language* also called *machine language* (language of 0s and 1s) to perform operations and were able to solve only one problem at a time. Each machine was fed with different binary codes and hence, were difficult to program. This resulted in lack of versatility and speed. In addition, to run on different types of computers, instructions must be rewritten or recompiled.

Examples: ENIAC, EDVAC and UNIVAC.

Characteristics of First-generation Computers

- These computers were based on vacuum tube technology.
- These were the fastest computing devices of their times (computation time was in milliseconds).
- These computers were very large and required a lot of space for installation.
- Since thousands of vacuum tubes were used, they generated a large amount of heat. Therefore, air conditioning was essential.
- These were non-portable and very slow equipments.
- They lacked in versatility and speed.
- They were very expensive to operate and used a large amount of electricity.
- These machines were unreliable and prone to frequent hardware failures. Hence, constant maintenance was required.
- Since machine language was used, these computers were difficult to program and use.
- Each individual component had to be assembled manually. Hence, commercial appeal of these computers was poor.

1.3.2 Second Generation (1956 to 1963): Transistors

Second-generation computers used transistors, which were superior to vacuum tubes. A *transistor* is made up of semiconductor material like germanium and silicon. It usually has three leads (see Figure 1.5) and performs electrical functions such as voltage, current or power amplification with low power requirements. Since a transistor is a small device, the physical size of computers was greatly reduced. Computers became smaller, faster, cheaper, energy efficient and more reliable than their predecessors. In second-generation computers, magnetic cores were used as the primary memory and magnetic disks as the secondary storage devices. However, they still relied on punched cards for the input and printouts for the output.

Figure 1.5 Transistor

One of the major developments of this generation includes the progress from machine language to assembly language. Assembly language uses mnemonics (abbreviations) for instructions rather than numbers, for example, ADD for addition and MULT for multiplication. As a result, programming became less cumbersome. Early high-level programming languages such as COBOL and FORTRAN also came into existence in this period.

Examples: PDP-8, IBM 1401 and IBM 7090.

Characteristics of Second-generation Computers

- These machines were based on transistor technology.
- These were smaller as compared to the first-generation computers.
- The computational time of these computers was reduced to microseconds from milliseconds.
- These were more reliable and less prone to hardware failure. Hence, they required less frequent maintenance.
- These were more portable and generated less amount of heat.
- Assembly language was used to program computers. Hence, programming became more time-efficient and less cumbersome.
- Second-generation computers still required air conditioning.
- Manual assembly of individual components into a functioning unit was still required.

1.3.3 Third Generation (1964 to Early 1970s): Integrated Circuits

The development of the integrated circuit, also called an IC, was the trait of the third-generation computers. An *IC* consists of a single chip (usually silicon) with many components such as transistors and resistors fabricated on it. ICs replaced several individually wired transistors. This development made computers smaller in size, reliable and efficient (Figure 1.6).

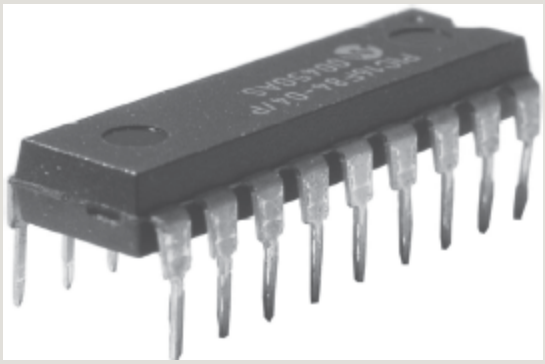


Figure 1.6 Integrated Circuit

Instead of punched cards and printouts, users interacted with third-generation computers through keyboards and monitors, and interfaced with the operating system. This allowed the device to run many different applications simultaneously with a central program that monitored the memory. For the first time, computers became accessible to mass audience because they were smaller and cheaper than their predecessors.

Examples: NCR 395 and B6500.

Characteristics of Third-generation Computers

- These computers were based on IC technology.
- These were able to reduce the computational time from microseconds to nanoseconds.
- These were easily portable and more reliable than the second-generation computers.
- These devices consumed less power and generated less heat. In some cases, air conditioning was still required.
- The size of these computers was smaller as compared to previous-generation computers.
- Since hardware rarely failed, the maintenance cost was quite low.
- Extensive use of high-level languages became possible.
- Manual assembling of individual components was not required, so it reduced the large requirement of labour and cost. However, highly sophisticated technologies were required for the manufacturing of IC chips.
- Commercial production became easier and cheaper.

1.3.4 Fourth Generation (Early 1970s to Till Date): Microprocessors

The fourth generation is an extension of third generation technology. Although, the technology of this generation is still based on the IC, these have been made readily available to us because of the development of the microprocessor (circuits containing millions of transistors). The Intel 4004 chip, which was developed in 1971, took the IC one step further by locating all the components of a computer (CPU, memory and I/O controls) on a minuscule chip. A microprocessor is built on to a single piece of silicon, known as *chip*. It is about 0.5 cm along one side and no more than 0.05 cm thick.

The fourth-generation computers led to an era of *Large Scale Integration* (*LSI*) and *VLSI* technology. LSI technology allowed thousands of transistors to be constructed on one small slice of silicon material, whereas VLSI squeezed hundreds of thousands of components on to a single chip. Ultra Large Scale Integration (ULSI) increased that number to the millions. This way computers became smaller and cheaper than ever before (Figure 1.7).

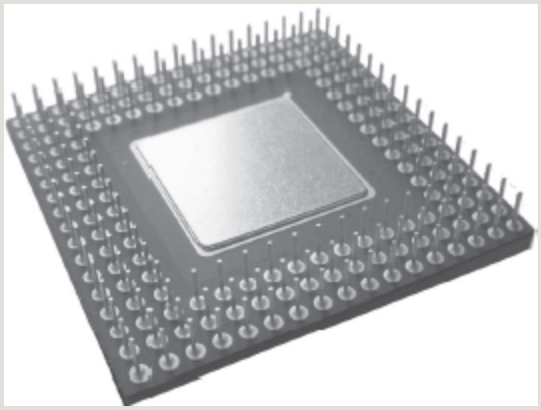


Figure 1.7 Microprocessor

The fourth-generation computers became more powerful, compact, reliable and affordable. As a result, it gave rise to the PC revolution. During this period, magnetic core memories were substituted by semiconductor memories, which resulted in faster random access main memories. Moreover, secondary memories such as hard disks became economical, smaller and bigger in capacity. The other significant development of this era was that these computers could be linked together to form networks, which eventually led to the development of the Internet. This generation also saw the development of the Graphical User Interfaces (GUIs), mouse and handheld devices. Despite many advantages, this generation required complex and sophisticated technology for the manufacturing of the CPU and the other components.

Examples: Apple II, Altair 8800 and CRAY-1.

Characteristics of Fourth-generation Computers

- These computers are microprocessor-based systems.

- These are very small in size.
- These are the cheapest among all the other generation computers.
- These are portable and quite reliable.
- These machines generate negligible amount of heat, hence do not require air conditioning.
- Hardware failure is negligible so minimum maintenance is required.
- The production cost is very low.
- The GUI and pointing devices enabled users to learn to use the computer quickly.
- Interconnection of computers led to better communication and resource sharing.

1.3.5 Fifth Generation (Present and Beyond): Artificial Intelligence

The dream of creating a human-like computer that would be capable of reasoning and reaching at a decision through a series of "what-if-then" analyses has existed since the beginning of computer technology. Such a computer would learn from its mistakes and possess the skill of experts. These are the objectives for creating the fifth generation of computers. The starting point for the fifth generation of computers had been set in the early 1990s. The process of developing fifth-generation computers is still in the development stage. However, the expert system concept is already in use. The *expert system* is defined as a computer system that attempts to mimic the thought process and reasoning of experts in specific areas. Three characteristics can be identified with the fifth-generation computers. These are:

- **Mega Chips:** Fifth-generation computers will use *Super Large Scale Integrated* (*SLSI*) chips, which will result in the production of microprocessors having millions of electronic components on a single chip. In order to store instructions and information, fifth-generation computers require a great amount of storage capacity. Mega chips may enable the computer to approximate the memory capacity of the human mind.
- **Parallel Processing:** Computers with one processor access and execute only one instruction at a time. This is called serial processing. However, fifth-generation computers will use multiple processors and perform parallel processing, thereby accessing several instructions at once and working on them at the same time.
- **Artificial Intelligence (AI):** It refers to a series of related technologies that try to simulate and reproduce human behavior,

including thinking, speaking and reasoning. AI comprises a group of related technologies: expert systems (ES), natural language processing (NLP), speech recognition, vision recognition and robotics.

1.4 CLASSIFICATION OF COMPUTERS

These days, computers are available in many sizes and types. Some computers can fit in the palm of the hand, while some can occupy the entire room. Computers also differ based on their data-processing abilities. Based on the physical size, performance and application areas, we can generally divide computers into four major categories: *micro, mini, mainframe* and *supercomputers* (Figure 1.8).

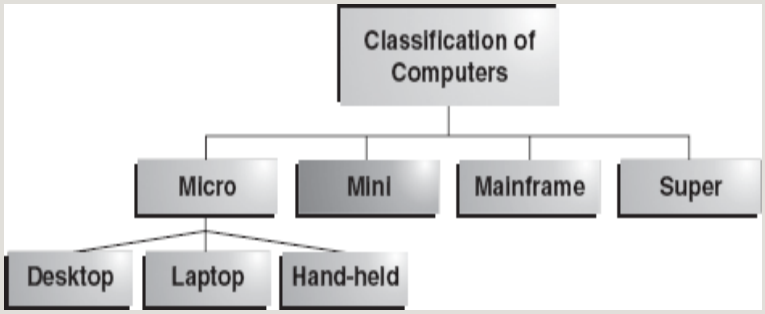


Figure 1.8 Classification of Computers

1.4.1 Microcomputers

A *microcomputer* is a small, low-cost digital computer, which usually consists of a microprocessor, a storage unit, an input channel and an output channel, all of which may be on one chip inserted into one or several PC boards. The addition of power supply and connecting cables, appropriate peripherals (keyboard, monitor, printer, disk drives and others), an operating system and other software programs can provide a complete microcomputer system. The micro-computer is generally the smallest of the computer family. Originally, these

were designed for individual users only, but nowadays they have become powerful tools for many businesses that, when networked together, can serve more than one user. IBM-PC Pentium 100, IBM-PC Pentium 200 and Apple Macintosh are some of the examples of microcomputers. Microcomputers include *desktop, laptop* and *hand-held* models such as Personal Digital Assistants (PDAs).

Desktop Computer: The *desktop computer*, also known as the *PC*, is principally intended for stand-alone use by an individual. These are the most-common type of microcomputers. These microcomputers typically consist of a system unit, a display monitor, a keyboard, an internal hard disk storage and other peripheral devices. The main reason behind the importance of the PCs is that they are not very expensive for the individuals or the small businesses. Some of the major PC manufacturers are APPLE, IBM, Dell and Hewlett-Packard (Figure 1.9).

Figure 1.9 Desktop Computer

Laptop: A *laptop* is a portable computer that a user can carry around. Since the laptop resembles a notebook, it is also known as the *notebook computer*. Laptops are small computers enclosing all the basic features of a normal desktop computer. The biggest advantage of laptops is that they are lightweight and one can use them anywhere and at anytime, especially when one is travelling. Moreover, they do not need any external power supply as a rechargeable battery is completely self-contained in them. However, they are expensive as compared to desktop computers (Figure 1.10).

Figure 1.10 Laptop

Hand-held Computers: A *hand-held* computer such as a PDA is a portable computer that can conveniently be stored in a pocket (of sufficient size) and used while the user is holding it. PDAs are essentially small portable computers and are slightly bigger than the common calculators. A PDA user generally uses a *pen* or *electronic stylus*, instead of a keyboard for input. As shown in (Figure 1.11), the monitor is very small and is the only

apparent form of output. Since these computers can be easily fitted on the top of the palm, they are also known as *palmtop* computers. Handheld computers usually have no disk drive; rather, they use small cards to store programs and data. However, they can be connected to a printer or a disk drive to generate output or store data. They have limited memory and are less powerful as compared to desktop computers. Some examples of hand-held computers are Apple Newton, Casio Cassiopeia and Franklin eBookMan.

Figure 1.11 Personal Digital Assistant

1.4.2 Minicomputers

In the early 1960s, Digital Equipment Corporation (DEC) started shipping its PDP series computer, which the press described and referred to as minicomputers. A *minicomputer* is a small digital computer, which normally is able to process and store less data than a mainframe but more than a microcomputer, while doing so less rapidly than a mainframe but more rapidly than a microcomputer. It is about the size of a two-drawer filing cabinet. Generally, these computers are used as desktop

devices that are often connected to a mainframe in order to perform the auxiliary operations (Figure 1.12).

Figure 1.12 Minicomputer

A minicomputer (sometimes called a *mid-range* computer) is designed to meet the computing needs of several people simultaneously in a small- to medium-sized business environment. It is capable of supporting from four to about 200 simultaneous users. It serves as a centralized storehouse for a cluster of workstations or as a network server. Minicomputers are usually multi-user systems so these are used in interactive applications in industries, research organizations, colleges and universities. They are also used for real-time controls and engineering design work. Some of the widely used minicomputers are PDP 11, IBM (8000 series) and VAX 7500.

1.4.3 Mainframes

A *mainframe* is an ultra-high performance computer made for high-volume, processor-intensive computing. It consists of a high-end computer processor, with related peripheral devices, capable of supporting large volumes of data processing, high-performance online transaction processing, and extensive data storage and retrieval. Normally, it is able to process and store more data than a minicomputer and far more than a microcomputer. Moreover, it is designed to perform at a faster rate than a minicomputer and at even more faster rate than a microcomputer. Mainframes are the second largest (in capability and size) of the computer family, the largest being the supercomputers. However, mainframes can usually execute many programs simultaneously at a high speed, whereas supercomputers are designed for a single process (Figure 1.13).

Figure 1.13 Mainframe

The mainframe allows its users to maintain a large amount of data storage at a centralized location and to access and process these data from different computers

located at different locations. It is typically used by large businesses and for scientific purposes. Some examples of the mainframe are IBM's ES000, VAX 8000 and CDC 6600.

1.4.4 Supercomputers

Supercomputers are the special-purpose machines, which are especially designed to maximize the numbers of floating point operations per second (FLOPS). Any computer below one gigaflop per second is not considered a supercomputer. A supercomputer has the highest processing speed at a given time for solving scientific and engineering problems. Essentially, it contains a number of CPUs that operate in parallel to make it faster. Its processing speed lies in the range 400–10,000 MFLOPS (millions of floating point operations per second). Due to this feature, supercomputers help in many applications including information retrieval and computer-aided designing (Figure 1.14).

Figure 1.14 Supercomputer

A supercomputer can process a great deal of data and make extensive calculations very quickly. It can resolve complex mathematical equations in a few hours, which would have taken many years when performed using a paper and pencil or using a hand calculator. It is the fastest, costliest and most powerful computer available today. Typically, supercomputers are used to solve multivariant mathematical problems of existent physical processes, such as aerodynamics, metrology and plasma physics. They are also required by the military strategists to simulate defence scenarios. Cinematic specialists use them to produce sophisticated movie animations. Scientists build complex models and simulate them in a supercomputer. However, a supercomputer has limited broad-spectrum use because of its price and limited market. The largest commercial uses of supercomputers are in the entertainment/advertising industry. CRAY-3, Cyber 205 and PARAM are some well-known supercomputers.

FACT FILE

India's Super Achievement

In 2003, India developed the PARAM Padma supercomputer, which marks an important step towards high-performance computing. The PARAM Padma computer was developed by India's Center for Development of Advanced Computer (C-DAC) and promises processing speeds of up to 1 teraflop per second (1 trillion processes per second).

A computer can be viewed as a system, which consists of a number of interrelated components that work together with the aim of converting data into information. In a computer system, processing is carried out electronically, usually with little or no intervention from the user.

The general perception of people regarding the computer is that it is an “intelligent thinking machine”. However, this is not true. Every computer needs to be instructed exactly what to do and how to do. The instructions given to computers are called *programs*. Without programs, *computers* would be useless. The physical parts that make up a computer (the CPU, input, output and storage unit) are known as *hardware*. Any hardware device connected to the computer or any part of the computer outside the CPU and working memory is known as a *peripheral*. Some examples of peripherals are keyboards, mouse and monitors.

1.5.1 Components of a Computer System

There are several computer systems in the market with a wide variety of makes, models and peripherals. In general, a computer system comprises the following components:

- **CPU:** This unit performs processing of instructions and data inside the computer.
- **Input Unit:** This unit accepts instructions and data.
- **Output Unit:** This unit communicates the results to the user.
- **Storage Unit:** This unit stores temporary and final results (Figure 1.15).

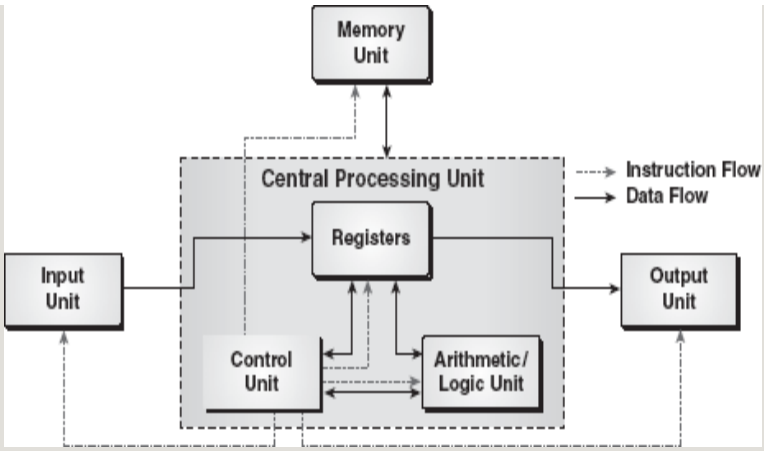


Figure 1.15 Components of a Computer System

Central Processing Unit: The CPU, also known as a *processor*, is the brain of the computer system that processes data (input) and converts it into meaningful information (output). It is referred to as the *administrative section* of the computer system that interprets the data and instructions, coordinates the operations, and supervises the instructions. The CPU works with data in discrete form, that is, either 1 or 0. It counts, lists, compares and rearranges the binary digits of data in accordance with the detailed program instructions stored within the memory. Eventually, the results of these operations are translated into characters, numbers and symbols that can be understood by the user. The CPU itself has three parts:

- **Arithmetic Logic Unit (ALU):** This unit performs the arithmetic (add, subtract) and logical operations (and, or) on the data made available to it. Whenever an arithmetic or logical operation is to be performed, the required data are transferred from the memory unit to the ALU, the operation is performed and the result is returned to the memory unit. Before the completion of the processing, data may need to be transferred back and forth several times between these two sections.
- **Control Unit:** This unit checks the correctness of the sequence of operations. It fetches the program instructions from the memory unit, interprets them and ensures correct execution of the program. It also

controls the I/O devices and directs the overall functioning of the other units of the computer.

- **Registers:** These are the special-purpose, high-speed temporary memory units that can hold varied information such as data, instructions, addresses and intermediate results of calculations. Essentially, they hold the information that the CPU is currently working on. The registers can be considered as the CPU's working memory, an additional storage location that provides the advantage of speed.

Note: *The circuits necessary to create a CPU for a PC are fabricated on a microprocessor*.

Input, Output and Storage Unit: The user must enter instructions and data into the computer system before any operation can be performed on the given data. Similarly, after processing the data, the information must go out from the computer system to the user. For this, every computer system incorporates the I/O unit that serves as a communication medium between the computer system and the user.

An *input unit* accepts instructions and data from the user with the help of input devices such as keyboard, mouse, light pen, etc. Since the data and instructions entered through different input devices will be in different form, the input unit converts them into the form that the computer can understand. After this, the input unit supplies the converted instructions and data to the computer for further processing.

The *output unit* performs just opposite to that of input unit. It accepts the output (which is in machine-coded form) produced by the computer, converts them into the user-understandable form and supplies the converted results to the user with the help of an output device such as printer, monitor and plotter.

Besides, a computer system incorporates a *storage unit* to store the input entered through the input unit before processing starts and to store the results produced by the computer before supplying them to the output unit. The storage unit of a computer comprises two types of

memory/storage: *primary* and *secondary*. The primary memory, also called the *main memory*, is the part of a computer that holds the instructions and data currently being processed by the CPU, the intermediate results produced during the course of calculations and the recently processed data. While the instructions and data remain in the main memory, the CPU can access them directly and quickly. However, the primary memory is quite expensive and has a limited storage capacity.

Due to the limited size of the primary memory, a computer employs the secondary memory, which is extensively used for storing data and instructions. It supplies the stored information to the other units of the computer as and when required. It is less expensive and has higher storage capacity than the primary memory. Some commonly used secondary storage devices are floppy disks, hard disks and tape drives (Figure 1.16).

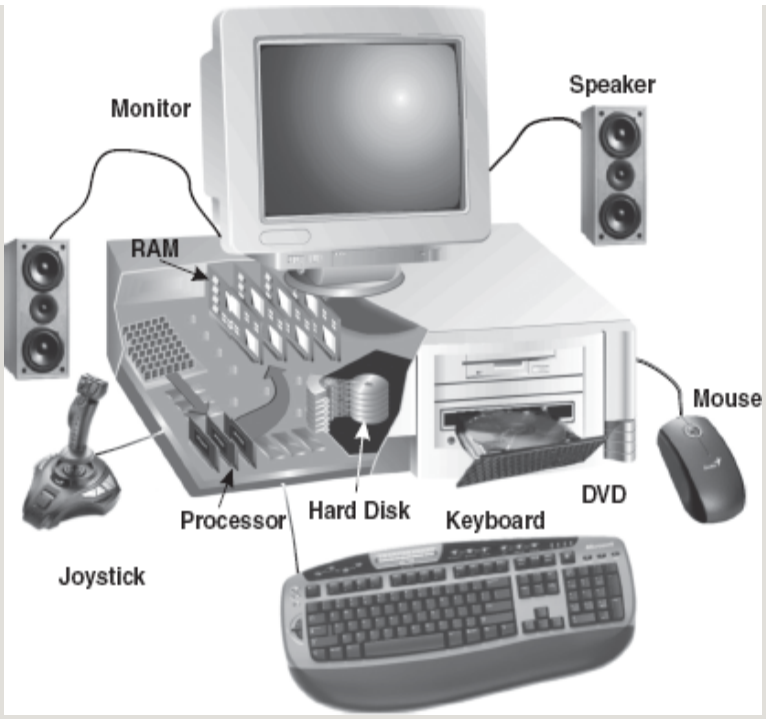


Figure 1.16 Typical Hardware and Peripherals in a Computer System

1.5.2 How Does a Computer Work?

A computer performs three basic steps to complete any task: *input, processing* and *output*. A task is assigned to a computer in a set of step-by-step instructions, which is known as a *program*. These instructions tell the computer what to do with the input in order to produce the required output. A computer functions in the following manner:

Step 1	The computer accepts the *input*. The computer input is whatever entered or fed into a computer system. The input can be supplied by the user (such as by using a keyboard) or by another computer or device (such as a diskette or CD-ROM). Some examples of input include the words and symbols in a document, numbers for a calculation, instructions for completing a process, and so on.

Step 2	The computer *processes* the data. During this stage, the computer follows the instructions using the data that have been input. Examples of processing include calculations, sorting lists of words or numbers and modifying documents according to user instructions.
Step 3	The computer produces *output*. Computer output is the information that has been produced by a computer. Some examples of computer output include reports, documents and graphs. Output can be in several formats, such as printouts, or displayed on the screen (Figure 1.17).

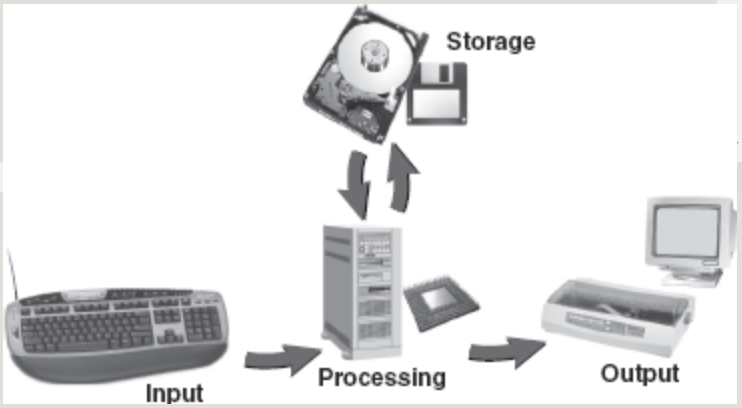


Figure 1.17 Basic Computer Operations

1.6 APPLICATIONS OF COMPUTERS

In the last few decades, computer technology has revolutionized the businesses and other aspects of human life all over the world. Practically, every company, large or small, is now directly or indirectly dependent on computers for data processing. Computer systems also help in the efficient operation of railway and airway reservation, hospital records, accounts, electronic banking and so on. Computers not only save time, but also save paper work. Some of the areas where computers are being used are listed below.

- **Science:** Scientists have been using computers to develop theories, to analyse and to test the data. The fast speed and the accuracy of the computer allow different scientific analyses to be carried out. They can be used to generate detailed studies of how earthquakes affect buildings or pollution affects weather pattern. Satellite-based applications would not have been possible without the use of computers. It would also not be possible to get the information of our solar system and the cosmos without computers.
- **Education:** Computers have also revolutionized the whole process of education. Currently, the classrooms, libraries and museums are utilizing computers to make the education much more interesting. Unlike recorded television shows, computer-aided education (CAE) and computer-based training (CBT) packages are making learning much more interactive.
- **Medicine and Healthcare:** There has been an increasing use of computers in the field of medicine. Now, doctors are using computers right from diagnosing the illness to monitoring a patient's status during complex surgery. By using automated imaging techniques, doctors are able to look inside a person's body and can study each organ in detail (such as CAT scans or MRI scans), which was not possible few years ago. There are several examples of special-purpose computers that can operate within the human body such as a cochlear implant, a special kind of hearing aid that makes it possible for deaf people to hear.
- **Engineering/Architecture/Manufacturing:** The architects and engineers are extensively using computers in designing and drawings. Computers can create objects that can be viewed from all the three dimensions. By using techniques like virtual reality, architects can explore houses that have been designed but not built. The manufacturing factories are using computerized robotic arms in order to perform hazardous jobs. Besides, computer-aided manufacturing (CAM) can be used in designing the product, ordering the parts and planning production. Thus, computers help in coordinating the entire manufacturing process.
- **Entertainment:** Computers are finding greater use in the entertainment industry. They are used to control the images and sounds. The special effects, which mesmerize the audience, would not have been possible without the computers. In addition, computerized animation and colourful graphics have modernized the film industry.
- **Communication:** E-mail or electronic mail is one of the communication media in which computers are used. Through an e-mail, messages and reports are passed from one person to one or more persons with the aid of computers and telephone lines. The advantage of this service is that while transferring the messages it saves time, avoids wastage of paper, and so on. Moreover, the person who is receiving the messages can read the messages whenever he is free and can save it, reply it, forward it or delete it from the computer.
- **Business Application:** This is one of the important uses of the computer. Initially, computers were used for batch processing jobs, where one does not require the immediate response from the computer. Currently, computers are mainly used for real-time applications (like at the sales counter) that require immediate response from the computer. There are various concerns for which computers are used such as in business forecasting, to prepare pay bills and personal records, in banking operations and data storage, in various types of life insurance business, and as an aid to management.

Businesses are also using the networking of computers, where a number of computers are connected together to share the data and the information. Use of an e-mail and the Internet has changed the ways of doing business.

- **Publishing:** Computers have created a field known as *Desktop Publishing* (*DTP*). In DTP, with the help of a computer and a laser printer one can perform the publishing job all by oneself. Many of the tasks requiring long manual hours, such as making a table of contents and an index, can be automatically performed using the computers and DTP software.
- **Banking:** In the field of banking and finance, computers are extensively used. People can use the Automated Teller Machine (ATM) services 24 hours a day in order to deposit and withdraw cash. When the different branches of the bank are connected through the computer networks, the inter-branch transactions, such as drawing cheques and making drafts, can be performed by the computers without any delay (Figure 1.18).

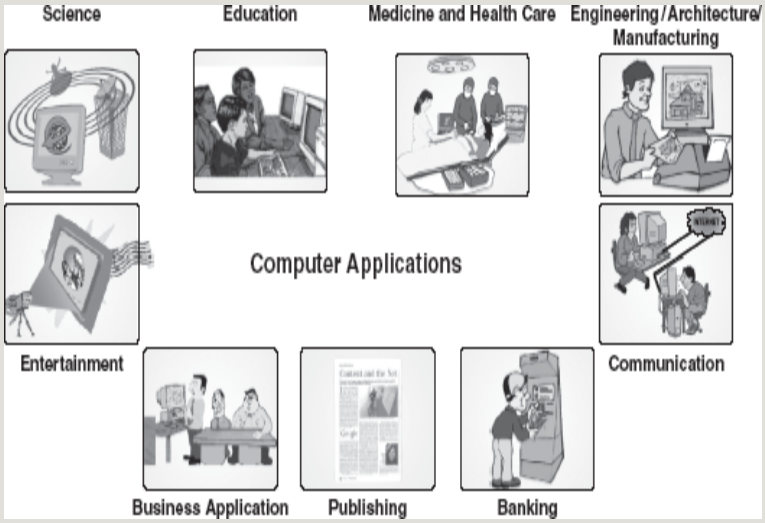


Figure 1.18 Application Areas of Computer

LET US SUMMARIZE

1. *Computer* is an electronic device that performs diverse operations with the help of instructions to process the data in order to achieve desired results. Speed, accuracy, reliability, versatility, diligence, storage capability and resource sharing characterize the computers.
2. Many devices, which humans developed for their computing requirements, preceded computers. Some of those devices were Sand Tables, Abacus, Napier Bones, Slide Rule, Pascaline, Stepped Reckoner, Difference Engine, Analytical Engine and Hollerith's Tabulator.

3. Computer development is divided into five main generations. With every generation, computer technology has fundamentally changed, resulting in an increasingly smaller, cheaper, more powerful, more efficient and reliable devices.
4. *First-generation computers* were vacuum tube based machines. These computers were very large, required a lot of space for installation, generated a large amount of heat were non-portable and have very slow equipments. In addition, these machines were unreliable and prone to frequent hardware failures.
5. *Second-generation computers* used transistors in place of vacuum tubes. Since a transistor is a small device, the physical size of computers was greatly reduced. Computers became smaller, faster, cheaper, energy-efficient and more reliable than their predecessors.
6. *Third-generation computers* were IC-based machines. The IC replaced several individually wired transistors, making computers smaller in size, reliable and efficient.
7. *Fourth-generation computers* use microprocessors (circuits containing millions of transistors) as their basic processing device. These computers are the most powerful, compact, reliable and affordable as compared to their predecessors.
8. *Fifth-generation computers* are still in the development stage. These computers will use megachips, which will result in the production of microprocessors having millions of electronic components on a single chip. They will use intelligent programming (AI) and knowledge-based problem-solving techniques.
9. A *microcomputer* is a small, low-cost digital computer, which usually consists of a microprocessor, a storage unit, and an input and output channel, all of which may be on one chip inserted into one or several PC boards. Microcomputers include desktop, laptop and hand-held models, such as PDAs.
10. A *minicomputer* is a small digital computer, which normally is able to process and store less data than a mainframe but more than a microcomputer, while doing so less rapidly than a mainframe but more rapidly than a microcomputer.
11. A *mainframe* is an ultrahigh performance computer made for high-volume, processor-intensive computing. It is capable of supporting large volumes of data processing, high-performance online transaction processing, and extensive data storage and retrieval.
12. *Supercomputers* are the special-purpose machines, which are specifically designed to maximize the number of FLOPS. Any computer below one gigaflop per second is not considered as a supercomputer.
13. A computer can be viewed as a system that comprises several units (*CPU, input unit, output unit* and *storage unit*). These individual units work together to convert data into information.
14. The *CPU* interprets, coordinates the operations and supervises the instructions. It has three parts: *ALU*, *CU* and *registers*. The ALU performs arithmetic (add, subtract) and logical operations (and, or) on the stored numbers. The CU checks the correctness of the sequence of operations, controls the I/O devices and directs the overall functioning of the other units of the computer. The registers are special-purpose, high-speed temporary memory units that can hold varied information such as data, instructions, addresses and intermediate results of calculations.

15. The *input unit* involves the receipt of data or instructions from the user, in a computer acceptable form. The computer takes in the data through input devices like keyboard, mouse, light pen, etc.
16. The *output unit* supplies the resulting data obtained from the data processing to the user. Monitors, printers and plotters are some of the examples of output devices.
17. During processing, the intermediate and the results of processing are held by the *storage unit* until the manipulation of the data is completed. When the data to be processed or the results produced by the processing are in large volumes, they are stored on various storage media like floppies, hard disks and tapes.
18. A computer performs three basic steps to complete any task: *input, processing* and *output*. A computer receives data as input, processes it, stores it and then produces output.
19. Computers have entered in every sphere of human life and found applications in various fields, such as medicine and healthcare, business, education, science, technology, government, entertainment, engineering and architecture.

EXERCISES

Fill in the Blanks

1. The basic component of first-generation computers was .....
2. The speed of a computer is calculated in .....
3. Third-generation computers were ..... based machines.
4. Keyboard is an ..... device.
5. Laptops are also known as .....
6. Computers can be classified as ....., ..., ... and ..... computers.
7. A computer performs three basic steps to complete any task, which are ....., ... and .....
8. PDA stands for .....
9. The CPU consists of ....., ... and .....
10. Physical components on which the data are stored permanently are called .....

Multiple-choice Questions

1. The development of computers can be divided ..... into generations.
    1. 3
    2. 4
    3. 5
    4. 6
2. Choose the odd one out.
    1. Microcomputer
    2. Minicomputer
    3. Supercomputer
    4. Digital computer
3. ..... is a very small computer that can be held in the palm of the hand.

1. PDA
2. PC
3. Laptop
4. Minicomputer

4. Analytical engine was developed by .....

1. Gottfried Wilhem Von Leibriz
2. Charles Babbage
3. Herman Hollerith
4. Joseph-Marie Jacquard

5. The main distinguishing feature of fifth-generation computers will be .....

1. Liberal use of microprocessors
2. Artificial Intelligence
3. Extremely low cost
4. Versatility

6. The computer that is not considered as a portable computer is .....

1. Laptop
2. PDA
3. Minicomputer
4. None of these

7. The CPU stands for .....

1. Central protection unit
2. Central processing unit
3. Central power unit
4. Central prerogative unit

8. UNIVAC is an example of .....

1. First-generation computer
2. Second-generation computer
3. Third-generation computer
4. Fourth-generation computer

9. The unit that performs the arithmetic and logical operations on the stored numbers is known as .....

1. Arithmetic logic unit
2. Control unit
3. Memory unit
4. Both (a) and (b)

10. The ..... is the “administrative” section of the computer system.

1. Input unit
2. Output unit
3. Memory unit
4. Central processing unit

State True or False

1. The ALU is responsible for performing the arithmetic and logical operations.
2. Microcomputers are more powerful than minicomputers.
3. Laptop is also known as a notebook.
4. Vacuum tubes were a part of third-generation computers.
5. EDVAC was a second-generation computer.
6. LSI and VLSI technology are part of fifth-generation computers.
7. A laptop is a portable computer.
8. Primary memory and main memory are synonyms.
9. Computer development is divided into four main generations.
10. PARAM is an example of a portable computer.

Descriptive Questions

1. Discuss the characteristics of computers.
2. What are the advantages of transistors over vacuum tubes?
3. Discuss various types of computers in detail.
4. List out various applications of computers.
5. Discuss various computer generations along with the key characteristics of the computers of each generation.
6. Discuss the basic organization of a computer system and explain the functions of various units of a computer system.

ANSWERS

Fill in the Blanks

1. Vacuum tube
2. Megahertz
3. Integrated circuits
4. Input
5. Notebook computers
6. Micro, Mini, Mainframe, Super
7. Input, Processing, Output
8. Personal digital assistant
9. ALU, CU, registers
10. Secondary storage devices

Multiple-choice Questions

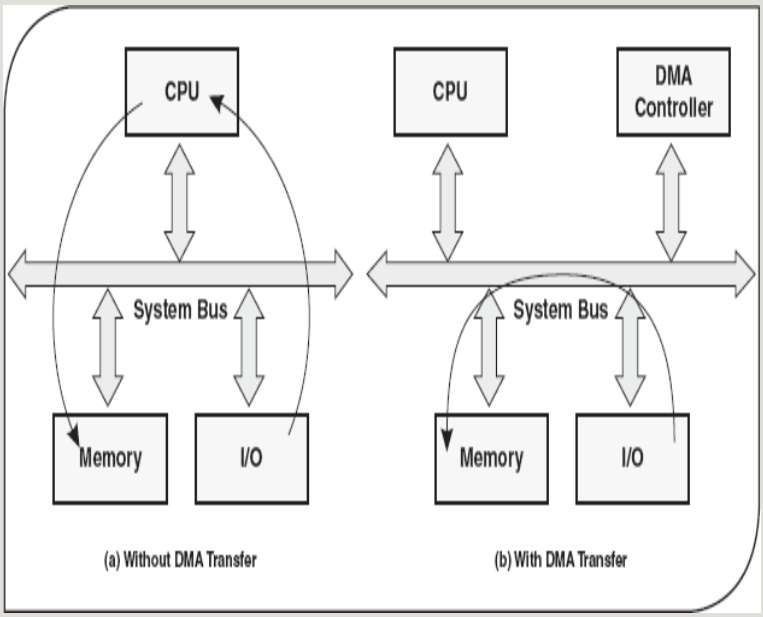
1. (c)
2. (d)
3. (a)
4. (b)
5. (b)
6. (c)
7. (b)
8. (a)
9. (a)
10. (d)

State True or False

1. True
2. False
3. True
4. False
5. False
6. False
7. True
8. True
9. False
10. False

Chapter 2

Computer Organization and Architecture



This chapter provides an insight into the basic science of computer architecture and organization. The architecture of a computer defines the manner in which the hardware components of the system work to meet functional objectives, and enabling these hardware components to work in a harmonized manner is called computer organization. The chapter begins with the structure and functioning of CPU. It then discusses the operations performed on the instruction and user data. Further, the communication among the processor, memory, and I/O devices is explained. CISC and RISC architectures are described along with their distinctions. The chapter also discusses the various hardware peripherals that can be added to a computer.

The chapter concludes with an overview on various number systems along with the coding schemes.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The central processing unit (CPU), referred to as the “brain” of the computer system, and its various components
- The units of a computer that communicate with each other to formulate a functional computer
- The instruction format, stating the addressing mode used for each operand in an instruction
- The instruction cycle, the basic function performed by the CPU, that is, the execution of a program
- The instruction set, which is a collection of operations
- The devices inside a computer that play an important role in the functioning of a properly designed and well-built computer
- The number system, which defines a set of values used to represent a quantity
- How to represent numeric, alphabetic and special characters in a computer-understandable format

2.1 INTRODUCTION

Just as a tall building has different levels of detail, including the number of storeys, the size of rooms, door and window placement, etc., each computer has a visible structure, which is referred to as its *architecture*. One can look at a computer's architecture at similar levels of hardware elements, which in turn depends on the type of the computer (personal computer, supercomputer, etc.) required. Therefore, when we talk about architecture in terms of a computer, it is defined as the science of selecting and interconnecting hardware components to

create computers that meet functional, performance and cost goals.

Extending the concept of the architecture and making these hardware components work in a harmonized manner to achieve a common objective in an environment is known as *computer organization*. Computer organization is like a group of people who work together to achieve a common goal, for example, a football team where different members of the team play specific roles in the game. The role of the goalkeeper is to safeguard the goal. Similarly, other members of the team (striker, mid-fielder and defender) have their respective roles to perform. However, the ultimate objective of the team as a whole is to defeat the opponent by scoring more goals than the opponent. Hence, the study of computer organization focuses more on the collective contribution from the hardware peripherals than individual electronic components.

2.2 CENTRAL PROCESSING UNIT

The central processing unit (CPU) is referred to as the "brain" of a computer system and it converts data (input) into meaningful information (output). It is a highly complex, extensive set of electronic circuitry, which executes stored program instructions. It controls all internal and external devices, performs arithmetic and logic operations, and operates only on binary data, that is, data composed of 1s and 0s. In addition, it also controls the usage of the main memory to store data and instructions, and controls the sequence of operations (Figure 2.1).

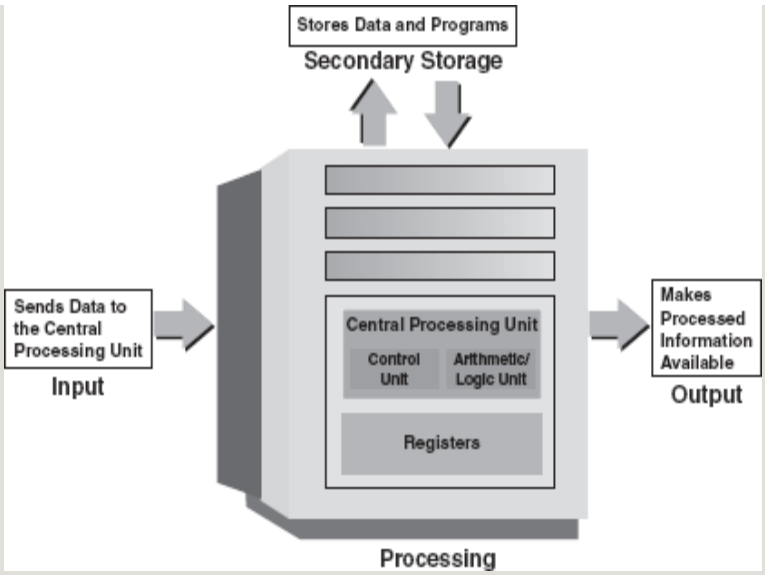


Figure 2.1 The Central Processing Unit

The CPU consists of three main subsystems, the *Arithmetic/Logic Unit (ALU)*, the *Control Unit (CU)* and the *registers*. The three subsystems work together to provide operational capabilities to a computer.

2.2.1 Arithmetic/Logic Unit

The arithmetic/logic unit (ALU) contains the electronic circuitry that executes all arithmetic and logical operations on the data made available to it. The data required to perform the arithmetic and logical functions are inputs from the designated registers. The ALU comprises two units: Arithmetic Unit (AU) and Logic Unit (LU).

Arithmetic Unit: The AU contains the circuitry that is responsible for performing the actual computing and carrying out arithmetic calculations, such as addition, subtraction, multiplication, and division. It can perform these operations at a very high speed.

Logic Unit: The LU enables the CPU to make logical operations based on the instructions provided to it. These operations are logical comparisons between data items. The LU can compare numbers, letters or special characters and then take action based on the result of the comparison. Logical operations of the LU test for three conditions:

- **Equal-to Condition**: In a test for this condition, the ALU compares two values to determine if they are equal. For example, if the number of tickets sold equals the number of seats in a cinema hall, then housefull is declared.
- **Less-than Condition**: To test this condition, the ALU compares values to determine if one is less than the other; for example, if a person purchases less than two tickets, then no discount rate is given.
- **Greater-than Condition**: In this type of comparison, the computer determines if one value is greater than the other; for example, if a person purchases greater than two tickets, then the discount rate is 5 per cent.

2.2.2 Registers

The registers are special-purpose, high-speed temporary memory units. These are temporary storage areas for holding various types of information, such as data, instructions, addresses and intermediate results of calculations. Essentially, they hold the information that the CPU is currently working on. The registers can be thought of as the CPU's working memory, a special additional storage location that offers the advantage of speed. The registers work under the direction of the CU to accept, hold, and transfer instructions or data, and perform arithmetic or logical comparisons at high speed. The CU uses a data-storage register in the similar way a store owner uses a cash register as a temporary, convenient place to store the transactions. As soon as a particular instruction or piece of data is processed, the next instruction immediately replaces it, and the information that results from the processing is returned to the main memory. Figure 2.2 reveals the types of registers present inside the CPU.

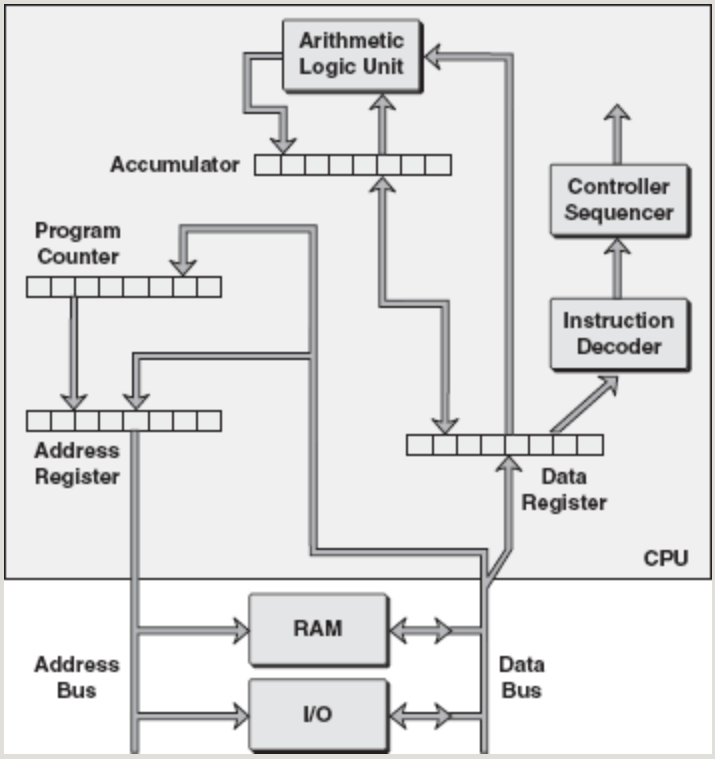


Figure 2.2 Register in the CPU

Instruction addresses are normally stored in consecutive registers and are executed sequentially. The CU reads an instruction from the memory by a specific address in the register and executes it. The next instruction is then fetched from the sequence and executed, and so on. This type of instruction sequencing is possible only if there is a counter to calculate the address of the instruction that has been executed. This counter is one of the registers, which stores intermediate data used during the execution of the instructions after they are read from the memory. Table 2.1 lists some of the important registers used in the CPU.

Table 2.1 Registers and Functions

Register Name	Function
Program Counter (PC)	The PC keeps track of the next instruction to be executed
Instruction Register (IR)	The IR holds the instruction to be decoded by the CU
Memory Address Register (MAR)	The MAR holds the address of the next location in the memory to be accessed
Memory Buffer Register (MBR)	The MBR is used for storing data either coming to the CPU or data being transferred by the CPU
Accumulator (ACC)	The ACC is a general-purpose register used for storing temporary results and results produced by the ALU
Data Register (DR)	The DR is used for storing the operands and other data

THINGS TO REMEMBER

Logical Condition

A computer can test for more than one condition. These are listed below:

- Equal to =
- Less than <
- Greater than >
- Less than or equal to <=
- Greater than or equal to >=
- Not equal to !=

The size or length of each register is determined by its function; for example, the MAR, which holds the address of the next location in the memory to be accessed, must have the same number of bits as the memory address.

The IR holds the next instruction to be executed and, therefore, should be of the same number of bits as the instruction.

2.2.3 Control Unit

The control unit (CU) of the CPU contains circuitry that uses electrical signals to direct the entire computer system to carry out, or execute, stored program instructions. This resembles an orchestra leader who himself does not play a musical instrument but directs other people to play the musical instrument in a harmonized manner. The CU also does not execute program instructions; rather, it directs other parts of the system to do so by communicating with both the ALU and the memory.

The CU controls the I/O devices and transfer of data to and from the primary storage. The CU itself is controlled by individual instructions in programs located in the primary storage. Instructions are retrieved from the primary storage, one at a time. For this, the CU uses the IR for holding the current instruction and an instruction pointer to hold the address of the next instruction. Each instruction is interpreted (decoded) so that it can be executed. Based on the instructions, the CU controls how other parts of the CPU and, in turn, the rest of the computer system should work so that the instructions are executed in a correct manner. An analogy can be considered between the CU and the traffic police; the control unit decides which action will occur just as the traffic police takes decisions on which lanes the traffic will move or stop.

Figure 2.3 illustrates how the CU instructs other parts of the CPU (ALU and registers) and the I/O devices on what to do and when to do. It also determines what data are needed, where they are stored and where to store the results of the operation, as well as sends the control

signals to the devices involved in the execution of the instructions. It administers the movement of a large amount of instructions and data used by the computer. In order to maintain the proper sequence of events required for any processing task, the CU uses clock inputs.

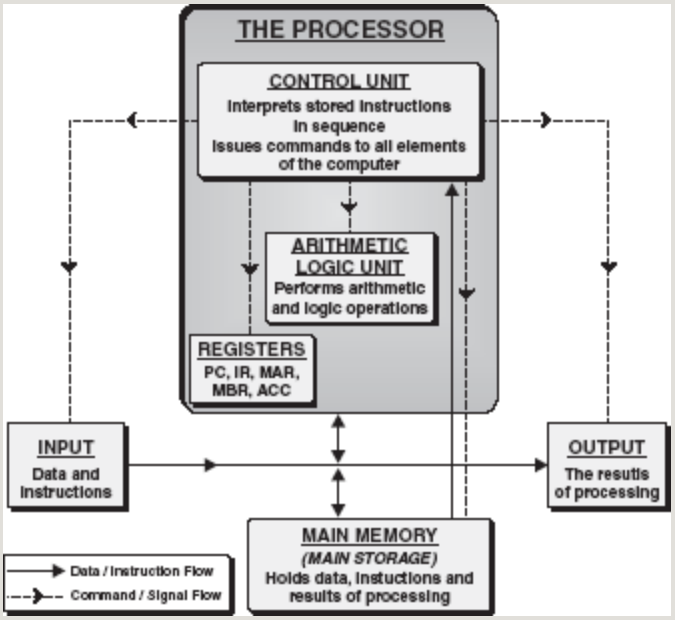


Figure 2.3 Control Unit

2.2.4 System Bus

A *bus* is a set of connections between two or more components/devices, which are designed to transfer several/all bits of a word from the source to the destination. It is a shared medium of information transfer. A bus consists of multiple paths, which are also termed *lines*, and each line is capable of transferring one bit at a time. Thus, to transmit 8 bits simultaneously over a bus, eight lines are required. In addition, some other lines are needed to control this transfer. A bus can be

unidirectional (transmission of data can be in only one direction) or bi-directional (transmission of data can be in both directions). In a shared bus, only one source can transmit data at one time, while one or more than one can receive that signal. A bus that connects to all three components (CPU, memory, I/O devices) is called a *system bus* (see Figure 2.4). A system bus consists of 50–100 separate lines. These lines are broadly categorized into three functional groups.

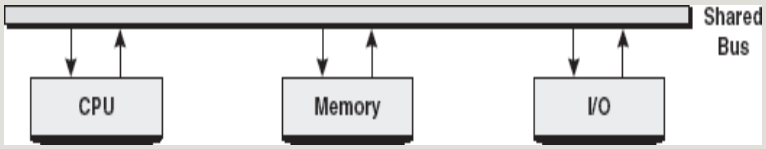


Figure 2.4 System Bus

Data Lines: Data lines provide a path for moving data between the system modules. Data lines are collectively known as a *data bus*. Normally, a data bus consists of 8, 16 or 32 separate lines. The number of lines present in a data bus is called the *width* of the data bus. Data bus width limits the maximum number of bits, which can be transferred simultaneously between two modules. The width of a data bus helps in determining the overall performance of a computer system.

Address Lines: Address lines are used to designate the source of data for a data bus. As the memory may be divided into a linear array of bytes or words, for reading or writing any information on to the memory, the CPU needs to specify the address of a particular location. This address is supplied by the address bus (address lines are collectively called an *address bus*). Thus, the width of the address bus specifies the maximum possible memory supported by a system; for example, if a system has a 16-

bit-wide address bus, it can support memory size of equal to 2^{16} (or 65,536) bytes.

Control Lines: Control lines are used to control the access to data and the address bus; this is required as a bus is a shared medium. The control lines are collectively called a *control bus* (see Figure 2.5). These lines are used for transmission of commands and timing signals (which validate data and addresses) between the system modules. Timing signals indicate whether data and address information is valid, whereas command signals specify which operations are to be performed. Some of the control lines of a bus are required to provide clock signals to synchronize operations and reset signals to initialize the modules. Control lines are also required for reading/writing to I/O devices or the memory. Control lines if used as a bus request indicate that a module needs to gain control of the bus. The bus grant control line is used to indicate whether the requesting module has been granted control of the bus.

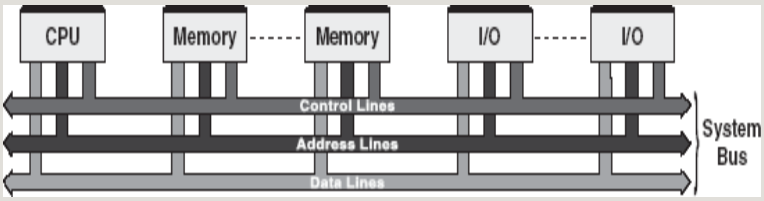


Figure 2.5 Bus Interconnection

Note: *Physically, a bus is a number of parallel electrical conductors. These circuits are normally imprinted on printed circuit boards. The bus normally extends across most of the system components, which can be tapped into the bus lines.*

2.2.5 Main Memory Unit

The memory is that part of the computer that holds data and instructions for processing. Logically, it is an integral

component of the CPU, but, physically, it is a separate part placed on the computer's motherboard. The memory stores program instructions or data for only as long as the program they pertain to is in operation. The CPU accesses the main memory in a random manner, that is, the CPU can access any location of this memory to either read information from it or store information in it. The primary memory is of two types: the first is called the *Random Access Memory (RAM)* and the second is the *Read Only Memory (ROM)*.

The RAM directly provides the required information to the processor. It can be defined as a block of sequential memory locations, each of which has a unique address determining the location and those locations contain a data element. Storage locations in the main memory are addressed directly by the CPU's instructions. It is volatile in nature, which means the information stored in it remains as long as the power is switched on. As soon as the power is switched off, the information contained in the RAM is lost.

The ROM stores the initial start-up instructions and routines in the Basic Input/Output System (BIOS), which can only be read by the CPU, each time it is switched on. The contents of the ROM are not lost even in the case of a sudden power failure, thus making it non-volatile in nature. The instructions in the ROM are built into the electronic circuits of the chip, called *firmware*. The ROM is also randomly accessible in nature, which means the CPU can randomly access any location within the ROM. Improvement in technology for constructing a flexible ROM comes in various types, namely Programmable Read Only Memory (PROM), Erasable Programmable Read Only Memory (EPROM) and Electrically Erasable Programmable Read Only Memory (EEPROM).

2.2.6 Cache Memory

The cache is a very high speed, expensive piece of memory, which is used to speed up the memory-retrieval process. Due to its higher cost, the CPU comes with a relatively small amount of cache compared with the main memory. Without the cache memory, every request from the CPU to access data would be sent to the main memory, which in response sends the data back to the CPU through the system bus. This is a slow process in computing terms. The idea of introducing cache is that this extremely fast memory would store data that are frequently accessed and, if possible, the data that are around it. This is to achieve the quickest possible response time to the CPU.

Let us consider an example of a library system for a better understanding of the cache concept. When a person comes into the library and asks for a book, the librarian searches for the requested book in the bookshelves, retrieves the book and delivers it to the person. After the person reads the book, it is returned to the library. At the same time, if any other person comes in requesting the same book, which is still placed on the librarian's desk, the librarian does not have to go through the process of searching the book in the bookshelves. This results in saving of precious time in retrieving the requested book and thus improving the efficiency of the work.

The above analogy can be related to the computer's cache system. The computer uses logic to determine which data are the most-frequently accessed and keeps them in the cache. A cache is a piece of very fast memory made from a high-speed static RAM that reduces the access time of the data. It is very expensive and generally incorporated in the processor, where valuable data and program segments are kept. The cache memory can be categorized

into three levels: *L1 cache*, *L2 cache* and *L3 cache* (see Figure 2.6).

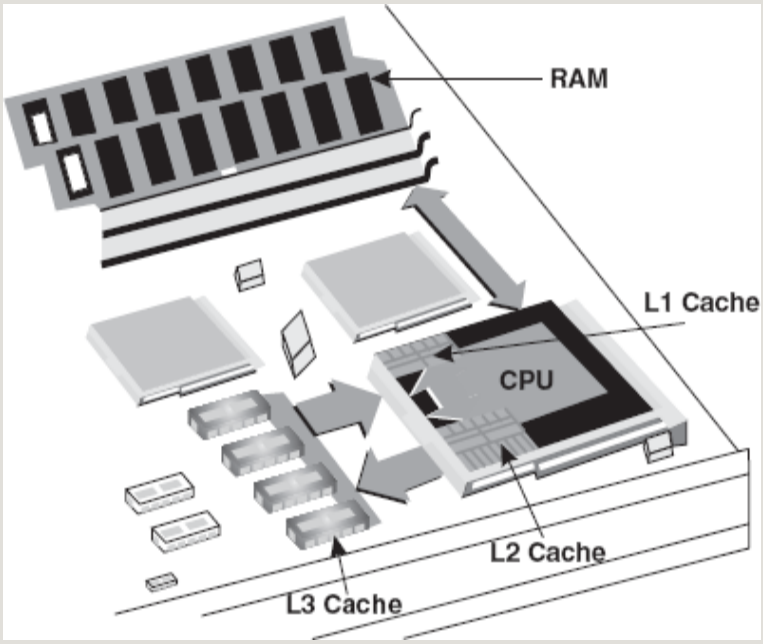


Figure 2.6 Various Levels of Cache

L1 Cache: This cache is closest to the processor and hence is termed as the primary or L1 cache. Each time the processor requests information from the memory, the cache controller on the chip uses special circuitry to first check if the requested data are already in the cache. If they are present, then the system is spared from time-consuming access to the main memory. In a typical CPU, the primary cache ranges in size from 8 to 64 KB, with larger amounts on the newer processors. This type of cache memory is very fast because it runs at the speed of the processor since it is integrated into it. There are two different ways that the processor can organize its primary cache: first, some processors have a single cache to handle both command instructions and program data (called a *unified cache*), while others have a separate

data and instruction cache called *split cache*. However, the overall performance difference between integrated and separate primary cache is not significant.

L2 Cache: The L2 cache is larger but slower in speed than L1 cache. It is used to see recent accesses that is not picked by the L1 cache and is usually 64 to 2 MB in size. The L2 cache is also found on the CPU. If the L1 and L2 cache are used together, then the missing information that is not present in the L1 cache can be retrieved quickly from the L2 cache.

L3 Cache: The L3 cache memory is an enhanced form of memory present on the motherboard of the computer. It is an extra cache built into the motherboard between the processor and main memory to speed up the processing operations. It reduces the time gap between the request and retrieval of the data and instructions, thereby accessing data much more quickly than the main memory. The L3 cache is being used with processors nowadays having more than 3 MB of storage in it.

2.3 COMMUNICATION AMONG VARIOUS UNITS

All units in a computer system work in conjunction with each other to formulate a functional computer system. To have proper coordination among these units (processor, memory and I/O devices), a reliable and robust means of communication is required. Let us discuss one of the most important functions in the computer system, that is, the communication between these units.

2.3.1 Processor to Memory Communication

The whole process of communication between the processor and memory can be divided into two steps, namely, information transfer from the memory to processor and writing information in the memory. The following sequence of events takes place when

information is transferred from the memory to the processor (see Figure 2.7 also):

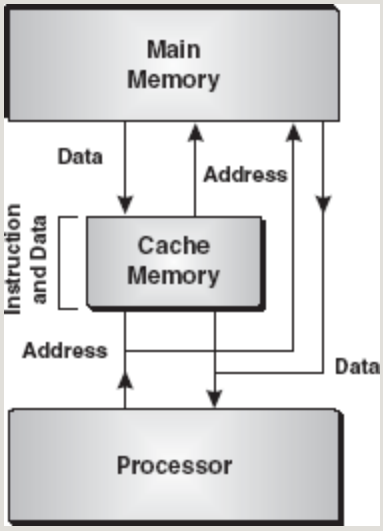


Figure 2.7 Processor-Memory

1. The processor places the address in the MAR through the address bus.
2. The processor issues a READ command through the control bus.
3. The memory places required data on the data bus, which are then transferred to the processor.

Based on the read time of the memory, a specific number of processor clock intervals are allotted for completion of this operation. During this interval, the processor is forced to wait. Similarly, the following sequence of events takes place when information is written in the memory:

1. The processor places the address in the MAR through the address bus.
2. The processor places the data to be written in the memory on the data bus.
3. The processor issues a WRITE command to the memory by the control bus.

4. The data are written in the memory at the address specified in the MAR.

The main concern in the processor-memory communication is the speed mismatch between the memory and processor. The memory speed is slower than the CPU's speed. Hence, the CPU is forced to wait for the data. As already discussed, this speed mismatch is reduced by using the small fast memory (called *cache*) as an intermediate buffer between the processor and memory.

2.3.2 Processor—I/O Devices Communication

The I/O units are connected to the computer system through the system bus. Each I/O device in a computer system is first met with the controller, called the *Direct Memory Access* (*DMA*) controller, which controls the operation of that device. The controller is connected to the buses to perform a sequence of data transfers on behalf of the CPU (see Figure 2.8). It is capable of taking over control of the system bus from the CPU, which is required to transfer data to and from the memory over the system bus. The DMA controller can directly access the memory and is used to transfer data from one memory location to another or from an I/O device to the memory and vice versa. The DMA controller can use the system bus only when the CPU does not require it or it should suspend the operations currently being processed by the CPU.

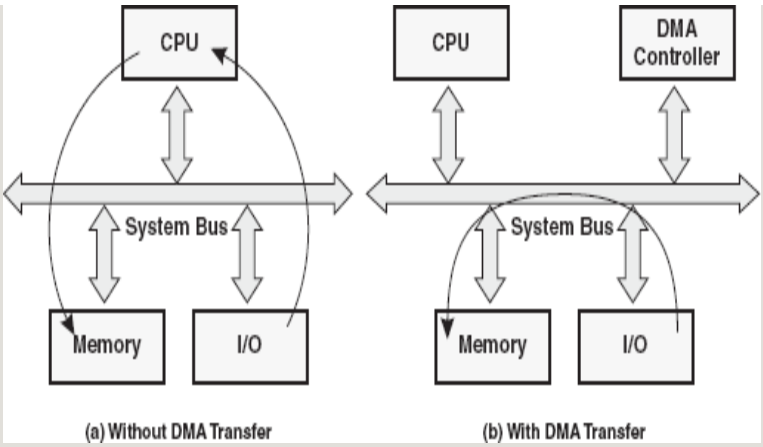


Figure 2.8 Processor—I/O Communication

With the DMA controller, a dedicated data-transfer device reads the incoming data from a device and stores that data in a system memory buffer for later retrieval by the CPU. The DMA controller allows peripheral devices to access the memory for both read and write operations without affecting the state of the computer's central processor. As a result, the data-transfer rate is significantly increased, thus improving the system efficiency. When a large amount of data is to be transferred from the CPU, the DMA controller can be used. The DMA controller allows the I/O unit to exchange data directly with the memory without going through the CPU except at the beginning (to issue the command) and at the end (to clean up after the command is processed). While the I/O operation is being performed by the DMA controller, the CPU can start execution of some other part of the same program or can start executing some other program. Thus, the DMA controller increases the speed of I/O operations by taking over buses and thus eliminating the CPU's intervention.

An instruction consists of an opcode and one or more operands, which may be addressed implicitly or explicitly. To define the layout of the bits allocated to these elements of instructions, an instruction format is used. The instruction format also indicates (implicitly or explicitly) the addressing mode used for each operand in that instruction. Figure 2.9 illustrates the format of an instruction for the processor. The first three bits represent the opcode and the last six bits represent the operand. The middle bit specifies whether the operand is a memory address or a number. When the bit is set to 1, the operand represents a number. Note that for most of the instruction sets, more than one instruction format is used. Over the years, a variety of instruction formats have been used, but designing of an instruction format involves many complex issues. Some of the key issues are discussed below:

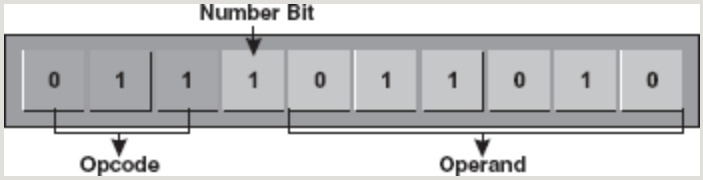


Figure 2.9 Instruction Format

1. **Instruction Length**: The core designing issue involved in an instruction format is the designing of the instruction length. The instruction length determines the flexibility of the machine. The decision on the length of an instruction depends on the memory size, memory organization and memory transfer length. There exists a trade-off between a desire of having a powerful instruction range and need of saving space.
2. **Allocation of Bits**: For the given instruction length, there is a trade-off between the number of opcodes and the power of the addressing capability. More opcode means more bits in the opcode field, for an instruction format of a given length, which reduces the number of bits available for addressing.

The basic function performed by the CPU is the execution of a program. The program to be executed is a set of instructions, which are stored in the memory. The CPU executes the instructions of the program to complete a given task. The CPU fetches instructions stored in the memory and then executes the fetched instructions within the CPU before it can proceed to fetch the next instruction from the memory. This process is continued until specified to stop. The instruction execution takes place in the CPU registers, which are used as temporary storage areas and have limited storage space. These CPU registers have been discussed earlier.

In the simplest form, the instruction processing consists of two cycles, *fetch cycle* and *execute cycle*, as shown in Figure 2.10. Here, the CPU fetches (reads) instructions from the memory, one at a time, and performs the operation specified by this instruction. The instruction fetching involves reading of an instruction from a memory location to the CPU; the execution of this instruction may involve several operations depending on the nature of instructions. The processing needed for a single instruction (fetch and execution) is referred to as the *instruction cycle*.

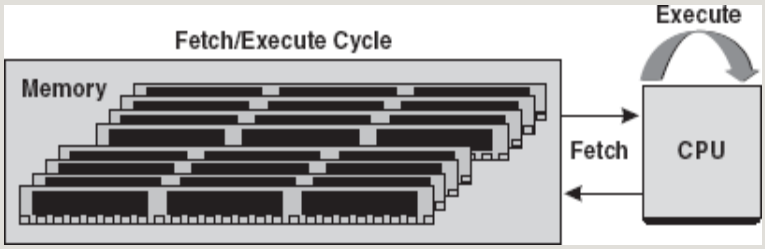


Figure 2.10 Two-step Instruction Cycle

Fetch Cycle: In the beginning, the address, which is stored in the PC, is transferred to the MAR. The CPU then transfers the instruction located at the address stored in the MAR to the MBR through the data lines connecting the CPU to memory. This transfer from the memory to the CPU is coordinated by the CU. To finish the cycle, the newly fetched instruction is transferred to the IR and, unless instructed otherwise, the CU increments the PC to point to the next address location in the memory.

Figure 2.11 illustrates the fetch cycle, which can be summarized in the following points:

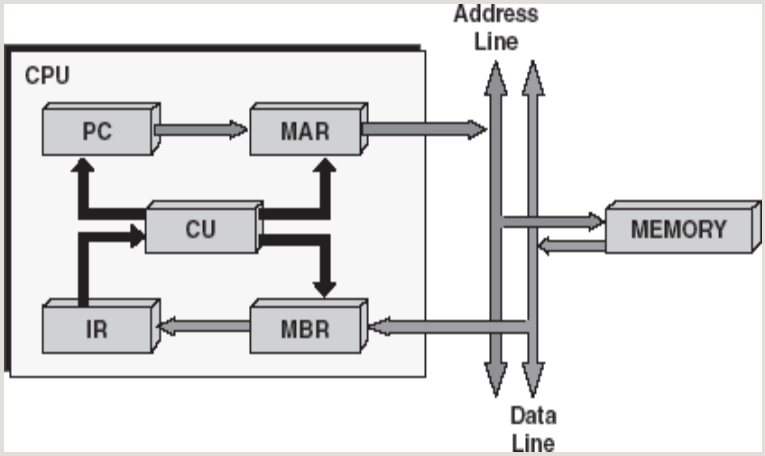


Figure 2.11 Fetch Cycle

1. PC=>MAR
2. MAR => MEMORY=>MBR
3. MBR=>IR
4. CU increments PC

After the CPU has finished fetching an instruction, the CU checks the contents of the IR and determines which type of execution is to be carried out next. This process is

known as the *decoding phase*. The instruction is now ready for the execution cycle.

Execution Cycle: Once an instruction has been loaded into the IR and the CU has examined and decoded the fetched instruction and determined the required course of action to take, the execution cycle can commence. Unlike the fetch cycle and the interrupt cycle, both of which have a set of instruction sequences, the execution cycle can contain some complex operations. The actions within the execution cycle can be categorized into the following four groups:

1. **CPU to Memory**: Data may be transferred from the memory to the CPU or from the CPU to the memory.
2. **CPU to I/O**: Data may be transferred from the I/O module to the CPU and vice versa.
3. **Data Processing**: The CPU may perform some arithmetic or logic operation on the data via the ALU.
4. **Control**: An instruction may specify that the sequence of operation may be altered; for example, the PC may be updated with a new memory address to reflect that the next instruction fetched should be read from this new location.

For simplicity, the example [LOAD ACC, memory] (illustrated in Figure 2.12) deals with one operation that can occur. The example [LOAD ACC, memory] can be classified as a memory-reference instruction. Instructions that can be executed without leaving the CPU are referred to as non-memory reference instructions.

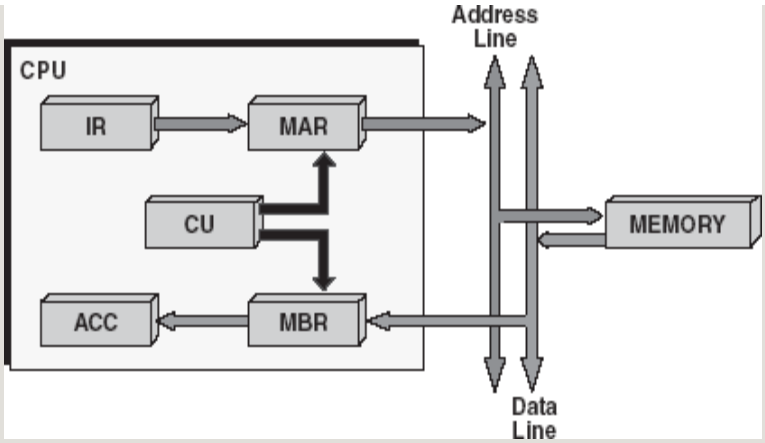


Figure 2.12 Execution Cycle

This operation loads the accumulator with data that is stored in the memory location specified in the instruction. The operation starts by transferring the address portion of the instruction from the IR to the MAR. The CPU then transfers the instruction located at the address stored in the MAR to the MBR via the data lines connecting the CPU to the memory. This transfer from the memory to the CPU is coordinated by the CU. To finish the cycle, the newly fetched data is transferred to the ACC.

The illustrated LOAD operation (see Figure 2.12)can be summarized in the following points:

1. IR [adderess portion] =>MAR
2. MAR =>MEMORY=>MBR
3. MBR =>ACC

After the execution cycle completes, the next instruction is fetched and the process starts again.

2.6 INSTRUCTION SET

Processors are built with the ability to execute a limited set of basic operations. The collections of these

operations are known as the *instruction set* of a processor. An instruction set is necessary so that a user can create machine language programs to perform any logical and/or mathematical operations. The instruction set is hardwired (embedded) in the processor, which determines the machine language for the processor. The more complex the instruction set is, the slower the processor works.

Processors differ from one another by their instruction set. If the same program can run on two different processors, they are said to be compatible; for example, programs written for IBM computers may not run on Apple computers because these two architectures (different processors) are not compatible. Since each processor has its unique instruction set, machine language programs written for one processor will normally not run on a different processor. Therefore, all operating systems and software programs are constructed within the boundaries of the processor's instruction set. Thus, the design of the instruction set for the processor becomes an important aspect of the computer architecture. Based upon the instruction sets, there are two common types of architectures: *Complex Instruction Set Computer* (*CISC*) and *Reduced Instruction Set Computer* (*RISC*).

2.6.1 CISC Architecture

Earlier, programming was done in low-level languages such as the machine language and assembly language. These languages were executed very quickly on computers, but were not easy for programmers to understand and code. To overcome these shortcomings and make programming more accessible to the masses, high-level languages were developed. These languages resembled the English language and were user-friendly. However, instructions in high-level languages were required to be converted into their equivalent low-level

languages before the processor can execute them. This conversion process was performed by the compiler. With continuous development in high-level languages, they became more powerful and provided more features (e.g. complex mathematical functions). Writing compilers for such high-level languages became increasingly difficult. Compilers had to translate complex subroutines into long sequences of machine instructions. The development of a compiler was a tricky, error-prone and time-consuming process.

To make the compiler development easier, the CISC was developed. The sole motive of manufacturers of CISC-based processors was to manufacture processors with a more extensive and complex instruction set. It shifted most of the burden of generating machine instructions to the processor. For example, instead of making a compiler to write long machine instructions for calculating a square root, a CISC processor would incorporate a hardwired circuitry for performing the square root in a single step. Writing instructions for a CISC processor is comparatively easy because a single instruction is sufficient to utilize the built-in ability. In fact, the first PC microprocessors were CISC processors because all the instructions that the processor could execute were built into the processors. As the memory was expensive in the early days of computers, CISC processors saved the memory because their instructions could be fed directly into the processor. Most of the PCs today include a CISC processor.

Advantages of the CISC Architecture: The advantages of the CISC architecture are as follows:

- At the time of their initial development, CISC machines used available technologies to optimize computer performance.
- CISC architecture uses general-purpose hardware to carry out commands. Therefore, a new command can be added into the chip without changing the structure of the instruction set.

- Microprogramming is as easy as an assembly language to implement and much less expensive than hardwiring the CU.
- As each instruction became more capable, fewer instructions could be used to implement a given task. This makes efficient use of the relatively slow main memory.
- As microprogram instruction sets can be written to match the constructs of high-level languages, the compiler does not have to be very complex.

Disadvantages of the CISC Architecture: The disadvantages of CISC architecture are as follows:

- Processors of early generation of computers were contained as a subset in succeeding versions, so the instruction set and chip hardware became complex with each generation of computers.
- Different instructions take different amount of clock time to execute and thus slow down the overall performance of the machine.
- CISC architecture requires continuous reprogramming of the on-chip hardware.
- CISC design includes the complexity of the hardware needed to perform many functions and the complexity of the on-chip software needed to make the hardware do the right thing.

2.6.2 RISC Architecture

The RISC is a processor architecture that utilizes a small, highly optimized set of instructions. The concept behind the RISC architecture is that a small number of instructions are faster in execution as compared to a single long instruction. To implement this, the RISC architecture simplifies the instruction set of the processor, which helps in reducing the execution time. Optimisation of each instruction in the processor is done through a technique known as *pipelining*. Pipelining allows the processor to work on different steps of the instruction at the same time; using this technique, more instructions can be executed in a shorter time. This is achieved by overlapping the fetch, decode and execute cycles of two or more instructions. To prevent more interactions with the memory or to reduce the access time, the RISC design incorporates a larger number of registers.

As each instruction is executed directly using the processor, no hardwired circuitry (used for complex instructions) is required. This allows RISC processors to be smaller, consume less power and run cooler than CISC processors. Due to these advantages, RISC processors are ideal for embedded applications such as mobile phones, PDAs and digital cameras. In addition, the simple design of a RISC processor reduces its development time as compared to a CISC processor.

Advantages of the RISC Architecture: The advantages of the RISC architecture are as follows:

- A simplified instruction set allows for a pipelined, superscalar design RISC processor to often achieve two to four times the performance of a CISC processor using comparable semiconductor technology and the same clock rates.
- As the instruction set of a RISC processor is simple, it uses less chip space. Extra functions such as memory management units or floating point arithmetic units can also be placed on the same chip. Smaller chips allow a semiconductor manufacturer to place more parts on a single silicon wafer, which can lower the per-chip cost significantly.
- Since the RISC architecture is simpler than the CISC architecture, it can be designed more quickly and can take advantage of other technological developments faster than the corresponding CISC designs, leading to greater leaps in performance between generations.

Disadvantages of the RISC Architecture: The disadvantages of RISC architecture are as follows:

- The performance of a RISC processor depends largely on the code that it is executing. If the compiler does a poor job of instruction scheduling, the processor can spend time waiting for the result of one instruction before it can proceed with subsequent instructions.
- Instruction scheduling makes the debugging process difficult. If scheduling (and other optimisations) is turned off, the machine-language instructions show a clear connection with their corresponding lines of source. However, once instruction scheduling is turned on, the machine language instructions for one line of source may appear in the middle of the instructions for another line of source code.
- RISC machines require very fast memory systems to feed instructions. RISC-based systems typically contain large memory cache usually on the chip itself.

2.6.3 Comparing the CISC with RISC

The CISC processor came with complex instructions sets, where decoding and execution of instructions was a complicated and time-consuming task. Moreover, with the development of high-level languages, using the instruction set posed problems in the compiler's design. As the CISC processors were less memory (very expensive earlier) intensive, it resulted in the rapid growth of CISC processors. With time, memory prices reduced drastically but CISC processors could not optimally use this availability of cheap memory. Manufacturers started working towards processors, which could run faster by using extra memory. This idea gave birth to the RISC architecture, which included small, highly optimized instructions but were more memory intensive.

The difference between the RISC approach and the CISC approach (see Figure 2.13) can be best explained by an example, which shows how each design carries out five multiplications tasks. The general steps required to perform the multiplications are as follows:

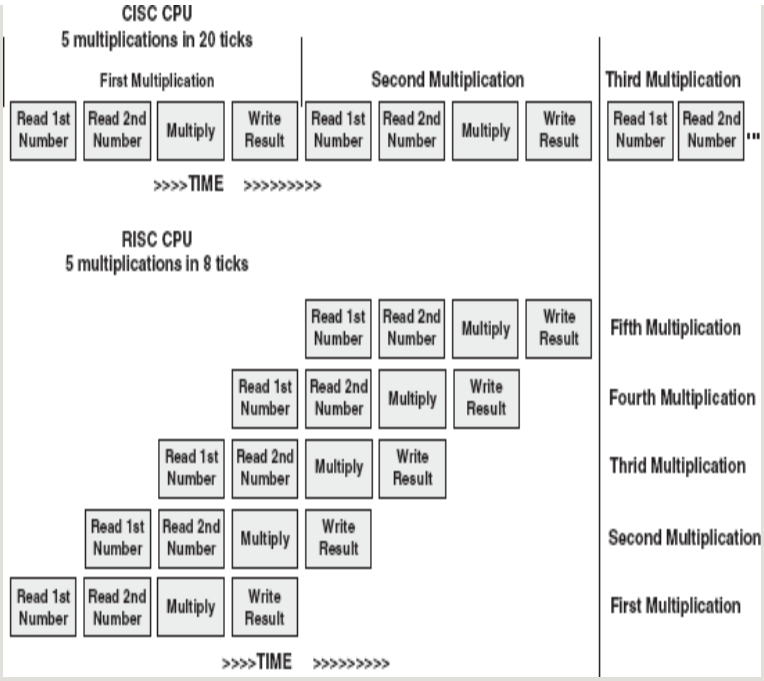


Figure 2.13 CISC versus RISC Operations

1. Read the first number out of the memory.
2. Read the second number out of the memory.
3. Multiply the two numbers.
4. Write the result back to the memory.
5. Repeat Steps 1–4 for each of the four remaining multiplications.

On a simple CISC-based CPU, the CPU is first configured to get (read) the numbers and then the numbers are read. Next, the CPU is configured to multiply the numbers and then the numbers are multiplied. Next, the CPU is configured to write the result to the memory. Finally, the numbers are written into the memory. To multiply five sets of numbers, the whole process must be repeated five times.

On a simple RISC CPU, the process is slightly different. A piece of hardware on the CPU is dedicated to read the first number. When this operation is complete, another piece of hardware reads the second number. After completion of this operation, another hardware performs

the multiplication and when it is completed, another hardware writes the result to the memory. If this operation happens five times in a row, the RISC hardware dedicated to obtaining the first number from the memory obtains the first number for the second operation and second number for the first operation. At the same time, the second number for the second operation is retrieved from the memory, while the first number for the third operation is obtained. As the first result is written back to the memory, the second multiplication is performed, while the second number for the third operation is read from the memory and the first number for the fourth operation is read from the memory, and so on.

2.7 INSIDE A COMPUTER

Computing machines are complex devices made from numerous electronic components. Many of these components are small, sensitive, expensive and operate with other components to provide better performance to the computing machines. Therefore, to ensure better performance and increase in the life of these components, they are placed inside a metal enclosure called the *system case* or *cabinet* (see Figure 2.14). The system case is a metal and plastic box that houses the main components of the computer. It protects electronics hardware against the heat, light, temperature, etc. It serves important roles in the functioning of a properly designed and well-built computer. Several areas where the system case plays an important role are:

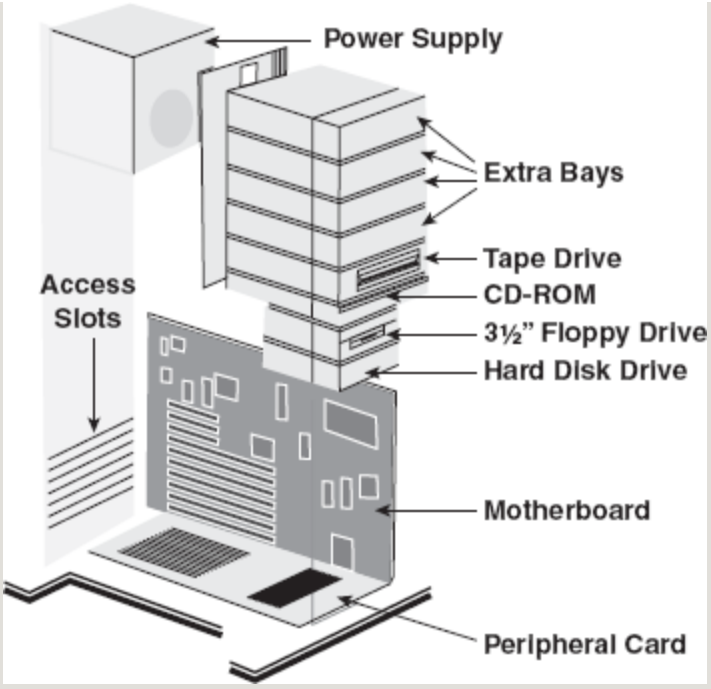


Figure 2.14 System Case

- **Structure**: The system case provides a rigid structural framework to the components, which ensures that everything fits together and works in a well-organized manner.
- **Protection**: The system case protects the inside of the system from physical damage and electrical interference.
- **Cooling:** The case provides a cooling system to the vital components. Components that run under cool temperature last longer and are less troublesome.
- **Organization and Expandability**: The system case is a key to the organization of physical systems. If a system case is poorly designed, upgradation or expansion of a peripheral is limited.
- **Status Display**: The system case contains lights or LEDs that provide varied information pertaining to the working of the system to the user.

The system case encloses all the components, which are essential in running the computer system. These components include the motherboard, processors, memory, power supply, expansion slots, cables, removable drives, etc.

FACT FILE

AC versus DC

AC (alternating current) is an electric current that reverses its direction in a circuit at regular intervals. AC is useful for appliances that work based on the fact that the current is constantly changing, like pumping motors. The current flowing from most wall outlets in our homes is AC. **DC** (direct current) is an electric current that flows in one direction only. DC is useful for appliances that run on a steady current, like emergency light.

2.7.1 Power Supply (SMPS)

Power supply or *Switched Mode Power Supply* (*SMPS*) is a transformer and voltage control device in a computer that furnishes power to all the electronic components by converting incoming AC into the low-voltage DC supply. When a computer is turned on, the power supply allows the converted electricity to travel to other components inside the computer. The modern-day power supply provides protection against surge and spikes in the power, which could damage vital components of the computer. Nowadays, the computer power supply is capable of providing different voltages, at different strengths, and manages additional signals for the motherboard. The power supply plays an important role in the following areas of the computer system (see Figure 2.15):

Figure 2.15 Power Supply

- **Stability**: A high-quality power supply with sufficient capacity to meet the demands of the computer provides years of stable power for the computer.
- **Cooling**: The power supply contains the main fan that controls the flow of air through the system case. This fan is a major component in the computer cooling system.
- **Expandability**: The capacity of the power supply determines the ability to add new drives to the system or upgrade to a more powerful motherboard or processor.

2.7.2 Motherboard

The *motherboard*, also known as the *system board*, is a large multilayered printed circuit board inside a computer. The motherboard contains the CPU, the BIOS ROM chip and the CMOS Setup information. It has expansion slots for installing different adapter cards like the video card, sound card, network interface card and modem. The circuit board provides a connector for the keyboard as well as housing to the keyboard controller chip. It possesses RAM slots for the system's RAM chips, and enables the system's chipset, controllers and underlying circuitry (bus system) to tie everything together. In a typical motherboard, the circuitry is imprinted on the surface of a firm planar surface and is usually manufactured in a single piece. The most

common design of the motherboard in today's desktop computers is the ATX design. In ATX designs, the computer components included are processor, co-processors (optionally), memory, BIOS, expansion slot and interconnecting circuitry. Additional components can be added to a motherboard through its expansion slot. Nowadays, they are designed to put peripherals as integrated chips directly onto the motherboard. Initially, this was confined to audio and video chips, but in recent times, the peripherals integrated in this way include SCSI, local area network (LAN) and RAID controllers. Though, there is cost benefit to this approach, the biggest downside is the restriction of future upgrade options. Figure 2.16 provides an insight into various components on motherboards.

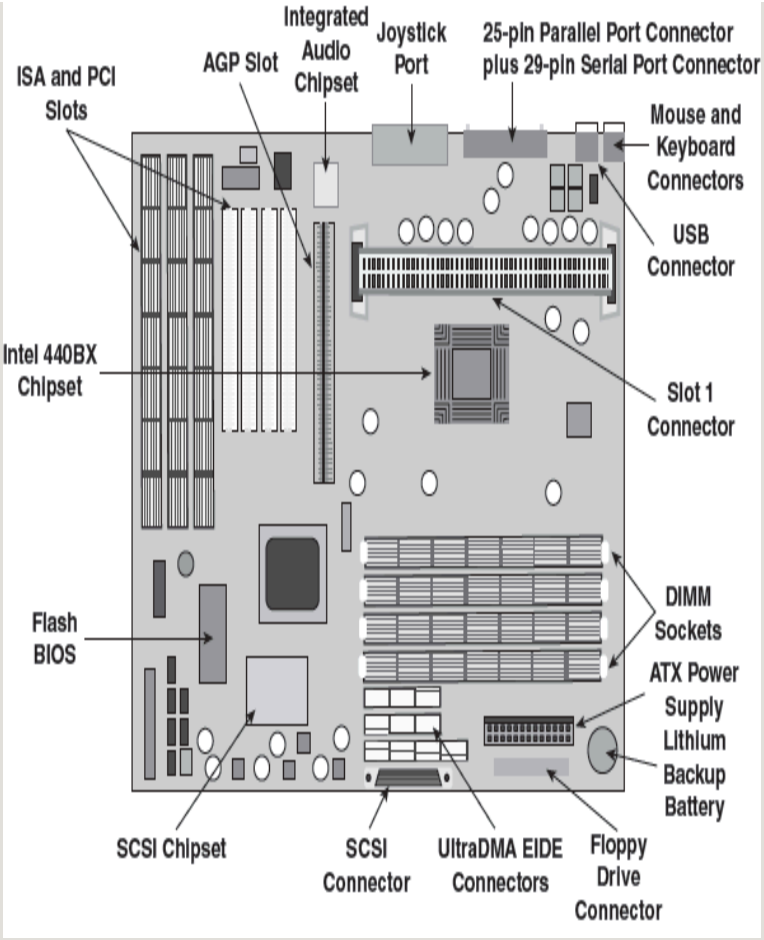


Figure 2.16 Motherboard

BIOS: The BIOS comprises a set of several routines and start-up instructions inside the ROM. This gives two advantages to the computer. First, the code and data in the ROM BIOS need not be reloaded each time the computer is started. Secondly, they cannot be corrupted by wayward applications that are accidentally written into the wrong part of the memory. As the machine is switched on, the control is transferred to the bootstrap procedure of the BIOS. This procedure inspects the computer to determine what hardware is fitted and then conducts a simple test (power-on self test) for normal functionality. If all the tests are passed, the ROM then determines the drive to boot the machine. Most

computers have the BIOS set to check for the presence of an operating system in the primary hard disk drive. Once the machine is booted, the BIOS serves a different purpose by presenting the DOS with a standardized Application Program Interface (API) for the computer hardware.

CMOS: The motherboard includes a separate block of memory called the *complementary metal oxide silicon* (*CMOS*) chip that consumes very low power. This chip is kept alive by a battery even when the computer's power is off. The function of the CMOS chip is to store basic information about the computer's configuration, such as the type of hard disks and floppy drives, memory capacity, etc. The other important data, which is kept in the CMOS memory, is the system time and date. The clock, CMOS chip and batteries are usually all integrated into a single chip.

2.7.3 Ports and Interfaces

Ports and interfaces are a generic name for the various “holes” (and their associated electronics) found at the back of the computer, using which external devices are connected to the computer's motherboard. Different interfaces and ports run at varying speeds and work best with specific types of devices (see Figure 2.17).

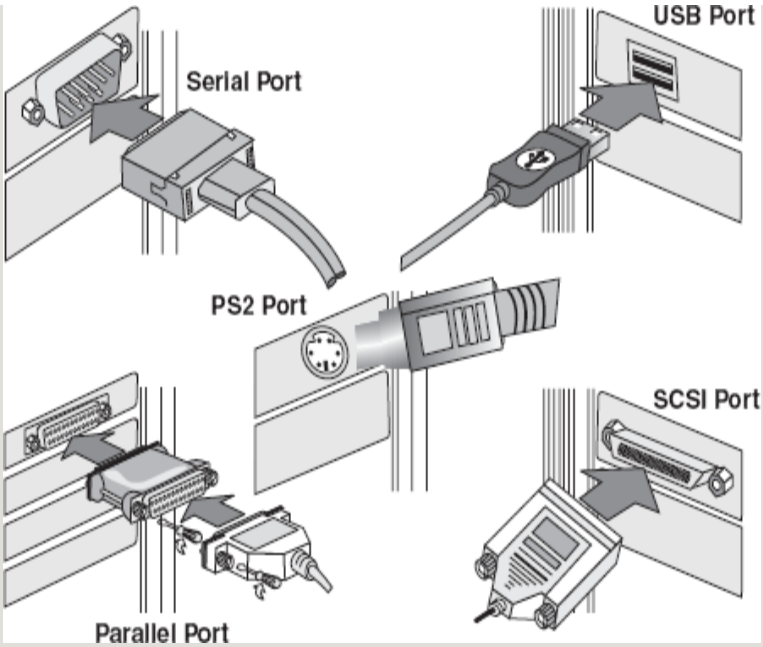


Figure 2.17 Different Types of Ports

- **PS/2 Ports**: It is a standard serial port connector used to plug the computer mouse and keyboards into the personal computer. It consists of six pins in a small, round-shape socket.
- **Serial Ports**: It is a general-purpose communications port through which data are passed serially, that is, one bit at a time. These ports are used for transmitting data over long distances. In the past, most digital cameras were connected to a computer's serial port in order to transfer images to the computer. However, because of its slow speed, these ports are used with the computer mouse and modem.
- **Parallel Port**: It is an interface on a computer, which supports transmission of multiple bits of data (usually 8 bits) at the same time. This port transmits data faster than a serial port and is exclusively used for connecting peripherals such as printers and CD-ROM drives.
- **SCSI Port:** These ports are used in transmitting data up to seven devices in the “daisy chain” fashion and at a speed faster than serial and parallel ports (usually 32 bits at a time). In the daisy chain fashion, several devices are connected in series to each other, so that data for the seventh device need to go through the entire six devices first. These ports are hardware interfaces, which include an expansion board that plugs into the computer called the *SCSI host adapter* or *SCSI* controller. Devices which can be connected to SCSI ports are hard-disk drives and network adapters.
- **Universal Serial Bus (USB) Port**: The USB port is a plug-and-play hardware interface for connecting peripherals such as the keyboard, mouse, joystick, scanner, printer and modem. It supports a maximum bandwidth of 12 MB per second and has the capability to connect up to 127 devices. With the USB port, a new device can be added to the computer without adding an adapter card. These ports are the replacement for the parallel and serial ports.

2.7.4 Expansion Cards

An *expansion card*, also called an *adapter card*, is a circuit board that provides additional capabilities to the computer system. Adapter cards are made up of large-scale integrated circuit components Sound Card installed on it. The cards are plugged into the expansion sockets present on the computer's motherboard to provide the computer an added functionality. Common available expansion cards connect monitors (for enhanced graphics) and microphones (for sound), each having a special purpose to perform (see Figure 2.18). However, nowadays most of the adapters come in-built on the motherboard and no expansion card is required, unless the need for high performance is required.

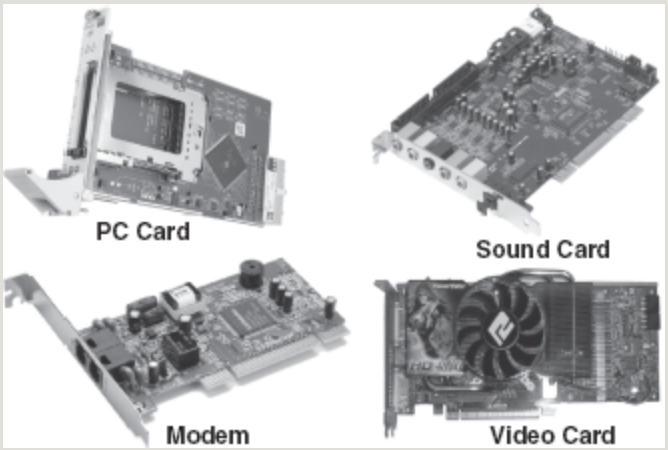


Figure 2.18 Expansion Card

- **Sound Cards**: An expansion card that allows the computer to output sound through connected speakers, to record sounds from a microphone and to manipulate sounds stored on the computer is called a sound card. It contains special circuits for operating the computer's sound and allows playback and recording of sound from the CD-ROM.
- **Video Cards**: A video card, also called the display adapter, is used for enhancing graphics images that are seen on the computer's monitor. The card converts the images created in the computer to electronic signals required by the monitor. Generally, a good card with a graphics accelerator is preferred for editing digital videos. There are different

video cards with varying capabilities related to the size of the monitor and total number of displayable colours.

- **Network Interface Card**: A network interface card is a computer circuit board that is installed in a computer so that it can be connected to other computers in a network. Personal computers and workstations on a LAN contain a network interface card specifically designed for transmitting data across LANs. Network interface cards provide a dedicated, full-time connection to a network.
- **Modem**: Modem is an expansion card that allows two computers to communicate over ordinary phone lines. It converts digital data from computers into analog data, transmits over the telephone lines and also converts incoming analog signals back to digital signals for the receiving computer. Modems do not provide high bandwidth for the data communication and as a result, they do not support high-speed Internet access as current modems can run up to 56 KBps.
- **PC Card**: A PC card is a removable device, approximately the size of a credit card, which is designed to plug into a Personal Computer Memory Card International Association (PCMCIA) slot. It is a standard formulated by the PCMCIA for providing expansion capabilities to computers. The PCMCIA standard supports input-output devices, memory, fax/modem, SCSI and networking products. The card fits into a notebook or laptop computer.

2.7.5 Ribbon Cables

Ribbon cables are wide, flat and insulated cables, which are flexible enough to fit into areas with little space. These cables are made up of numerous tiny wires (traces and electronic pathways) called the bunch, where one *bunch* carries data/information around to different components on the motherboard and another bunch connects these components to the various devices attached to the computer. These cables connect the hard drive, floppy drive and CD-ROM drive to the connectors on the motherboard and control the drives by getting and sending data from and to them. These cables connect different external devices, peripherals, expansion slots, I/O ports and drive connections to the rest of the computer (see Figure 2.19)

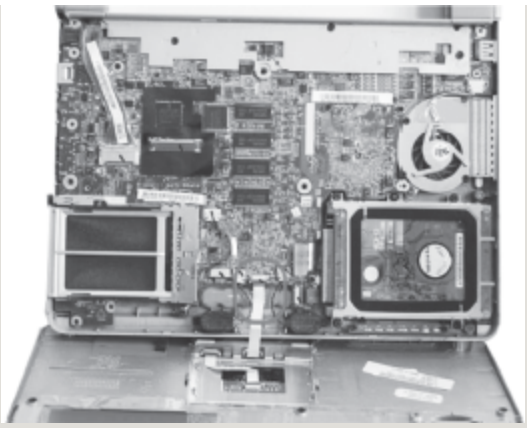


Figure 2.19 Ribbon Cables

2.7.6 Memory Chips

The memory is the place where the computer holds programs and data that are currently in use. The system memory 4.25" 72-pins SIMM on the motherboard is arranged in groups called the *memory banks*. The number of the memory banks and their configurations vary from computer to computer because these are determined by the CPU and the 168-pins DIMM way it receives information. The speed of Fi220 SIMM and DIMM RAM Chips the CPU determines the number of memory sockets required in a bank. For the main memory, either of the two types of memory chips is used: *Single In-Line Memory Modules* (*SIMM*) or *Dual In-Line Memory Modules* (*DIMM*) (see Figure 2.20).

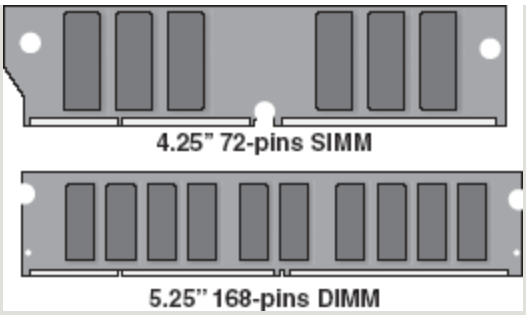


Figure 2.20 SIMM and DIMM RAM Chips

SIMM: SIMM are small circuit boards designed to accommodate surface-mount memory chips. A typical SIMM chip comprises a number of RAM chips on a printed circuit board (PCB), which fits into a SIMM socket on a computer's motherboard. These chips are packed into small plastic or ceramic dual inline packages (DIPs), which are assembled into a memory module. A typical motherboard offers four SIMM sockets capable of taking either single-sided or double-sided SIMMs with module sizes of 4, 8, 16, 32 or even 64 MB. When 32-bit SIMM chips are used with processors, they have to be installed in pairs, with each pair of modules making up a memory bank. These chips support 32-bit data paths and are originally used with 32-bit CPUs. The CPU then communicates with the memory bank as one logical unit. SIMM chips usually come in the following two formats:

- A 30-pin SIMM used in older system boards, which deliver 1 byte of data.
- A larger 72-pin SIMM used in modern PCs, which deliver 4 bytes of data (plus parity) in every memory request.

DIMM: With the increase in speed and bandwidth capability, a new standard for the memory was adopted called *dual in-line memory module* (*DIMM*). These chips have 168 pins in two (or dual) rows of contacts; one on each side of the card. With the additional pins, a CPU retrieves information from the DIMM chip at 64 bits as compared to a 32- or 16-bit transfer with SIMMs. Some

of the physical differences between 168-pin DIMMs and 72-pin SIMMs include the length of the module, the number of notches on the module and the way the module is installed. The main difference between the two is that on a SIMM, opposing pins on either side of the board are tied together to form one electrical contact, while on a DIMM, opposing pins remain electrically isolated to form two separate contacts. DIMMs are often used in computer configurations that support a 64-bit or wider memory bus (like Intel's Pentium 4).

2.7.7 Storage Devices

Disk drives are the important components present inside the system case. These drives are used to read and write information to and from the processor. The three most common disk drives located inside a system case are the hard drive, floppy disk drive and CD-ROM. These drives are high-storage devices, which enable the user to store large amount of data. Out of these drives, the hard disk drive provides the largest storage space for saving. All the vital applications ranging from the operating system to word processor are stored in the hard disk drive. The hard disk drive is costly and not robust enough to transfer data physically; therefore, CD-ROMs and floppy disks are used as an alternative means to transfer data physically.

2.7.8 Processors

Processor, often called the CPU, is the central component of the computer. It is referred to as the brain of a computer responsible for carrying out operations in an efficient and effective manner. A processor holds the key for carrying out all the processing and computational work. Every work that is done by the user on the computer is performed either directly or indirectly by the processor. The following factors should be considered while choosing a processor of a computer system:

- **Performance**: The processor's capabilities dictate the maximum performance of a system. It is the most important single determinant of system performance (in terms of speed and accuracy) in the computer.
- **Speed**: The speed of a processor defines how fast it can perform operations. There are many ways to indicate speed, but the most obvious way to measure is through the internal clock speed of the CPU. The faster the speed of the internal clock of the processor, the faster the CPU will work and, therefore, hardware will be more expensive.
- **Software Support**: New and faster processors support resource-consuming software in a better manner. For example, new processors such as the Pentium 4 enable the use of specialized software, which were not supported on earlier machines.
- **Reliability and Stability**: The reliability of the computer system directly depends on the type and quality of the processor.
- **Energy Consumption and Cooling**: Although processors consume relatively little power compared to other system devices, a newer processor consumes a great deal of power resulting in the impact on everything from the cooling method selection to overall system reliability.
- **Motherboard Support**: The type of processor used in the system is a major determining factor of the chipset used on the motherboards. The motherboard, in turn, dictates many facets of the system's capabilities and performance.

2.8 DATA REPRESENTATION IN COMPUTERS

Since the early days of human civilisation, people have been using their fingers, sticks, etc., for counting things. The need for counting probably originated when man started to use animals for domestic purposes and practise animal breeding for fulfilling his needs and requirements. As daily activities became more complex, numbers became more important in trade, time, distance and in all other spheres of human life. Ever since people discovered that it was necessary to count objects, they have been looking for easier ways of counting. To count large numbers, man soon started to count in groups and various number systems were formed.

As manual counting had a limited role for carrying out a simple computing task, computation that was more complex made humans to depend on the machines to perform the computing task efficiently and accurately. With the advancement of machines, different number systems were formed to make the task simple, accurate

and fast. These number systems worked on the principle of a digital logic design present in the modern-day computer system and opened a gateway to overcome complex computation barriers. In a precise manner, a *number system* defines a set of values used to represent "quantity". Generally, one talks about a number of students attending a class and a number of modules taken by each student as well as uses numbers to represent grades achieved by students in tests. Quantifying values and items in relation to each other is helpful for us to make sense of our environment. The number system can be categorized into two broad categories:

- **Non-Positional Number Systems**: In ancient times, people used to count with their fingers. When fingers became insufficient for counting, stones and pebbles were used to indicate the values. This method of counting is called the *non-positional number system*. It was very difficult to perform arithmetic operations with such a number system, as it had no symbol for zero. The most common non-positional number system is the Roman number system. These systems are often clumsy and it is very difficult to do calculations for large numbers.
- **Positional Number Systems**: A positional number system is any system that requires a finite number of symbols/digits of the system to represent arbitrarily large numbers. When using these systems, the execution of numerical calculations becomes simplified because a finite set of digits are used. The value of each digit in a number is defined not only by the symbol, but also by the symbol's position. The most popular positional number system being used today is the decimal number system.

Base (or Radix) of System: The word base (or radix) means the quantity of admissible marks used in a given number system. The admissible marks are the characters such as Arabic numerals, Latin letters or other recognisable marks, which are used to present the numerical magnitude of a "quantity." The decimal number system originated in India. This number system has 10 as the base and is indicated by a subscript 10 (decimal number) following the value of the number. For example, (7592)10 represents a number (7592) in the base 10 number system. Similarly, we have number systems with base other than 10. For example, (214)8 represents a number (214) in the base 8 number system.

For a computer, everything is in the digital form (binary form) whether it is a number, alphabet, punctuation mark or instruction (Table 2.2). Let us illustrate this with the help of an example. Consider the word “INDIA” that appears on the computer screen as a series of alphabetic characters. However, for the computer, it is a combination of numbers. On the basis of the American Standard Code for Information Interchange (ASCII) value (discussed later in this chapter) of each character, it appears to the computer as follows:

01001001	01001110	01000100	01001001	01000001
I	N	D	I	A

Table 2.2 Types of Number System

Number System	Radix Value	Set of Digits	Example
Decimal	$R = 10$	(0, 1, 2, 3, 4, 5, 6, 7, 8, 9)	$(25)_{10}$
Binary	$R = 2$	(0, 1)	$(11001)_2$
Octal	$R = 8$	(0, 1, 2, 3, 4, 5, 6, 7)	$(31)_8$
Hexadecimal	$R = 16$	(0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F)	$(19)_{16}$

2.8.1 Types of Number System

Generally, a user is least aware of the fact that actual operations in a computer are done with a binary number system. Traditionally, the two possible states for the binary system are represented by the digits 0 and 1. Long before the introduction of octal and hexadecimal numbers, programmers used a convenient method of handling large binary numbers in either 3-bit or 4-bit groupings. Later, the actual machine code for the computer instructions was replaced by mnemonics, which comprised of three to four letters of the assembly language for a particular CPU. It was also possible to use more than one base number for writing

instructions/programs in the assembly languages, so programmers made sure their assemblers could understand octal, hexadecimal and binary numbers. Octal and hexadecimal numbers are more compact than binary numbers, which a computer understands. Thus, octal and hexadecimal numbers prevent unwieldy strings that were written in binary. For example, the three-digit decimal number 513 requires 10 digits in pure binary (1000000001) but only three (201) in hexadecimal. Table 2.2 lists the types of number system.

Binary Number System: The digital computer provides accurate solutions to the problems by performing arithmetic computations. The numbers are not expressed as decimal numbers within the computer because it is not suitable for machine processes. Computers are not only powered by electricity, but also compute with electricity. They shift voltage pulses around internally. When numbers are represented in a computer's memory by means of small electrical circuits, a number system with only two symbols is used. These symbols are ON or OFF state of the circuit. This system of representing numbers is known as the *binary number system*. Circuits allow electricity to flow or to be blocked depending on the type of the circuit. A computer circuit is made up of transistors, which have only two states, ON and OFF. ON is interpreted as 1, while OFF as 0. Similar to the decimal system, the position of a digit in a number indicates its value. Instead of ones, tens, hundreds, thousands, etc., as in the decimal system, the columns in the binary system contain ones, twos, fours, eights, etc. Each column represents a number in powers of 2; starting with 20 for the leftmost column, the power is incremented by 1 for each additional column to the right. In other words, each place in the number represents two times (2×'s) the place to its right. Table 2.3 represents the first 10 decimal numbers in binary.

Table 2.3 Representing Decimal Numbers in Binary

Decimal Numbers	Binary Numbers				
	2^4	2^3	2^2	2^1	2^0
	16	8	4	2	1
0					0
1					1
2				1	0
3				1	1
4			1	0	0
5			1	0	1
6			1	1	0
7			1	1	1
8		1	0	0	0
9		1	0	0	1

Octal Number System: The octal number system is a base 8 system, having eight admissible marks: 0, 1, 2, 3, 4, 5, 6 and 7 with no 8 or 9 in the system. This system is a positional notation number system. The octal number system uses powers of 8 to determine the number (see Table 2.4).

Table 2.4 Representing Binary and Decimal in Octal

Decimal Number	Binary Number	Octal Number
0	0000	$000\ (0 \times 8^{0})$
1	0001	$001\ (1 \times 8^{0})$
2	0001	$002\ (2 \times 8^{0})$
3	0011	$003\ (3 \times 8^{0})$
4	0100	$004\ (4 \times 8^{0})$
5	0101	$005\ (5 \times 8^{0})$
6	0110	$006\ (6 \times 8^{0})$
7	0111	$007\ (7 \times 8^{0})$
8	1000	$010\ (1 \times 8^{1} + 0 \times 8^{0})$
9	1001	$011\ (1 \times 8^{1} + 1 \times 8^{0})$

Hexadecimal Number System: The hexadecimal number system uses 0–9 and A–F to represent a number, where A is equivalent to decimal 10 and the largest hexadecimal digit F is equivalent to decimal 15. The base for a hexadecimal number is 16 (see Table 2.5).

Table 2.5 Comparing Decimal, Binary, Octal and Hexadecimal

Decimal	Binary	Octal	Hexadecimal
0	0000	000	0
1	0001	001	1
2	0010	002	2
3	0011	003	3
4	0100	004	4
5	0101	005	5
6	0110	006	6
7	0111	007	7
8	1000	010	8
9	1001	011	9
10	1010	012	A
11	1011	013	B
12	1100	014	C
13	1101	015	D
14	1110	016	E
15	1111	017	F

2.8.2 Conversion Between Number Bases

Computers and other digital systems process information as their primary function. Therefore, it is necessary to have methods and systems for representing information in forms, which can be manipulated and stored electronically. As discussed earlier, internally, a computer uses binary numbers for data representation, whereas, externally, it uses decimal numbers. It means a computer performs the decimal to binary conversion and vice versa. In general, any number in one number system can be represented in another number system.

Converting Decimal to Binary, Octal and Hexadecimal: The method used for the conversion of a decimal number into other number systems is often done using the “remainder” method. This method involves the following steps:

1. Divide the decimal number by the base of the target number system. That is, to convert decimal to binary, divide the decimal

number with 2 (the base of the binary number system), 8 for octal and 16 for hexadecimal.
2. Note the remainder separately. In case of hexadecimal, if the remainder exceeds 9, then convert the remainder into an equivalent hexadecimal form; for example, if the remainder is 10, then note the remainder as A.
3. Continually repeat the process of dividing until the quotient is zero and keep writing the remainders after each step of division.
4. Finally, when no more division can occur, write down the remainders in reverse order.

Example 1: Determine the binary equivalent of $(36)_{10}$.

2	36	Remainder	Least Significant Bit (LSB)
2	18	0	↑
2	9	0	
2	4	1	
2	2	0	
2	1	0	
	0	1	Most Significant Bit (MSB)

Taking remainders in the reverse order, we have 100100.
Thus, the binary equivalent of $(36)_{10}$ is $(100100)_{2}$.

Example 2: Determine the octal equivalent of $(359)_{10}$.

8	359	Remainder	Least Significant Bit (LSB)
8	44	7	↑
8	5	4	
8	0	5	Most Significant Bit (MSB)

Thus, the octal equivalent of $(359)_{10}$ is $(547)_{8}$.

Example 3: Determine the hexadecimal equivalent of $(5112)_{10}$.

16	5112	Remainder	Least Significant Bit (LSB)
16	319	8 = 8	
16	19	15 = F	
16	1	3 = 3	
16	0	1 = 1	Most Significant Bit (MSB)

Thus, the hexadecimal equivalent of $(5112)_{10}$ is $(13F8)_{16}$.

Converting Binary, Octal and Hexadecimal to Decimal: The method used for the conversion of a binary, octal or hexadecimal number to decimal number involves each digit of the binary, octal or hexadecimal number to be multiplied by its weighted position and then each of the weighted values is added together to get the decimal number.

Example 1: Determine the decimal equivalent of $(11010)_2$.

Binary Number	1	1	0	1	0
Weight of Each Bit	2^4	2^3	2^2	2^1	2^0
Weighted Value	$2^4 \times 1$	$2^3 \times 1$	$2^2 \times 1$	$2^1 \times 1$	$2^0 \times 0$
Solved Multiplication	16	8	0	2	0

Sum of weight of all bits = 16 + 8 + 0 + 2 + 0 = 26. Thus, the decimal equivalent of $(11010)_2$ is $(26)_{10}$.

Example 2: Determine the decimal equivalent of $(456)_8$.

Binary Number	1	1	0	1	0
Weight of Each Bit	2^4	2^3	2^2	2^1	2^0
Weighted Value	$2^4 \times 1$	$2^3 \times 1$	$2^2 \times 1$	$2^1 \times 1$	$2^0 \times 0$
Solved Multiplication	16	8	0	2	0

Sum of weight of all bits = 256 + 40 + 6 = 302. Thus, the decimal equivalent of $(456)_8$ is $(302)_{10}$.

Example 3: Determine the decimal equivalent of $(B14)_{16}$.

Hexadecimal Number	B = 11	1	4
Weight of Each Bit	16^2	16^1	16^0
Weighted Value	$16^2 \times 11$	$16^1 \times 1$	$16^0 \times 4$
Solved Multiplication	2816	16	4

Sum of weight of all bits = 2816 + 16 + 4 = 2836. Thus, the decimal equivalent of $(B14)_{16}$ is $(2836)_{10}$.

Converting Among Binary, Octal and Hexadecimal: Converting among binary, octal and hexadecimal can be accomplished easily without converting to decimal first, as the base numbers of all three systems (2, 8 and 16, respectively) are powers of 2. Any octal digit can be written as a group of three binary digits, while a hexadecimal number will comprise four binary digits.

Example 1: Determine the octal equivalent of $(010111)_2$

Binary Number	010	111
Octal Number	2	7

The octal equivalent of $(010111)_2$ is $(27)_8$.

Example 2: Determine the hexadecimal equivalent of $(11001011)_2$.

Binary Number	1100	1011
Decimal Number	12	11
Hexadecimal Number	C	B

The hexadecimal equivalent of $(11001011)_2$ is $(CB)_{16}$.

Example 3: Determine the binary equivalent of $(231)_8$.

Octal Number	2	3	1
Binary Value	010	011	001

Thus, the binary equivalent of $(231)_8$ is $(010011001)_2$.

Example 4: Determine the binary equivalent of $(5AF)_{16}$.

Hexadecimal Number	5	A	F
Binary Value	0101	1010	1111

Thus, the binary equivalent of $(5AF)_{16}$ is $(010110101111)_2$.

Converting Between Octal and Hexadecimal: The method used for the conversion of an octal number to hexadecimal number is accomplished by the following steps:

1. Convert each octal digit to 3-bit binary form.
2. Combine all the 3-bit binary numbers.
3. Segregate the binary numbers into the 4-bit binary form by starting the first number from the right bit (LSB) towards the number on the left bit (MSB).
4. Finally, convert these 4-bit blocks into their respective hexadecimal symbols.

Example 1: Determine the hexadecimal equivalent of (2327)8.

Octal Number	2	3	2	7
Binary Value	010	011	010	111

Combining the 3-bit binary blocks, we have 010011010111. Separating the group of binary numbers (from the left side) into the 4-bit binary number and by converting these blocks into their respective hexadecimal symbols, we have

0100	1101	0111
4	D	7

Thus, the hexadecimal equivalent of $(2327)_8$ is $(4D7)_{16}$.

The method used for the conversion of a hexadecimal number to octal number is the same as that used for the octal to hexadecimal conversion, except that each hexadecimal digit is converted into 4-bit binary form and then after grouping of all the 4-bit binary blocks, it is

converted into the 3-bit binary form. Finally, these 3-bit binary forms are converted into octal symbols.

Example 2: Determine the octal equivalent of $(2B6)_{16}$.

Hexadecimal Number	2	B	6
Binary Value	0010	1011	0110

Combining all the 4-bit binary blocks, we have 001010110110. Separating the group of binary numbers into the 3-bit binary blocks and by converting these blocks into octal symbols, we have

001	010	110	110
1	2	6	6

Thus, the octal equivalent of $(2B6)_{16}$ is $(1266)_8$.

2.9 CODING SCHEMES

In today's technology, the binary number system is used by the computer system to represent the data in the computer understandable format. Numeric data (0, 1, 2,..., 9) is not the only form of data, which is handled by the computer. The alphanumeric data (it is a string of symbols of the letters A, B, C,..., Z or the digits 0, 1, 2,..., 9) and some special characters such as =, –, +, *, /, (,), etc., are also processed by the computer. There are a lot of ways to represent the numeric, alphabetic and special characters in the computer's internal storage area. In computers, the code is made up of fixed-size groups of binary positions. Each binary position in a group is assigned a specific value, for example, 8, 4, 2 or 1. In this way, every character can be represented by a combination of bits that is unique. Moreover, data can also be arranged in a way that is very simple and easy to

decode or transmitted with varying degrees of redundancy for error detection and correction. There are many coding schemes available for representing characters. The most commonly used coding systems are *ASCII* code and *Unicode*.

American Standard Code for Information Interchange Code: The standard binary code for the alphanumeric characters is the ASCII. This code was originally designed as a 7-bit code. Several computer manufacturers cooperated to develop this code for transmitting and processing data. They made use of all eight bits providing 256 symbols. Nevertheless, IBM had not changed the original set of 128 codes so that the original instructions and data could still work with the new character set. The ASCII is commonly used in the transmission of data through data communication and is used almost exclusively to represent the data internally in the microcomputers. In the ASCII, the upper-case letters are assigned codes beginning with the hexadecimal value 41 and continuing sequentially through the hexadecimal value 5A and lower-case letters are assigned hexadecimal values of 61 through 7A. The decimal values 1–9 are assigned the zone code 0011 in the ASCII. Table 2.6 of the ASCII coding chart shows upper-case and lower-case alphabetic characters and numeric digits 0–9. The standard ASCII code defines 128-character codes (0–127) of which the first 32 are control codes (non-printable) and other 96 are representable characters.

Table 2.6 ASCII Coding Chart

Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char
0	00	Null	32	20	*Space*	64	40	@	96	60	`
1	01	Start of heading	33	21	!	65	41	A	97	61	a
2	02	Start of text	34	22	"	66	42	B	98	62	b
3	03	End of text	35	23	#	67	43	C	99	63	c
4	04	End of transmit	36	24	$	68	44	D	100	64	d
5	05	Enquiry	37	25	%	69	45	E	101	65	e
6	06	Acknowledge	38	26	&	70	46	F	102	66	f
7	07	Audible bell	39	27	'	71	47	G	103	67	g
8	08	Backspace	40	28	(	72	48	H	104	68	h
9	09	Horizontal tab	41	29	)	73	49	I	105	69	i
10	0A	Line feed	42	2A	*	74	4A	J	106	6A	j
11	0B	Vertical tab	43	2B	+	75	4B	K	107	6B	k
12	0C	Form feed	44	2C	,	76	4C	L	108	6C	l
13	0D	Carriage return	45	2D	-	77	4D	M	109	6D	m
14	0E	Shift out	46	2E	.	78	4E	N	110	6E	n
15	0F	Shift in	47	2F	/	79	4F	O	111	6F	o
16	10	Data link escape	48	30	0	80	50	P	112	70	p
17	11	Device control 1	49	31	1	81	51	Q	113	71	q
18	12	Device control 2	50	32	2	82	52	R	114	72	r
19	13	Device control 3	51	33	3	83	53	S	115	73	s
20	14	Device control 4	52	34	4	84	54	T	116	74	t
21	15	Neg acknowledge	53	35	5	85	55	U	117	75	u
22	16	Synchronous idle	54	36	6	86	56	V	118	76	v
23	17	End trans. block	55	37	7	87	57	W	119	77	w
24	18	Cancel	56	38	8	88	58	X	120	78	x
25	19	End of medium	57	39	9	89	59	Y	121	79	y
26	1A	Substitution	58	3A	:	90	5A	Z	122	7A	z
27	1B	Escape	59	3B	;	91	5B	[	123	7B	{
28	1C	File separator	60	3C	<	92	5C	\	124	7C	|
29	1D	Group separator	61	3D	=	93	5D	]	125	7D	}
30	1E	Record separator	62	3E	>	94	5E	^	126	7E	~
31	1F	Unit separator	63	3F	?	95	5F	_	127	7F	□

Unicode: Before the invention of the Unicode, hundreds of different encoding systems were used. No single encoding system was large enough to represent vast number of characters. Even for a simple language like English, no single encoding system was adequate for all

the letters, punctuation and technical symbols in common use. Moreover, these encoding systems also conflicted with one another. Therefore, to overcome these issues, the Unicode encoding system was developed.

THINGS TO REMEMBER

UTF Formats

Unicode characters are divided into two basic transformation formats, namely Unicode Transformations Format-8 (UTF-8) and UTF-16. UTF-8 is a lossless encoding of Unicode characters. This format encodes each Unicode character as a variable number of octets (1–4), where the number of octets required depends on the number assigned to the character.

In the UTF-16 encoding, characters are symbolized using either one or two unsigned 16-bit integers, depending on the character value for storage or transmission through data networks.

The Unicode is a universal character-encoding standard for the interpretation of text for computer processing. It offers a constant way of encoding multilingual plain text. The standard provides the capacity to encode all the characters used in different languages around the world. To keep character coding simple and proficient, the Unicode standard allocates a unique numeric value and name to each character. The objective behind the Unicode was to use a single 16-bit encoding that provides code points for more than 65,000 characters and to support characters in major languages of the world. The Unicode standard incorporates punctuation marks,

mathematical symbols, technical symbols, arrows and many more characters.

Even though a character may be used in more than one language, it is defined only once in the Unicode; for example, the Latin capital letter “A” is mapped once, even though it is used in English, German and Japanese. On the other hand, it was decided that the Cyrillic capital letter “A” is a different character from the Latin capital letter “A”, even though the two letters look a lot like each other. The reason behind decisions like these is interesting to linguists, but usually not important to programmers.

LET US SUMMARIZE

1. The *CPU* or the *processor* is a chip inside the system plugged onto the motherboard and controls all internal and external devices as well as performs arithmetic and logic operations. The unit consists of three main subsystems, the CU, the ALU and the registers.
2. The functions of a processor include carrying out arithmetic and logic functions, controlling the use of the main memory to store data and instructions, and controlling the sequence of operations and all the parts of the computer system.
3. The set of wires used for interconnection of various units is known as the *system bus*. The system bus is divided into three logical units, namely the address bus, the data bus and the control bus.
4. The *registers* are special-purpose, high-speed temporary memory units. They hold various types of information pertaining to data, instructions, addresses and the intermediate results of calculations. Some of the important registers within the CPU are the PC, IR, MAR, MBR, ACC and DR.
5. The ALU carries out arithmetic and logical operations on the data made available to it. The ALU contains the circuitry that is responsible for performing the actual computing and carrying out the arithmetic calculations. On the other hand, the importance of the ALU is to facilitate the CPU to make logical operations based on the instructions provided to it.
6. The CPU controls the I/O devices and transfers data to and from the primary storage. It repeats a set of four basic operations: fetching, decoding, executing and storing.
7. The main memory unit is a collection of registers, logically integrated to the processor but physically separate from the processor. It is implemented by two types of memory technologies, namely *RAM* and *ROM*.
8. A cache memory, sometimes called a *cache store*, is a portion of the memory made up of the high-speed static RAM (SRAM) instead of the slower and comparatively cheaper dynamic RAM (DRAM) used for the main memory. It is very expensive and

generally incorporated in the processor, where valuable data and program segments are kept.

9. In the processor to memory communication, the whole process can be divided into two steps, namely information transfer from the memory to the processor and writing information in the memory.
10. In the processor to I/O devices communication, the I/O units are connected to the computer through the system bus. In addition, it buffers the flow of data from the device to the processor and vice versa.
11. The *instruction cycle* is sequence of well-defined instructions in the form of programs, which consist of the following steps: fetch cycle (fetching the instruction from the memory), decode cycle (decoding the instruction), execute cycle (executing the instruction) and store cycle (storing the result back to the memory).
12. The *instruction set* of a processor is a limited set of basic operations built into the processor. Based upon the instruction sets, there are two common types of architectures: CISC and RISC.
13. A *system case* or *cabinet* is a metal and plastic box that houses the main components of the computer. It protects all the vital electronic components against heat, light and temperature.
14. A *power supply* or *SMPS* is a transformer and voltage control device in the computer that furnishes power to all the electronic components by converting incoming AC supply into the low-voltage DC supply. When a computer is turned on, the power supply allows the converted electricity to travel to other components inside the computer.
15. A *motherboard* is a large multilayered printed circuit board inside a computer. It contains the CPU, BIOS ROM chip, CMOS setup information and is equipped with expansion slots for installing different adapter cards, the connector for the keyboard, slots for the system's RAM, etc.
16. *Ports* and *interfaces* are a generic name for the various sockets, found at the back of the computer, using which external devices are connected to the computer's motherboard.
17. An *expansion card* is a circuit board that provides additional capabilities to the computer system. These cards are made up of large-scale integrated circuit components installed on it. The cards are plugged into the expansion sockets present in the computer's motherboard to give the computer an added functionality.
18. *Ribbon cables* are wide, flat and insulated cables, which are flexible to fit into areas with little space. These cables connect the hard drive, floppy drive and CD-ROM drive to the connectors on the motherboard and control the drives by getting and sending data from and to them.
19. A *number system* defines a set of values used to represent a “quantity”. These are of two types: non-positional and positional number systems.
20. In a non-positional number system, special symbols or characters are used to indicate the values. It is very difficult to perform arithmetic operations with such a number system, as it has no symbol for zero.
21. In a positional number system, the value of each digit in a number is not only defined by the symbols, but also defined by the symbol's position. These symbols are called digits.

22. The positional number system, which is being used nowadays, is called as the *decimal number system*. Apart from this number system, there are some other positional number systems, such as the binary number system, octal number system and hexadecimal number system.
23. The *base* or *radix* of the number system tells about the number of symbols or digits used in the system. The base of the decimal number system is 10, of binary number system is 2, of octal number system is 8 and of hexadecimal number system is 16.
24. In computers, the coding scheme is made up of fixed-size groups of binary positions, where each binary position in a group is assigned a specific value. Some of the most commonly used coding systems are the *ASCII* code and *Unicode*.

EXERCISES

Fill in the Blanks

1. The ..... controls the I/O devices and transfer of data to and from the primary storage.
2. ..... enables the processor to access data quickly whenever they are needed.
3. The ..... informs the CPU about the location of the data residing in the memory.
4. The ..... is responsible for recognizing which operation the bit pattern represents.
5. Decoding and executing instructions are complicated and time-consuming in the ..... processor architecture.
6. A metal and plastic box that houses the main components of the computer and protects against the heat, light, temperature and other means is called .....
7. ..... is a large multilayered printed circuit board inside a computer that contains CPU, BIOS ROM chip and the CMOS setup information
8. ..... is a general-purpose communications port through which data are passed slowly over long distances.
9. ..... is an expansion card that allows two computers to communicate over ordinary phone lines.
10. The number systems can be categorized into two broad categories: ..... and .....

Multiple-choice Questions

1. The processor is a ..... chip plugged onto the motherboard in a computer system.
    1. LSI
    2. VLSI
    3. ULSI
    4. XLSI
2. The set of wires, which carry information in a controlled manner, is called .....
    1. System bus
    2. Public bus
    3. Private bus
    4. None of these

3. The ALU works on the instructions and data held in the .....
    1. Notebook
    2. Registers
    3. Copy Pad
    4. I/O devices
4. A register that keeps track of the next instruction to be executed is called a .....
    1. Program Counter
    2. Instruction register
    3. Accumulator
    4. Data register
5. The fastest memory in a computer system is .....
    1. ROM
    2. RAM
    3. Cache
    4. None of these
6. In the sequence of events that takes place in an instruction cycle, the first cycle is .....
    1. Store cycle
    2. Execute cycle
    3. Fetch cycle
    4. Decode cycle
7. ..... is a type of the processor architecture that utilizes a small, highly optimized set of instructions.
    1. CISC
    2. RISC
    3. VISC
    4. LISC
8. RISC processors are ideal for embedded applications, such as mobile phones and PDAs, because .....
    1. They are smaller in size and consume less power.
    2. They are large in size and consume less power.
    3. They are smaller in size and consume more power.
    4. They are larger in size and consume large amount of power.
9. The concept of the CISC architecture is to accomplish the task in .....
    1. As longer lines of code as possible
    2. As few lines of code as possible
    3. Both (a) and (b)
    4. None of these
10. An expansion card that allows the computer to output sound through connected speakers is .....
    1. Video Card
    2. Network Interface Card
    3. PC Card
    4. Sound Card

State True or False

1. A processor only operates on binary data, that is, data composed of 1s and 0s.
2. The address bus informs the CPU about the location of the data residing in the memory.
3. Registers are special-purpose, high-speed permanent memory units.

4. A program counter is a general-purpose register used for storing temporary results and results produced by the ALU.
5. The L2 cache is usually 64 KB–2 MB in size.
6. A power supply is a transformer that converts power into the low-voltage AC supply.
7. A USB port is a plug-and-play hardware interface that supports maximum bandwidth of 12 Mbit per second.
8. The PC card is a removable device, approximately the size of a credit card.
9. SIMM chips are faster than DIMM chips.
10. The Unicode is a character-encoding standard developed by the Unicode Consortium.

Descriptive Questions

1. What do you understand by the CPU? Describe in details the various units of the CPU.
2. What are registers in a CPU? Name five registers with their functions.
3. Write a detailed note on the instruction cycle describing the various steps involved.
4. Compare the RISC and CISC architecture briefly. Also, discuss their advantages and disadvantages.
5. What is a system bus? Name the various units of the system bus.
6. With an appropriate example, explain the conversion of:
    1. Binary to octal and vice versa
    2. Binary to hexadecimal and vice versa
    3. Octal to hexadecimal and vice versa
7. What are expansion cards? How many types of expansion cards can be used in a computer system?
8. Write short notes on:
    1. Motherboard
    2. Power supply
    3. Ports
    4. Ribbon cables
9. Perform the following conversions:
    1. Convert binary 00011011 to decimal
    2. Convert decimal 278 to binary
    3. Convert decimal 0.625 to binary
10. What is a radix or base of the system? With the help of this system, brief the various types of number systems.

ANSWERS

Fill in the Blanks

1. CU
2. Cache
3. Address bus
4. Decode cycle
5. CISC
6. System case
7. Motherboard
8. Serial port
9. Modem

10. Non-positional number system, positional number system

Multiple-choice Questions

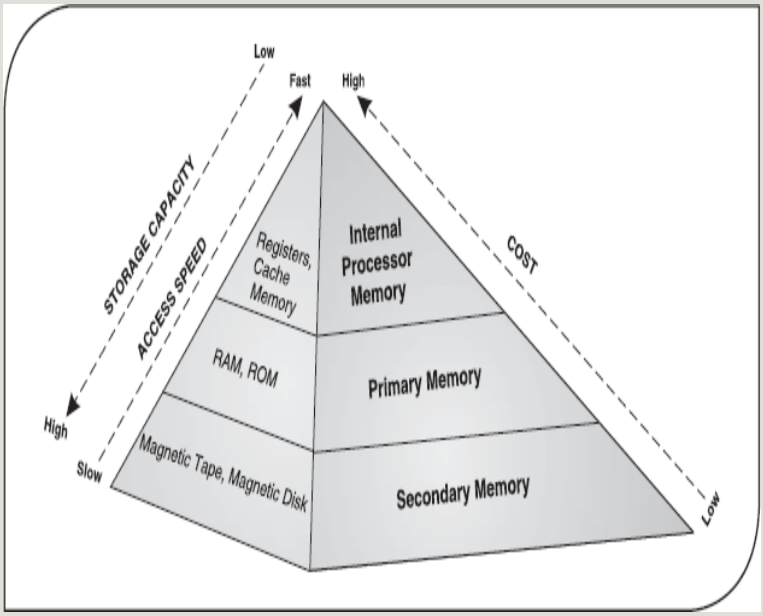
1. (b)
2. (a)
3. (b)
4. (a)
5. (c)
6. (c)
7. (b)
8. (a)
9. (b)
10. (d)

State True or False

1. True
2. True
3. False
4. False
5. True
6. False
7. True
8. True
9. False
10. True

Chapter 3

Computer Memory and Storage



This chapter introduces the reader to memory and storage present inside a computer system. A computer memory refers to the electronic holding place for instructions and data where the computer's microprocessor can reach quickly. Computer storage refers to the permanent computer memory that stores all the data files and instructions even after the computer system is turned off. The chapter begins with a brief outlook on memory and its representation. Next, memory hierarchy along with RAM and ROM is studied. Furthermore, classification of storage devices, magnetic tapes, magnetic disks, and optical disks are examined in detail. Finally, the chapter concludes by

providing an outlook on mass storage devices along with their types.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The three fundamental types of memory in a computer system
- How random access memory (RAM) stores and keeps track of the data currently being processed in a computer
- Read only memory (ROM), a special chip that stores start-up instructions to help the computer in the booting process
- Memory, the most essential part of computer processing, and how the CPU uses memory
- Different types of secondary storage devices
- How magnetic tape holds data recording in the form of tiny segments of magnetized and demagnetized part on the surface of the tape
- Magnetic disk, the most popular and widely used storage medium for direct access secondary storage
- Various types of magnetic disks and how read/write operation is performed on these disks
- Optical disk, a new storage medium that is gaining popularity
- Various types of optical disks—CD-ROM and DVD-ROM
- Magneto-optical storage devices, which include the basic principle of both magnetic and optical storage
- How mass storage device provides vast amount of storage capacity, and its different types

3.1 INTRODUCTION

Computers are used to perform various tasks in science, engineering, business, education, entertainment and many other fields. They work at high speed, can handle large volumes of data with great accuracy and have the

ability to carry out a specified sequence of operations without human intervention. The CPU handles the processing of data and after processing, presents the results with the help of output devices. However, the CPU requires memory to process the data, hold the intermediate results and to store the output. *Computer memory* refers to the electronic holding place for instructions and data where the processor can reach quickly. It can be classified into two broad categories: *primary memory* (to process the data and hold the intermediate results) and *secondary memory* (to store the output).

The primary memory allows the computer to store data for immediate manipulation and to keep track of what is currently being processed. The major limitation of this type of memory is that it is volatile. It means that when the power is turned off, the contents of primary memory are lost forever. Hence, to store the data permanently, a computer requires some nonvolatile storage medium like a hard disk. This kind of storage is known as *secondary memory*. Such memories store all the data (files) and instructions (computer programs) even after the power is turned off. The secondary storage devices have a larger storage capacity; they are less expensive as compared to primary storage devices, but slow in comparison.

Note: *When we talk about memory, we generally refer to the primary memory only, and when we talk about storage, we refer to secondary memory.*

3.1.1 Memory Representation

All quantities, physical or otherwise, are measured in units. For example, length is measured in metres and mass in grams. Likewise, for measuring computer memory, a standard unit is required. Digital computers work on only two states: ON (1) and OFF (0). These two values are represented by two different voltages within the circuit. For example, 0 volt represents a false value

(0), and +5 volt represents a true value (1). Each of these values (either 0 or 1) is called a *binary digit* or *bit* and can be considered a symbol for a piece of information. Although the smallest unit of data that a computer can deal with is a bit, computers generally do not deal with a single bit. Instead, they deal with a group of eight bits, which is referred to as a byte. A byte can have 256 different bit patterns, and thus can represent 256 different symbols. Various units used to measure computer memory are as follows:

- **Bit:** It is the smallest unit of data on a machine and a single bit can hold only one of two values: 0 or 1. Bit is represented by a lower case b.
- **Byte:** A unit of eight bits is known as a byte. Hence, a byte is able to contain any binary number between 00000000 and 11111111. It is represented by an upper case B.
- **Kilobyte:** In a decimal system, kilo stands for 1000, but in a binary system, kilo refers to 1024. Therefore, a kilobyte is equal to 1024 bytes. It is usually represented as KB.
- **Megabyte:** It comprises 1024 kilobytes, or 1,048,576 bytes. However, since this number is hard to remember, a megabyte can be thought of as a million bytes. Megabyte is the standard unit of measurement for RAM and is represented as MB.
- **Gigabyte:** It consists of 1024 megabytes (1,073,741,824 bytes). It is the standard unit of measurement for hard disks and is often represented as GB.
- **Terabyte:** It refers to 1024 gigabytes. Often represented as TB, terabyte memory is usually associated with super computers only.

Note: *In modern computers, groupings of bytes (usually 2 or 4), called computer words can represent larger "chunks" of information.*

3.2 MEMORY HIERARCHY

The processor is the "brain" of the computer where all the essential computing takes place. But unlike a human brain, a computer processor has very limited memory. Thus, it has to rely on other kinds of memories to hold data and instructions and to store results. The memory in a computer system is of three fundamental types:

- **Internal Processor Memory:** This memory is placed within the CPU (processor) or is attached to a special fast bus. Internal memory usually includes cache memory and special registers, both of which can be directly accessed by the processor. This memory is used for temporary storage of data and instructions on which the CPU is currently working. Processor memory is the fastest among all the

memories but is the most expensive also. Therefore, a very diminutive part of internal processor memory is used in the computer system. It is generally used to compensate for the speed gap between the primary memory and the processor.

- **Primary Memory:** Random access memory (RAM) and read only memory (ROM) fall under the category of the primary memory, also known as *main memory*. Every computer comes with a small amount of ROM, which contains the boot firmware (called *BIOS*). This holds enough information to enable the computer to check its hardware and load its operating system into its RAM at the time of system booting. RAM is the place where the computer temporarily stores its operating system, application programs and current data so that the computer's processor can reach them quickly and easily. It is volatile in nature, that is, when the power is switched off, the data in this memory are lost. Unlike RAM, ROM is non-volatile. Even when the computer is switched off, the contents of the ROM remain available.
- **Secondary Memory:** Also known as *auxiliary memory*, secondary memory provides backup storage for instructions (computer programs) and data. The most commonly used secondary storage devices are magnetic disk and magnetic tapes. These are the least expensive and also have much larger storage capacity than the primary memory. The instructions and data stored on secondary storage devices are permanent in nature. They can only be removed if the user wants it so or if the device is destroyed. Secondary memory can also be used as *overflow memory* (also known as *virtual memory*), when the capacity of the main memory is surpassed. Note that unlike processor memory and main memory, secondary memory is not directly accessible to the processor. Firstly, the data and instructions from the secondary memory have to be shifted to the main memory and then to the processor.

Figure 3.1 illustrates the memory hierarchy. The CPU accesses memory according to a distinct hierarchy. When the data come from permanent storage (for example, hard disk), they first go in RAM. The reason behind it is that if the CPU has to access the hard disk constantly to retrieve every piece of required data, it would operate very slowly. When the data are kept in primary memory, the CPU can access them more quickly. Subsequently, the CPU stores the required pieces of data and instructions in processor memory (cache and registers) to process the data.

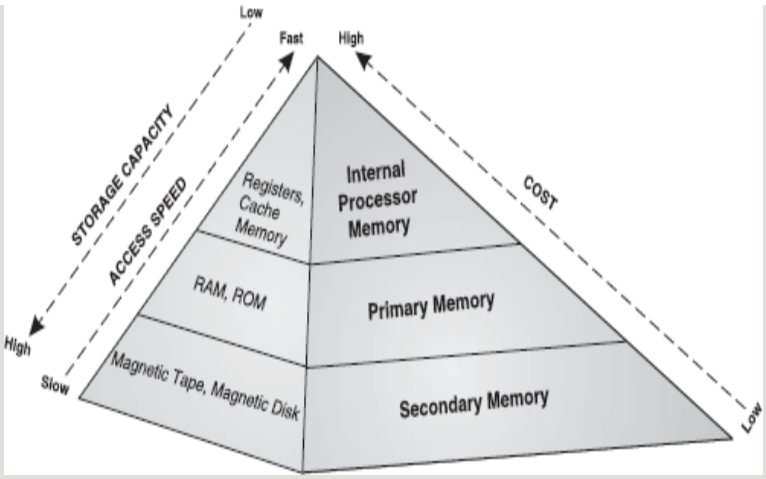


Figure 3.1 Memory Hierarchy

3.3 RANDOM ACCESS MEMORY

RAM is like the computer's scratch pad. It allows the computer to store data for immediate manipulation and to keep track of what is currently being processed. It is the place in a computer where the operating system, application programs and data in current use are kept so that they can be accessed quickly by the computer's processor. RAM is much faster to read from and write to than the other kinds of storage in a computer, like the hard disk or floppy disk. However, the data in RAM stay there only as long as the computer is running. When the computer is turned off, RAM loses all its contents. When the computer is turned on again, the operating system and other files are once again loaded into RAM. When an application program is started, the computer loads it into RAM and does all the processing there. This allows the computer to run the application faster. Any new information that is created is kept in RAM and since RAM is volatile in nature, one needs to continuously save the new information to the hard disk.

Let us take a simple example of why RAM is used by the computer. Whenever a user enters a command from the

keyboard, the CPU interprets the command and instructs the hard disk to “load” the command or program into main memory (see Figure 3.2). Once the data are loaded into memory, the CPU is able to access them much quickly. The reason behind this is that the main memory is much faster than secondary memory. The process of putting things that the CPU needs in a single place from where it can get them more quickly is similar to placing various documents, which the user needs, into a single file folder. By doing so, the user finds all the required files handy and avoids searching in several places every time he needs them.

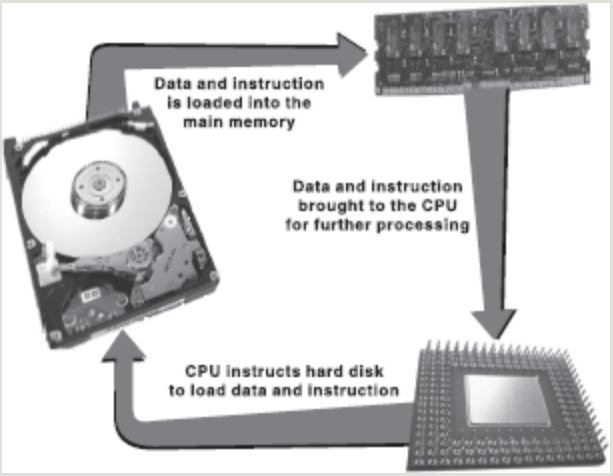


Figure 3.2 Interaction of Memory and Storage with Processor

3.3.1 Types of RAM

RAM is of two types:

- **Static RAM (SRAM):** The word “static” indicates that the memory retains its contents as long as power is being supplied. However, as soon as the power goes down, the data are lost. This makes SRAM a volatile memory as opposed to ROM. SRAM does not need to be “refreshed” (pulse of current through all the memory cells) periodically. It is very fast but much more expensive than DRAM (Dynamic RAM). SRAM is often used as cache memory due to its high speed.

- **Dynamic RAM (DRAM):** It is named so because it is very unstable. The data continue to move in and out of the memory as long as power is available. Unlike SRAM, DRAM must be continually refreshed in order to maintain the data. This is done by placing the memory on a refresh circuit that rewrites the data several hundred times per second. DRAM is used for most system memory because it is inexpensive and small.

The primary difference between SRAM and DRAM is the life of the data they store. SRAM retains its contents as long as electrical power is supplied to the chip. If the power is turned off, its contents are lost. On the other hand, DRAM must be continuously refreshed after about every 15 microseconds. This is true even when power is supplied constantly. SRAM chips are not as dense as DRAM chips, that is, the total number of cells in the SRAM chip is less than that on DRAM chip. SRAM is beneficial because it is fast, has low latency (the time lag between a request and the action being performed), and does not need to be refreshed. However, it is large and expensive, requires more power to operate, and produces a lot of heat. DRAM is simple, small, and space efficient. It may be slower and may have a longer latency than SRAM, but it is still very useful. Typical access time of SRAM is 25 nanoseconds while of DRAM 60 nanoseconds.

SRAM is useful for low amount of memory. Anything over 4 MB is very bulky. SRAM is good for internal memory in processors, and cache, but DRAM is best for the system's main memory. DRAM is used where its small size and power efficiency outweigh its slowness as compared to SRAM. SRAM is less dense than DRAM (fewer bits per unit area) and is, therefore, not suitable for high-capacity, low-cost-per-megabyte applications. The power consumption of SRAM varies widely depending on its speed. Fast SRAM is much more power-hungry than DRAM and some ICs can consume power of the order of a watt at full speed. Slow SRAM can have very low power consumption in the region of a microwatt. Currently, the technology does not exist to

produce small SRAMs so that they can replace DRAMs. Thus, DRAM is still used in computers.

FACT FILE

RAM and ROM

RAM is also called read/write memory because, unlike ROM that does not allow any write operation, RAM allows the CPU to read as well as write data and instructions.

3.4 READ ONLY MEMORY

Just as a human being needs instructions from the brain to perform actions in a certain event, a computer also needs special instructions every time it is started. This is required because during the start up operation, the main memory of the computer is empty due to its volatile property so there have to be some instructions (special boot programs) stored in a special chip that could enable the computer system to perform start up operations and transfer the control to the operating system. This special chip, where the start up instructions are stored, is called *ROM*. It is non-volatile in nature, that is, its contents are not lost when the power is switched off. The data and instructions stored in ROM can only be read and used but cannot be altered, thereby making ROM much safer and secure than RAM. ROM chips are used not only in the computer but also in other electronic items like washing machines and microwave ovens.

Generally, designers program ROM chips at the time of manufacturing circuits. Burning appropriate electronic fuses to form patterns of binary information does the programming. These patterns of binary information are meant for specific configurations, which is why different

categories of computers are meant for performing different tasks. For example, a micro program called *system boot program* contains a series of start-up instructions to check for the hardware, that is, I/O devices, memory and operating system in the memory. These programs deal with low-level machine functions and are alternate for additional hardware requirement. ROM performs the necessary BIOS (basic input output system) function to start the system and then transfers the control over to the operating system.

ROM can have data and instructions written into it only one time. Once a ROM chip is programmed, it cannot be reprogrammed or rewritten. If it is erroneous, or the data need to be reorganized, one has to replace it with the new chip. Thus, the programming of ROM chips should be perfect, having all the required data at the time of its manufacturing. Note that in some instances, ROM can be changed using certain tools. For example, flash ROM (a type of ROM) is non-volatile memory that occasionally can be changed, such as when a BIOS chip must be updated. The ROM chips consume very little power, are extremely reliable, and in the case of most small electronic devices, contain all the necessary programming to control the device.

3.4.1 Types of ROM

Memories in the ROM family are distinguished by the methods used to write data on them and the number of times they can be rewritten. This classification reflects the evolution of ROM devices from “hard-wired” to programmable to erasable-and-programmable. One common feature of all these devices is their ability to retain data and programs even during a power failure. ROMs come in following varieties:

- **Masked ROM:** The very first ROMs, known as masked ROMs, were hard-wired devices that contained a pre-programmed set of data or instructions. The contents of such ROMs had to be specified before

chip production so the actual data could be used to arrange the transistors inside the chip.

- **Programmable ROM (PROM):** Creating a ROM chip from scratch is a time-consuming and an expensive process. For this reason, developers created a type of ROM known as programmable read only memory (PROM), which can be programmed. Blank PROM chips can be bought economically and coded by the users with the help of a special device known as *PROM-programmer*. However, once a PROM has been programmed, its contents can never be changed. As a result, PROM is also known as *one-time programmable* (*OTP*) device. Like other ROMs, PROM is also non-volatile. However, it is more fragile than other ROMs as a jolt of static electricity can easily cause the fuses in the PROM to burn out, thus changing the bit pattern from 1 to 0. Nevertheless, blank PROMs are economical and are great for prototyping the data for a ROM before committing to the costly ROM fabrication process.
- **Erasable Programmable ROM (EPROM):** An EPROM is programmed in exactly the same manner as a PROM. However, unlike PROM, an EPROM can be erased and reprogrammed repeatedly. It can be erased by simply exposing the device to a strong source of ultraviolet light for a certain amount of time. Note that an EPROM eraser is not selective; it will erase the entire EPROM. Although EPROM is more expensive than PROM, its ability to be reprogrammed makes it more useful.
- **Electrically Erasable Programmable ROM (EEPROM):** This type of ROM can be erased by an electrical charge and then written to by using slightly higher-than-normal voltage. EEPROM can be erased one byte at a time, rather than erasing the entire chip with ultraviolet light. Hence, the process of reprogramming is flexible, but slow. Also, changing the contents does not require any additional committed equipment. As these chips can be changed without opening a casing, they are often used to store programmable instructions in devices like printers.
- **Flash ROM:** A flash ROM also called *flash BIOS or flash memory*, is a type of constantly powered non-volatile memory that can be erased and reprogrammed in blocks. It is a variation of EEPROM, which, unlike flash memory, is erased and rewritten at the byte level. Flash memory is often used to hold the control code such as the BIOS in a personal computer. When BIOS needs to be changed or rewritten, the flash memory can be written in block (rather than byte) sizes, thus making it easier to update. Flash memory gets its name because the microchip is organized so that a section of memory cells are erased in a single action or “flash”. Flash memory is used in digital cellular phones, digital cameras and other devices.

3.5 RAM, ROM AND CPU INTERACTION

Picture a bulletin board under the glass at the back of a classroom. One way to think of ROM is that it is similar to the hard-copy notes placed under the glass. At the end of the day, they remain unchanged. The next day, the notes are exactly the way they were the day before. Students are able to only read them. RAM, on the other

hand, can be thought of as a blackboard, which starts out blank and during the day, information is written on it, read from it, and even erased from it. When something is erased, new data can then be written on the same place on the board. When students go home at the end of the day, the blackboard is washed clean and whatever data was on it goes away forever. This is what happens when a computer's power is turned off; RAM no longer has the electrical current available to sustain the data in its memory cells. ROM is more like your long-term memory; the things you remember from your past. When you wake up in the morning, you get ready for school/office and know the address of your destination. Similarly, when computer “wakes up”, it searches for start-up routines from ROM BIOS and then hands over the control to the operating system to function properly.

The most essential part of computer processing is the memory. From the moment the computer is turned ON and until it is shut down, the CPU constantly uses memory. A typical scenario is listed below and illustrated in Figure 3.3:

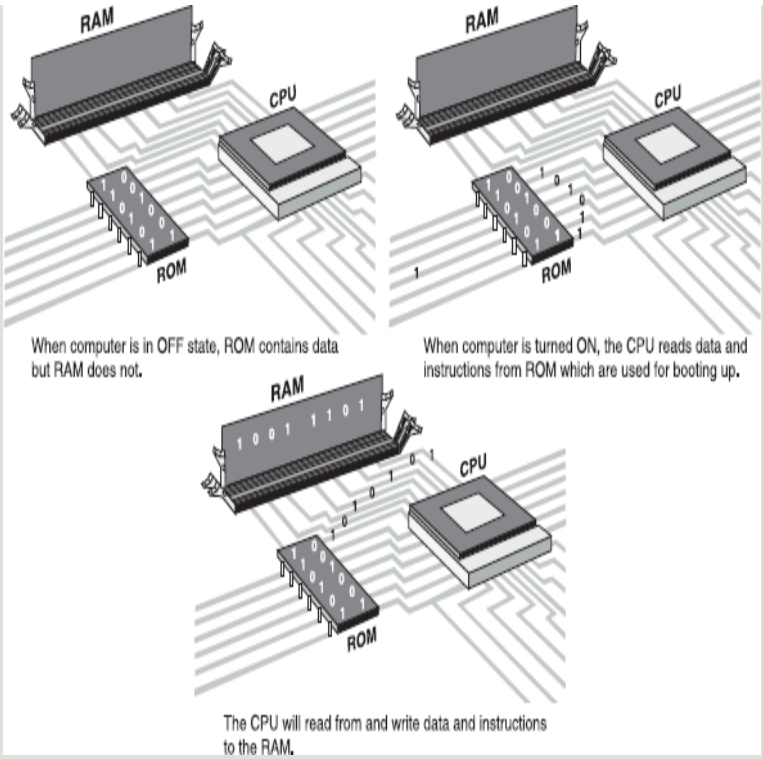


Figure 3.3 RAM, ROM and CPU Interaction

Step 1	The computer is switched ON.
Step 2	CPU loads data and instructions from ROM and checks whether all the major components like processor and hard disk are functioning properly.
Step 3	CPU loads BIOS (basic input/output system) from ROM to determine the machine's fundamental configuration and environment. The information stored in ROM BIOS chip determines what peripherals the system can support.
Step 4	CPU loads the operating system from the secondary storage (hard disk) into RAM. This allows the CPU immediate access to the operating system, which enhances the performance and functionality of the overall system.

Step 5	When an application is opened, it is loaded into RAM and any file that is opened for use in that application is also loaded into RAM.
Step 6	After processing, when the user saves the file and closes the respective application, the file is written to the specified location on the secondary storage device. After that the file(s) and the application are “flushed out” from the RAM.

Every time something is loaded or opened, it is placed into RAM so that the CPU can access that information more easily and promptly. The CPU requests the required data from RAM, processes it, and writes new data back to RAM in a continuous cycle. In most computers, this shuffling of data between the CPU and RAM happens millions of times every second. When the application is closed, the application and any other accompanying files are usually erased from RAM to make space for the new data.

3.6 TYPES OF SECONDARY STORAGE DEVICES

Secondary storage devices facilitate storing of data and instructions permanently. The data stored on a secondary storage device can be accessed depending upon how it is stored on the device. Primarily, there are two methods of accessing data from the secondary storage devices (see Figure 3.4):

Figure 3.4 Sequential and Direct Access

- **Sequential Access:** Sequential access means the computer system must search the storage device from the beginning until the desired data is found. The most common sequential access storage device is magnetic tape where data is stored and processed sequentially. Suppose, a tape contains information regarding employees of an organization. Now, to look for employee number 100's information, the computer will have to start with employee number 1 and then go past 2, 3 and so on, until it finally comes to 100. The sequential access method is quite simple than other methods but searching for data is slow.
- **Direct Access:** Direct access, also known as *random access*, means that the computer can go directly to the location, where the data that the user wants, are stored. The most common direct access storage devices are magnetic disk and optical disk. In these devices, the data are stored as sequentially numbered blocks. Thus, one can access block 12, then access block 78, then block 2 and so on. The direct access method is ideal for applications like airline reservation systems or computer-based directory assistance operations. In these cases, there is no fixed pattern of requests for data.

Based on the access method, secondary storage devices can be classified as shown in Figure 3.5.

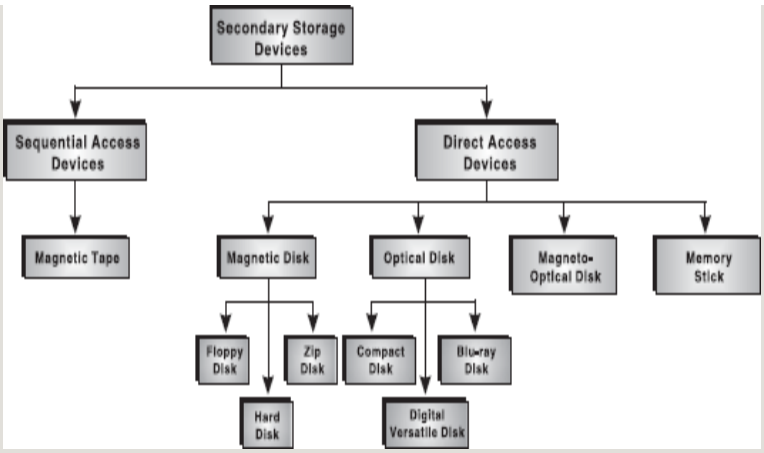


Figure 3.5 Classification of Secondary Storage Devices

3.7 MAGNETIC TAPE

Magnetic tape appears similar to the tape used in music cassettes. It is a plastic tape with magnetic coating on it. The data is stored in the form of tiny segments of magnetized and demagnetized portions on the surface of the material. The magnetized portion of the surface refers to the bit value “1” whereas the demagnetized portion refers to the bit value “0”. Magnetic tapes are available in different sizes, but the major difference between different magnetic tape units is the speed at which the tape is moved past the read/write head and the tape's recording density. The amount of data or the number of binary digits that can be stored on a linear inch of tape is the recording density of the tape.

Magnetic tapes are very durable and can be erased as well as reused. They are an inexpensive and reliable storage medium for organizing archives and taking backups. However, they are not suitable for data files that need to be revised or updated often because data on them are stored in a sequential manner and the user needs to advance or rewind the tape every time to the position where the requested data starts. Tapes are also slow due to the nature of the media. If the tape stretches

too much, then it will render itself unusable for data storage and may result in data loss. The tape now has a limited role because disk has proved to be a superior storage medium. Today, the primary role of the tape drive is limited to backing up or duplicating the data stored on the hard disk to protect the system against loss of data during power failures or computer malfunctions.

3.7.1 Magnetic Tape Organization

The magnetic tape is divided into vertical columns (*frames*) and horizontal rows (*channels* or *tracks*) as shown in Figure 3.6. The data is stored in a string of frames with one character per frame, and each frame spans multiple tracks (usually seven or nine tracks). Thus, a single bit is stored in each track, that is, one byte per frame. The remaining track (seventh or ninth) stores the parity bit. When a byte is written to the tape, the number of 1s in the byte is counted, the parity bit is then used to make the number of 1s even (even parity) or odd (odd parity). When the tape is read again, the parity bit is checked to see if any bit has been lost. In case of odd parity, there must be an odd number of 1s represented for each character and an even number of 1s in case of even parity.

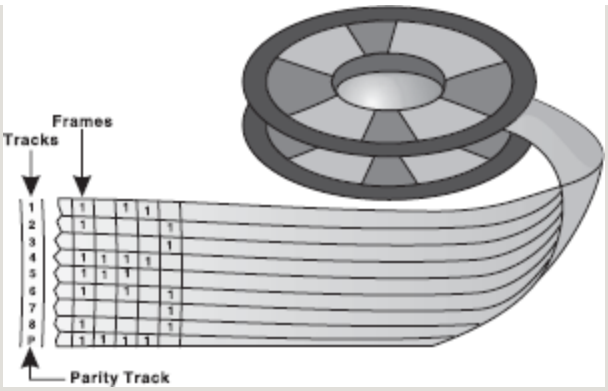


Figure 3.6 Representing Data in Magnetic Tape

Magnetic tape drive uses two reels: *supply reel* and *take-up reel*. Both reels are mounted on the hubs and the tape moves from the supply reel to the take-up reel. Figure 3.7 shows the basic tape drive mechanism. The magnetic-oxide-coated side of the tape passes directly over the read/write head assembly, thus making contact with the heads. As the tape passes under the read/write head, the data can be either read and transferred to the primary memory or read from the primary memory and written onto the tape.

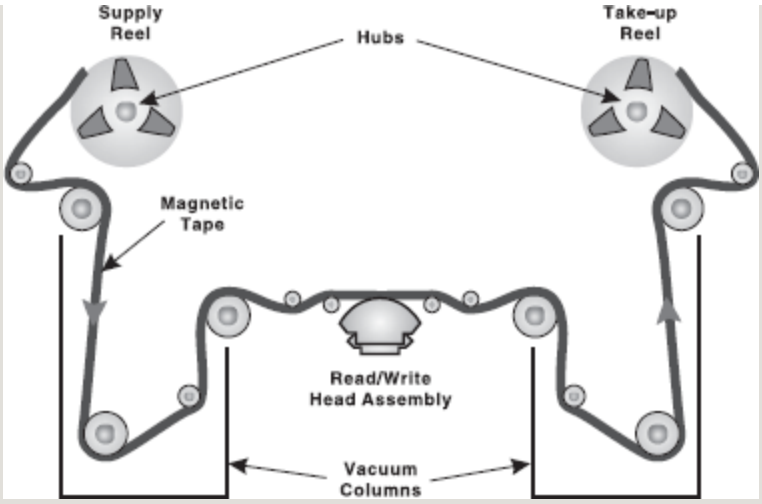


Figure 3.7 Basic Tape Drive Mechanism

A magnetic tape, as shown in Figure 3.8, is physically marked to indicate the location from where reading and writing on the tape is to begin (*BOT or beginning of tape*) and end (*EOT or end of tape*). The length of the tape between BOT and EOT is referred to as the usable recording (reading/writing) surface. BOT/EOT markers are usually made up of reflective-type of short silver strips. These markers are sensed by an arrangement of lamps and/or photodiode sensors to indicate the location from where reading and writing is to begin and end. On a magnetic tape, data are recorded in the form of blocks where each block consists of a grouping of data (known as *records*) that is written or read in a continual manner. Between these blocks, the computer automatically reserves some blank space called *inter-block gap* (*IBG*). One block may contain one or more records that are again separated by blank spaces (usually 0.5 inch) known as *inter-record gap* (*IRG*). Whenever an IRG is reached while reading data from a moving tape, the moving tape is stopped. It remains immobile until the record is processed.

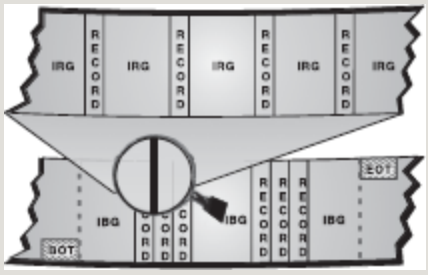


Figure 3.8 Information Format of Magnetic Tape

3.7.2 Advantages of Magnetic Tapes

- Magnetic tapes hold high data recording density thereby resulting in low cost per bit of storage.
- Magnetic tapes have virtually unlimited storage capacity because as many tapes as required can be used to store a large amount of data.
- Magnetic tapes are portable because they are compact in size, lightweight and removable. Due to these properties, they are also easy to handle and store.
- Magnetic tapes represent a very inexpensive mode of offline data storage and a large amount of data can be stored in a small storage space.

3.7.3 Disadvantages of Magnetic Tapes

- Since magnetic tapes are sequential in nature, they are not suitable in situations where data access is required in a random order. Moreover, data transmission in magnetic tapes is also slow as compared to the magnetic disks.
- Magnetic tapes should be kept in a dust-free environment and away from corrosive gases and chemicals as they can cause tape-reading errors.
- Since magnetic tapes use parity bit to check the data, the data on such devices are difficult to recover even if a minor bit error occurs.
- Magnetic tapes are not flexible as compared to other media types when file updating requires record insertion or deletion. One more drawback of magnetic tapes is that they wear out.

3.8 MAGNETIC DISK

Magnetic disks are the widely used and popular medium for direct access secondary storage. They offer high storage capacity and reliability and have the capability to access the stored data directly. A magnetic disk consists

of a plastic/metal circular plate/platter, which is coated with magnetic oxide layer. On a disk, data are represented as magnetized spots. A magnetized spot represents 1 and the absence of a magnetized spot represents 0. To read the data, the magnetized spots on the disk are converted into electrical impulses, which are then transferred to the processor. Writing data onto the disk is accomplished by converting the electrical impulses received from the processor into magnetized spots on the disk. The data in a magnetic disk can be erased and reused virtually infinitely. The disk is designed to reside in a protective case or cartridge to shield it from the dust and other external interference.

3.8.1 Storage Organization of a Magnetic Disk

The surface of a disk is divided into imaginary tracks and sectors as illustrated in Figure 3.9. *Tracks* are concentric circles where the data are stored, and are numbered from the outermost to the innermost ring, starting with zero. *Sectors* refer to the number of fixed-size areas (imaginary pie slices) that can be accessed by one of the disk drive's read/write heads, in one rotation of the disk, without the head having to change its position. An intersection of a track and a disk sector is known as *track sector*. Generally, a disk has eight or more disk sectors per track. However, different types of magnetic disks may have a different number of tracks. Today disks are marked (tracks and sectors) on both surfaces, hence they are also known as double-sided disks. Each sector is uniquely assigned a disk address before a disk drive can access a piece of data. The disk address comprises sector number, track number and surface number (if double-sided disks are used). The track sectors are grouped into a collection known as *cluster*. It refers to the basic allocation unit for storage on a disk, consisting of one or more track sectors. It is also the minimum amount of disk space used by a single file.

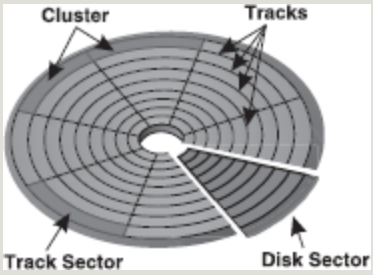


Figure 3.9 Organization of Disk Surface

Frequently, multiple disks (platters) are maintained and used together to create a large disk-storage system. Typically, two or more platters are stacked on top of each other with a common spindle, which rotates them. There is a gap between the platters, making room for the magnetic read/write head. There is a read/write head for each side of each platter and all the heads are attached to a single assembly called a *disk arm assembly*, which can move them towards the central spindle or towards the edge (see Figure 3.10). All the read/write heads are on an equal diameter track on the different platters at one time. The tracks of equal diameter on different platters form a *cylinder*.

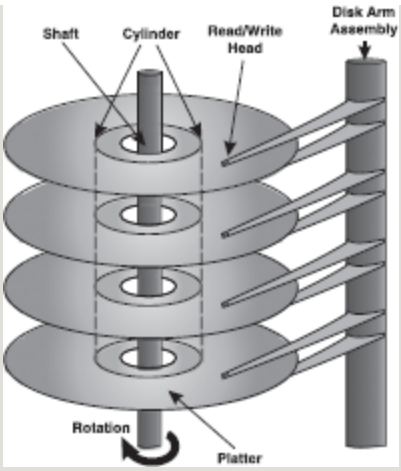


Figure 3.10 Disk Pack and Cylinder

3.8.2 Accessing Data from Magnetic Disk

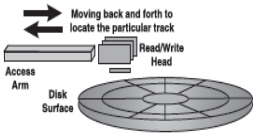
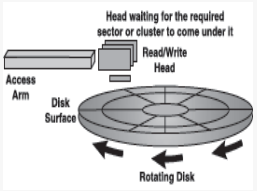
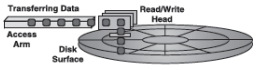
Data in a magnetic disk are recorded on the surface of the circular tracks with the help of the read/write head, which is mounted on the access arm assembly (see Figure 3.10). These heads can be in multiple numbers to access the adjacent tracks simultaneously and making a disk faster. The access arm assembly can be positioned in both inward and outward directions so that the read/write head can move on the horizontal surfaces of the disk. In case of multiple disk packs, each disk surface has its own read/write head, which works in harmony with other heads to record the data. Therefore, information is stored on the tracks, constituting a cylindrical shape through a disk pack. The process of accessing data comprises three steps:

1. **Seek:** As soon as the disk unit receives the read/write command, the read/write heads are positioned on the specific track on the disk platter. The time taken in doing so is known as *seek time*. It is the average time required to move the heads to the desired track on the disk. Seek times of modern disks may range between 2 and 15 milliseconds but the seek time of most common disks is 9 milliseconds.
2. **Rotate:** Once the heads are positioned on the desired track, the head of the specific platter is activated. Since the disk is rotated constantly, the head has to wait for the required sector or cluster

(desired data) to come under it. This waiting time is known as *rotational delay time* or *latency* of the disk. The rotational latency of a disk with 7200 rpm is 4.17 milliseconds.

3. **Data Transfer:** Once the read/write head is positioned over the desired sector, the data can be transferred to or from the disk to primary memory. The rate at which the data is read from or written to the disk is known as *data transfer rate*. It is measured in kilobits per second (kbps). The data transfer rate depends upon the rotational speed of the disk. If the disk has a rotational speed of 6000 rpm (rotations per minute), having 125 sectors and 512 bytes/sector, the data transfer rate per rotation will be 125 × 512 = 64,000 bytes. Hence, the total transfer rate per second will be 64,000 × 6000/60 = 6,400,000 bytes/second or 6.4 MB/second (see Table 3.1).

Table 3.1 Accessing Data

Step	Measured As	Illustration
Seek	Seek Time	
Rotate	Rotational Delay	
Data Transfer	Data Transfer Time	

The combined time (seek time, latency and data transfer time) is known as the *access* time of the magnetic disk. Generally, the access time can be described as the period of time that elapses between a request for data from disk or memory and the desired data arriving at the requesting device. *Memory access time* refers to the time it takes to transfer a character from memory to or from

the processor, while *disk access time* refers to the time it takes to place the read/write heads over the requested data. RAM may have an access time of 9–70 nanoseconds, while hard disk access time could be 10–40 milliseconds.

3.9 TYPES OF MAGNETIC DISKS

All magnetic disks come in the form of round platters. These disks are available in different sizes, shapes and designs. Some are attached to the read/write head assembly whereas some are available in the form of removable disks. Broadly, magnetic disks can be classified into three types: floppy disk, hard disk and zip disk.

3.9.1 Floppy Disk

A floppy disk is a round, flat piece of Mylar plastic coated with ferric oxide (a rust-like substance containing tiny particles capable of holding a magnetic field) and encased in a protective plastic cover (disk jacket). It is a removable disk and is read and written by a floppy disk drive (FDD), which is a device that performs the basic operation on a disk, including rotating the disk and reading and writing data onto it. The disk drive's read/write head alters the magnetic orientation of the particles, where orientation in one direction represents "1" and orientation in the other represents "0".

Traditionally, floppy disks were used on personal computers to distribute software, transfer data between computers and create small backups. Earlier, 5¼-inch floppy disks were used. Later, a new format of 3½-inch floppy disk came into existence, which has larger storage capacity and supports faster data transfer as compared to 5¼-inch floppy disks. Floppy diskettes are small, inexpensive, readily available, easy to store and have a good shelf life if stored properly. They also possess the write-protect feature, which allows the users to protect a

diskette from being written on. To write-protect a diskette, the user has to shift a slide lever towards the edge of the disk, uncovering a hole. The key advantage of floppy disk is that it is portable.

Read/Write Operation of a Floppy Disk: To read and write data onto a floppy disk, FDD is used. The drive (see Figure 3.11) is made up of a box with a slot (having a drive gate) into which a user inserts the disk. When the user inserts a disk into the FDD, the drive grabs the disk and spins it inside its plastic jacket. Also, the drive has multiple levers that get attached to the disk. One lever opens the metal plate, or shutter, to expose the data access area. Other levers and gears move two read/write heads until they almost touch the diskette on both sides. The drive's circuit board receives instructions for reading/writing the data from/to disk through the *floppy drive controller*. If the data are to be written onto the disk, the circuit board first verifies that no light is visible through a small window in the floppy disk. If the photo sensor on the opposite side of the floppy disk detects a beam of light, the floppy drive detects the disk to be write-protected and does not allow recording of data.

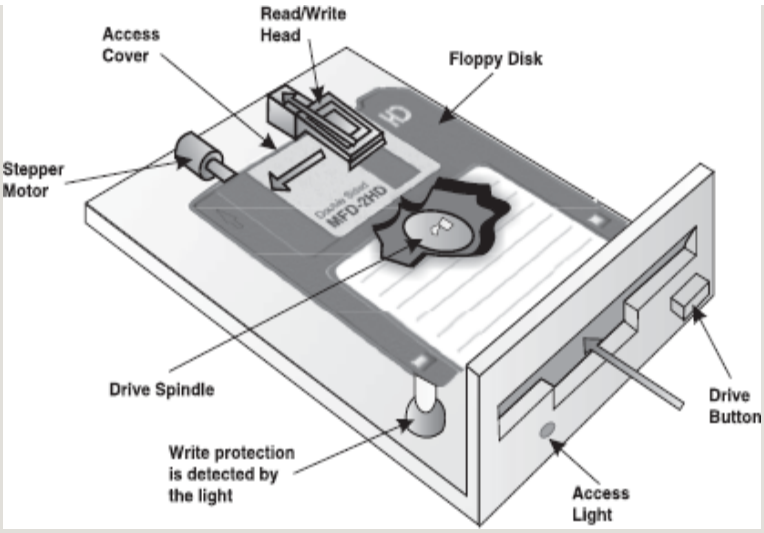


Figure 3.11 Floppy Disk Drive

The circuit board translates the instructions into signals that control the movement of the disk and the read/write heads. A motor located beneath the disk spins a shaft that engages a notch on the hub of the disk, causing the disk to spin. When the heads are in the correct position, electrical impulses create a magnetic field in one of the heads to write data to either the top or bottom surface of the disk. Similarly, on reading the data, electrical signals are sent to the computer from the corresponding magnetic field generated by the metallic particle on the disk.

Since the floppy disk head touches the diskette, both media and head wear out quickly. To reduce wear and tear, personal computers retract the heads and stop the rotation when a drive is not reading or writing. Consequently, when the next read or write command is given, there is a delay of about half a second while the motor gathers maximum speed.

3.9.2 Hard Disk

The *hard disk*, also called the *hard drive* or fixed disk, is the primary storage unit of the computer. It consists of a stack of disk platters that are made up of aluminium alloy or glass substrate coated with a magnetic material and protective layers (see Figure 3.12). They are tightly sealed to prevent any dust particle, which causes head crash, from getting inside. A hard disk can be external (removable) or internal (fixed) and can hold a large amount of data. The capacity, that is, the amount of information that a hard disk can store, is measured in bytes. A typical computer today comes with 80–320 GB of hard disk. The storage capacity of hard disk has increased dramatically since the day it was introduced. The hard disk speed is measured in terms of access time (typically in milliseconds). A hard disk with lower access time is faster than a hard disk with higher access time; the lower the access time, the faster the hard disk.

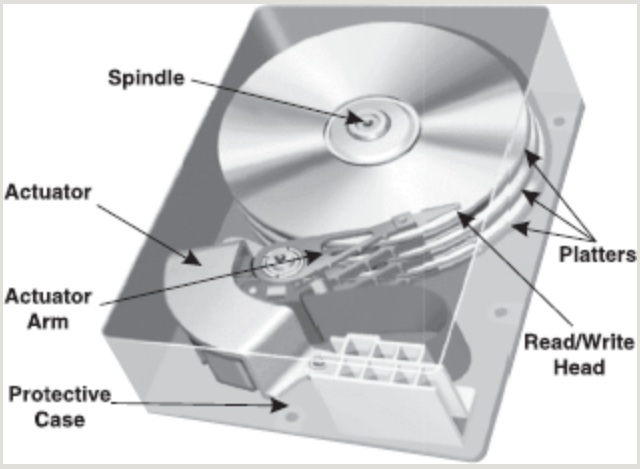


Figure 3.12 Hard Disk

Read/Write Operation of a Hard Disk: A hard disk uses round, flat disks (platters) made up of glass or

metals which are coated on both sides with a special material designed to store information in the form of magnetic patterns. Each platter has its information recorded in tracks, which are further broken down into smaller sectors. Making a hole in the centre of platters and stacking them onto a spindle mount the platters. The platters rotate at high speed, driven by a special motor connected to the spindle. Special electromagnetic read/write heads are mounted onto sliders and are used to either record data onto the disk or read data from it. The sliders are mounted onto arms, all of which are mechanically connected into a single assembly and positioned over the surface of the disk by a device called *actuator*. Each platter has two heads, one on the top of the platter and one on the bottom, so a hard disk with three platters would have six surfaces and six heads.

Data are recorded onto the magnetic surface of the disk in exactly the same way as they are on floppies. However, a disk controller is attached to the hard disk drive that handles the read/write commands issued by the operating system. Each read/write command specifies a disk address that comprises the surface number, track number and sector number. With this information, the read/write head moves to the desired sector that data can be read from or written to. Usually, the next set of data to be read is sequentially located on the disk.

Note that unlike floppy drives, in which the read/write heads actually touch the surface of the material, the heads in most hard disks float slightly off the surface. Nevertheless, the distance between the head and the disk surface is much less compared to the thickness of a human hair (see Figure 3.13). When the heads accidentally touch the media, either because the drive is dropped or bumped hard or because of an electrical malfunction, the surface becomes scratched. Any data stored where the head has touched the disk is lost. This is

called a *head crash*. To help reduce the possibility of a head crash, most disk controllers park the heads over an unused track on the disk when the drive is not being used by the CPU.

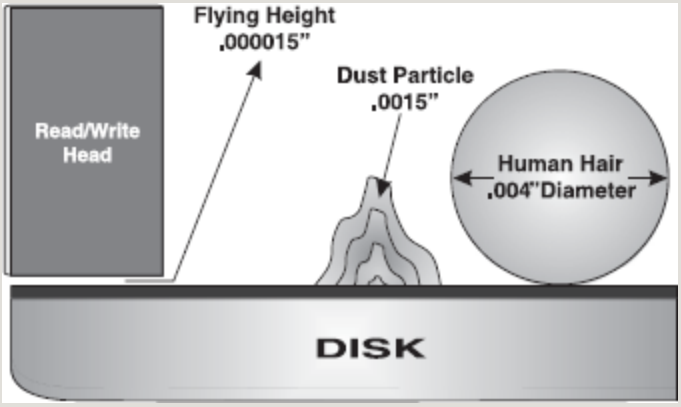


Figure 3.13 Distance between Head and Disk Surface

3.9.3 Zip Disk

Zip disk is a removable storage device that started with 100 MB storage capacity but today it can store 250–750 MB of data. Zip disk's drive unit is measured as 18 × 13 × 4 cm and weighs about half a kilogram (see Figure 3.14). It has rubber feet to stabilize the unit in either vertical or horizontal position. The substrate for the disk is made of a plastic material coated with magnetic oxide particles. There are two indicator lights: green for power and amber for disk access. It has an eject button, but no ON/OFF switch. It comes complete with drive, connection cable, power supply, operating software and a starter 100 MB disk with a set of zip tools software. On installation, the zip drive is automatically assigned a logical drive letter, and you can transfer files to and from the zip disk as if you are accessing a hard disk or floppy disk.

Figure 3.14 Zip Disk Drive

Zip disk is similar to the 3½-inch floppy diskette in diameter. However, one should keep in mind that zip drive cannot read or write floppy disks and similarly a floppy drive cannot read/write data from/to a zip disk. Either the zip drive can be built into the computer or it can be connected to the computer via the parallel port. There are different versions of zip disks and drives available in the market. These versions can hold 100, 250 and 750 MB of data. As a zip disk can hold as much as 70 times the information as compared to a floppy disk, it can be used to store heavy graphics, music or presentation files.

Read/Write Operation of a Zip Disk: A zip disk looks quite similar to a floppy disk. Both are about the same size but vary in their storage capacity. The magnetic-coated Mylar zip disk, which is also called a *cookie*, is encased in a hard plastic material having a metal plate cover that slides open to provide read/write access to the cookie. When a zip disk is inserted into the zip drive, the metal plate slides to expose the small portion of the cookie. The metal hub that is attached to the centre of the cookie spins at 3000 rpm (rotations per minute) along with the motor. The two heads on either

side of the cookie touch the surface of the disk to perform read/write operations. In a zip drive, the read/write heads are about one-tenth the size of a floppy drive head resulting in a higher data recording density. The small size of the zip drive heads allows writing 2118 tracks per inch. Zip drives, like a hard disk, use zone recording so that the same recording density is used throughout the disk. In contrast, the conventional floppy is divided into sectors radially in which the outermost tracks use up more surface area than the innermost tracks, and thus waste the recording surface.

3.9.4 Advantages of Magnetic Disks

- Magnetic disks follow direct access mode for reading and writing the data files, thereby making an ideal device for accessing frequently accessed data.
- Magnetic disks are used for both online and offline storage of data. Hard disk is used as an online storage whereas floppy and zip disks are used as offline storage.
- Magnetic disks are easily moveable from one place to another because of their small size.
- The data transfer rate of disks is much higher than magnetic tapes.
- Due to low cost and high data recording densities, cost per bit in magnetic disks is minimum.
- The storage capacity of magnetic disks is virtually unlimited as numbers of such disks can be added to store data.
- Magnetic disks are less prone to corruption of data as they can withstand temperature and humidity change much better as compared to magnetic tapes.

3.9.5 Disadvantages of Magnetic Disks

- Magnetic disks must be stored in a dust-free environment in order to protect them from crashing.
- Magnetic disks are not ideal devices to use in scenarios where the file access required is of sequential nature rather than direct or random nature.
- Magnetic disks are more expensive than the magnetic tapes.
- Magnetic disks are more susceptible to breach of security and access gain to sensitive online disk files from remote terminals.

3.10 OPTICAL DISK

Apart from magnetic tapes and magnetic disks, a new storage medium, which is gaining popularity, is the

optical disk. An optical disk is a flat, circular, plastic disk coated with material on which bits may be stored in the form of highly reflective areas and significantly less reflective areas, from which the stored data may be read when illuminated with a narrow-beam source, such as a laser diode. These disks are capable of storing enormously high amounts of data in a limited amount of space. The optical disk storage system consists of a rotating disk coated with a thin layer of metal (aluminium, gold or silver) that acts as a reflective surface and a laser beam, which is used as a read/write head for recording data onto the disk.

Optical disk comes in various sizes and capacities. A compact disk (CD) with 700 MB capacity and 12 cm diameter is the most popular means of optical storage. In a single-track optical disk, storage capacity is calculated by the multiple of number of sectors and number of bytes per sector. Since the storage capacity of an optical disk is huge, the cost per bit of storage is very low.

3.10.1 Storage Organization of Optical Disk

Unlike magnetic disk (which has several layers of concentric layers), an optical disk consists of a single long track in spiral shape as shown in Figure 3.15. This track starts from the outer edge and spirals inward to the centre of the disk. The spiral shape of the track makes the optical disk suitable for storing large blocks of sequential data onto it, such as music. Even though random access is possible, it is usually slower than magnetic disk (especially hard disk). The reason behind it is that in case of magnetic disk, it is easy to locate the magnetic disks tracks since they are located at a fixed distance from the centre.

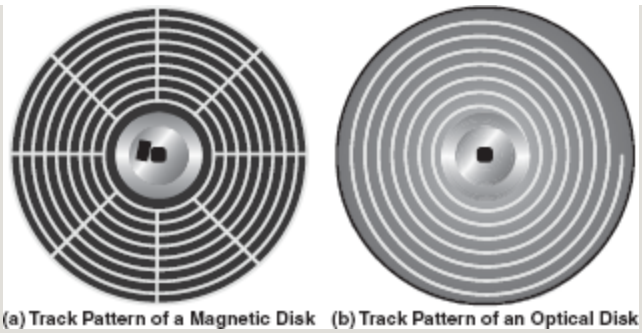


Figure 3.15 Comparing Track Patterns

These tracks are further divided into small sectors of the same length, regardless of their location on the disk's surface. Such type of disk formatting allows the data to be stored in even more densely as compared to magnetic disks. With more data recording density, the optical drives have a more complicated mechanism as compared to the magnetic disks drives. Here, rotation speed of the disk varies inversely with the radius of the disk. The disk moves slowly when data are being read near the edges and moves fast when close to the centre.

Since read/write operations are performed through laser beams, no access arm movement is required that is used in the case of the magnetic disk. Hence, many people assume that the data access time is even faster than the magnetic disk. However, this is not true because compared to the magnetic disks, where the tracks are arranged in the form of concentric circles, in an optical disk the tracks are organized in a spiral fashion. This results in the slower random access time than in the concentric circles. Generally, the access time for an optical disk ranges from 10 to 40 milliseconds which is very low as compared to a floppy disk that has an access time of several hundred milliseconds.

3.10.2 Access Mechanism of Optical Disk

The laser beam technology used for reading/writing of data on the disk surface of the optical disk uses two laser beam sources of different intensities. The greater intensity laser beam is used to write on the recording surface by turning it ON and OFF at a varying rate so that tiny pits are burnt into the metal coating of the disk. The lesser intensity beam is used to read that stored data, which is strongly reflected by the coated surface called *land* and weakly reflected by the burnt surface called *pit*. The changes in pattern are detected by a photo sensor and converted into digital signals (see Figure 3.16).

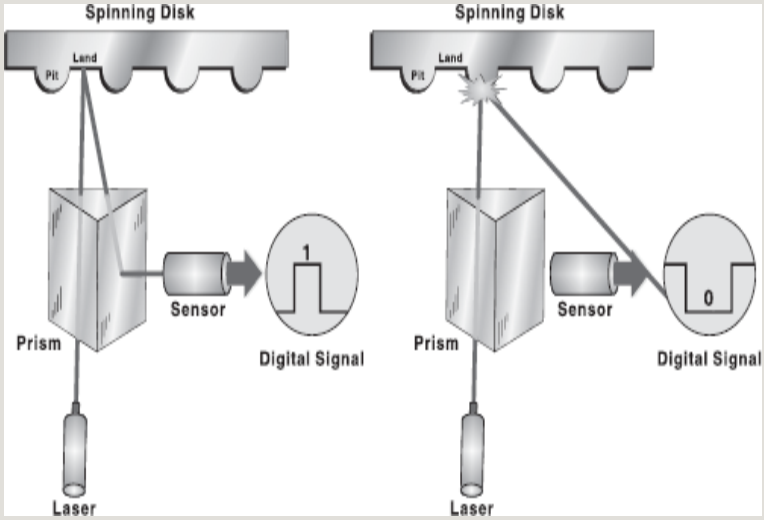


Figure 3.16 Read Operation in an Optical Disk

3.11 TYPES OF OPTICAL DISKS

Optical disks come in several varieties, which are made in somewhat different ways for different purposes. Compact disk (CD), digital versatile disk (DVD) and blu-ray disk (BD) are the three forms of optical disks. Note that these technologies are not compatible with one

another. Each requires a different type of disk drive and disk. Even within one category, there are many competing formats.

3.11.1 Compact Disk

Compact disk is the most popular and the least expensive type of optical disk. It was originally intended only for storing music (in the form of digital audio) and can record about 80 minutes of uninterrupted playing time. A CD is capable of being used as a data storage device along with storing of digital audio.

FACT FILE

What's in a Name?

CD-Rs and CD-RWs are made up of a polycarbonate substrate, a thin metal coating and a protective outer layer. In a CD-R, a layer of organic polymer dye between the polycarbonate and metal layers serves as the recording medium. The laser creates marks (*burns*) in the dye layer that mimic the reflective properties of the pits and lands (lower and higher areas) of the CD. Due to this, CD writing is also known as CD *Burning*.

A CD is a shiny, silver colour metal disk of 12 cm diameter. A typical optical disk is made up of three layers: a *polycarbonate base* through which light can pass, a *layer* of *aluminium*, and a protective *layer of acrylic* on top of that. The pits of CD are typically 0.5 microns wide, 0.83 to 3 microns long, and 0.15 microns deep. A CD has one track that spirals from the centre to the outside edge. If one could remove the track from a standard 12 cm CD, it would stretch for 3.5 miles. The single track is divided into sectors of equal length and

density. Files are stored on these particular contiguous sectors.

Compact disks are available in various formats: *CD-ROM* (*compact disk-read only memory*), *CD-R* (*compact disk-recordable*), and *CD-RW* (*compact disk-rewritable*) disks. A CD-ROM disk comes with pre-recorded data by the manufactures and can be read but cannot be altered. CD-R is a type of WORM (write once-read many) disk that allows you to record your own data. Once written, the data on the CD-R can be read but cannot be altered. A CD-RW disk is rewritable version of CD-R that means, it allows writing, erasing and rewriting of data several times. The data recorded on all CD formats can be read using the CD-ROM drive; however, to write data on CD-R and CD-RW disks, one needs a special peripheral device known as CD-writer (or *CD-burner*).

CD-ROM drives are characterized by the *spin rate*. The spin rate is the rotation speed of the disk and it influences the information retrieval speed (access time). Currently quad, hex and octal-speed CD-ROM drives are available where quad means 4x, hex means 6x and octal means 8x. Thus, octal-speed drive is the fastest and, as a result, the most expensive.

Reading Data from a CD-ROM: A CD is made up of three coatings, namely, *polycarbonate plastic, aluminium* and an *acrylic* coating to protect the disk from external scratches and dust. The polycarbonate plastic is stamped with millions of tiny indentations (pits), moulded in the spiral-shaped track. A CD drive reads information from the CD's spiral track of pits and lands, starting from the centre of the disk and moving to the outer edge, as shown in Figure 3.17. A light is beamed from a semiconductor laser through the bottom of the polycarbonate layer, and the aluminium coating monitors the light being reflected. Since the CD is read

through the bottom of the disk, each pit appears as an elevated bump to the reading light beam. Light striking the land areas (the areas without bumps) is reflected normally and detected by a photodiode. As the disk rotates at speeds between 200 and 500 rpm, the light bounces off the pits causing the frequency of the light to change. The reflected light then passes through a prism and onto a photo sensor. Light reflected from a pit is 18^{o} out of phase with the light from the lands, and the differences in intensity are measured by the photoelectric cells, which converts into a corresponding electrical pulse.

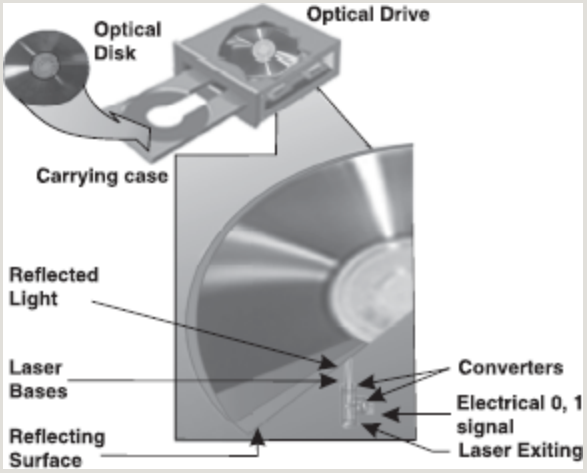


Figure 3.17 Reading Data from CD-ROM

Writing Data to a CD: On a new CD-R disk, the entire surface of the disk is reflective; the laser can shine through the dye and reflect off the gold layer. Hence, for a CD-R disk to work, there must be a way for a laser to create a non-reflective area on the disk. A CD-R disk, therefore, has an extra layer that the laser can modify. This extra layer is a greenish dye. When you write data to

a CD-R, the writing laser (which is much more powerful than the reading laser) heats up the dye layer and changes its transparency (see Figure 3.18). The change in the dye creates the equivalent of a non-reflective bump. The decomposition of the dye in the pit area through the heat of the laser is irreversible (permanent). Therefore, once a section of a CD-R is written, it cannot be erased or rewritten. However, both CD and CD-R drives can read the modified dye as a bump later on.

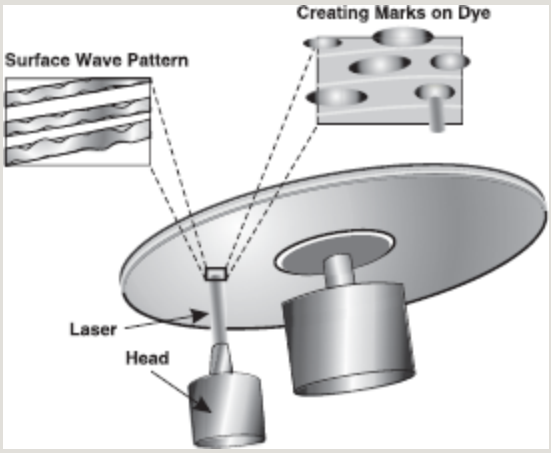


Figure 3.18 Writing Data to a CD-R

In contrast to CD-R disk, CD-RW disk is erasable and rewritable because it uses phase-changing material on its recording layer, usually an alloy of silver, tellurium, indium and antimony metals. Phase-changing material changes its state when heated to a high temperature (above its melting point) and can be converted back to its original state when heated at a temperature slightly below its melting point.

In CD-RW disk, the recording layer has a polycrystalline structure initially. While writing to the disk, the laser heats up the selected areas to a very high temperature (above the melting point), which melts the crystals into a non-crystalline amorphous phase. These areas have lower reflectance than the remaining crystalline areas. This difference in reflectance helps in reading the recorded data as in the case of CD-R disk.

To erase data on a CD-RW disk, a process called *annealing* is used. During this process, the area on the layer that has been changed to the amorphous phase (during writing) is converted back to its original crystalline state by heating it to a temperature slightly below the melting point of the phase-changing material.

3.11.2 Digital Versatile Disk

DVD, initially called *Digital Video Disk*, is a high-capacity data storage medium. At first glance, a DVD can easily be mistaken for a CD as both are plastic disks 120 mm in diameter and 1.2 mm thick, and both rely on lasers to read data. However, the DVD's seven-fold increase in data capacity over the CD has been largely achieved by tightening up the tolerances throughout the predecessor system. In DVD, the tracks are placed closer together, thereby allowing more tracks per disk. The DVD's *track pitch* (the distance between each) is reduced to 0.74 micron, less than half of CD's 1.6 micron. The pits, in which the data is stored, are also a lot smaller, thus allowing more pits per track. The minimum pit length of a single layer DVD is 0.4 micron as compared to 0.834 micron for a CD. With the number of pits having a direct bearing on capacity levels, DVD's reduced track pitch and pit size, as shown in Figure 3.19, alone give DVDs four times the storage capacity of CDs.

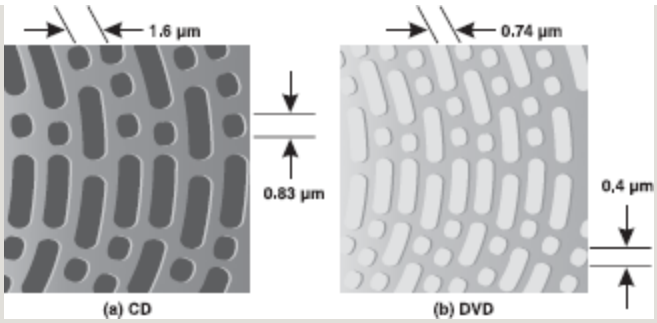


Figure 3.19 Comparing Track Pitch and Pit Length

Note: *Like CDs, DVDs are also available in different formats: DVD-ROM, DVD-R, and DVD-RW.*

The packing of as many pits as possible onto a disk is, however, the simple part; the real technological breakthrough of DVDs was with its laser. Smaller pits mean that the laser has to produce a smaller spot, and DVD achieves this by reducing the laser's wavelength from the 780 nm (nanometres) infrared light of a standard CD, to 635 nm or 650 nm. Secondly, the DVD specification allows information to be scanned from more than one layer of a DVD simply by changing the focus of the read laser. Instead of using an opaque reflective layer, it is possible to use a translucent layer with an opaque reflective layer behind carrying more data. This does not double the capacity because the second layer cannot be quite as dense as the single layer, but it does enable a single disk to deliver 8.5 GB of data without having to be removed from the drive and turned over.

An interesting feature of DVD is that the disk's second data layer can be read from the inside of the disk out, as well as from the outside in. In standard density CDs, the information is always stored first near the hub of the disk. The same is true for single-layer and dual-layer

DVD, but the second layer of each disk can contain data recorded “backwards”, or in a reverse spiral track. With this feature, it takes only an instant to refocus a lens from one reflective layer to another. On the other hand, a single-layer CD that stores all data in a single spiral track takes longer to relocate the optical pickup to another location or file on the same surface.

DVD allows for double-sided disks (see Figure 3.20). To facilitate the focusing of the laser on the smaller pits, manufacturers used a thinner plastic substrate thereby reducing the depth of the plastic layer that the laser has to travel through to reach the pits. This reduction resulted in disks that were 0.6 mm thick—half the thickness of a CD. However, since these thinner disks were too thin to remain flat and withstand handling, manufacturers bonded two disks back-to-back, resulting in disks that are 1.2 mm thick. This bonding doubles the potential storage capacity of a disk. Note that single-sided disks still have two substrates, even though one is not capable of holding data.

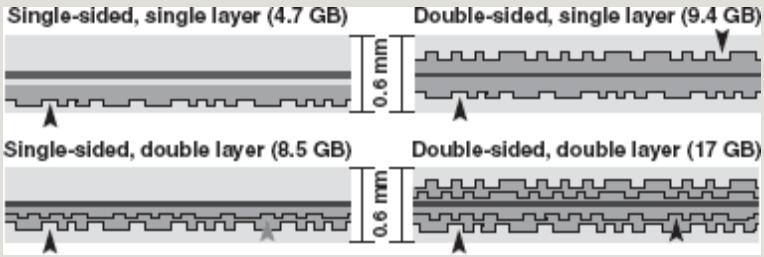


Figure 3.20 Different Types of DVDs

3.11.3 Blu-ray Disk

Blu-ray disk is an optical storage device, which is used to record and playback high definition video and audio as well as store images and other data (Figure 3.21). It is

developed by Blu-ray Disk Association (BDA), which is a group of world's leading consumer electronics, personal computers and media manufacturers including Thomson, Apple, Panasonic, Sony and Samsung. It uses blue-violet laser having shorter wavelength (405 nm) than a red laser (650 nm) used by DVDs. Due to this shorter wavelength, the laser can be focused more precisely on the small spot thereby resulting in storage capacity 10 times that of DVD. Like CDs and DVDs, Blu-ray disks are also available in different formats:

Figure 3.21 A Blu-ray Disk

FACT FILE

DVD and Blu-ray Disk

The organization of both Blu-ray disk and DVD is similar; however, the way the data are stored on them differs. In DVDs, the data reside between the two polycarbonate layers. Due to the layer on the top of data, the substrate layer may refract the laser beam into two separate beams, which is known as *birefringence*. If the beam is split too widely, the disk-reading operation cannot be performed. Blu-ray disk overcomes the birefringence by placing the

data on the top of the polycarbonate layer, and thus prevents the disk-reading problem.

- **BD-ROM:** It comes with prerecorded content that can only be read.
- **BD-R:** It is a WORM type of disk on which you can record data only once.
- **BD-RW:** It is similar to BD-R disk but the difference is that it is rewritable. This means data can be erased and recorded a number of times on the same disk.
- **BD-RE:** It is also a rewritable disk but is used only for high definition audio/video and television recording.

Blu-ray disks are available in two sizes: standard (12 cm) and mini (8 cm). Each disk can have either a single layer or dual layers, depending on which the data storage capacity of the disk differs. For instance, the storage capacity of standard disk with single layer is 25 GB and with dual layer is 50 GB. On the other hand, the storage capacity of mini disk with single layer is 7.8 GB and with dual layer is 15.6 GB.

Blu-ray disk allows not only recording of programs but the programs can be manipulated also. One can edit the programs recorded on the disk or change their order. The user can also create playlists of the programs stored on disks so that they can be randomly accessed. Moreover, it facilitates recording one program while simultaneously watching another program.

3.11.4 Advantages of Optical Disks

- Optical disks possess large capacity to store data/information in the form of multimedia, graphics and video files. They can store more data in less amount of space as compared to magnetic tapes and floppy or zip disks.
- The life span for data storage in optical disks is considered to be more, about 10–20 years as compared to magnetic disks, which have a comparatively lesser life span.
- Optical disks hold more data-recording density as compared to other storage media; therefore, they have low cost per bit of storage.
- Optical disk is not affected by magnetic field.
- An optical disk is tougher than tapes/floppy disks. It is physically harder to break or melt.

- Due to their small size and lightweight, these disks are easily portable and stored.

3.11.5 Disadvantages of Optical Disks

- Optical disk is not as easy to write as a floppy disk. One needs to use both software and hardware for writing optical disks.
- They possess slow data access speed as compared to the magnetic disks.
- The drive mechanism of an optical disk is more complicated than the magnetic and floppy disks.

3.12 MAGNETO-OPTICAL STORAGE DEVICES

The magneto-optical (MO) systems include basic principles of both magnetic and optical storage systems. This system writes magnetically (with thermal assist) and reads optically. Magnetic optical disk is a plastic or a glass disk coated with a compound (often a ternary alloy of terbium ferric cobalt, or TbFeCo) with special properties (see Figure 3.22). Since such materials are easily oxidized, dielectric barrier layers are used to protect the MO layer from oxidation. At the same time, the barriers together with the reflector coating act as an optical signal enhancement. MO disks function along opto-thermic magnetic principles. The thin film structure that gives the format its unlimited rewritability is based on a MO alloy layer enveloped by a barrier layer on each side.

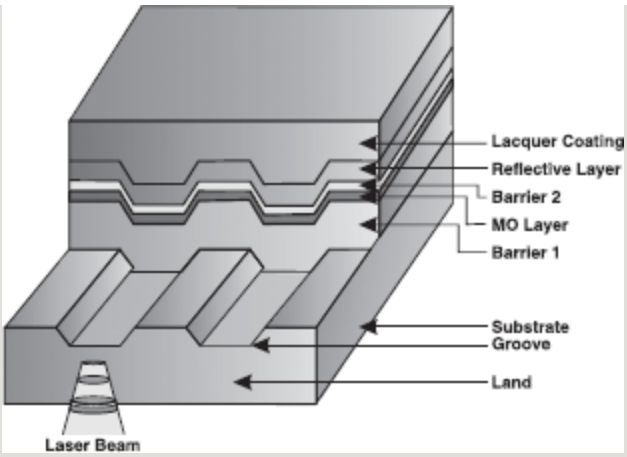


Figure 3.22 Magneto-Optical Disk Structure

Presently, MO disks are available in two formats: 5¼-inch and 3½-inch. The larger form-factor MO disks are capable of holding data about as much as the standard CD-ROM. Under pressure from the inexpensive and relatively fast CD-R and CD-RW, MO drives seem to be losing ground. On the other hand, some of the principles of the MO technology (thermally-assisted magnetic recording) may find their way into the most advanced magnetic storage devices of the future. Thus, the existence of the erasable optical disks has been made feasible due to this technology.

3.12.1 Basics of MO Reading

Traditional magnetic recording systems use currents induced in the magnetic heads by the changing magnetic fluxes on the disk surface to read the data. However, magneto-optical systems use polarized light to read the data from the disk (see Figure 3.23). The changes in light polarization occur due to the presence of a magnetic field on the surface of the disk. This phenomenon is known as the *Kerr effect*, where the polarity of the reflected light is

altered depending on the orientation of the magnetic particles. If a beam of polarized light is shined on the surface, the light polarization of the reflected beam will change slightly, (typically less than 0.50) if it is reflected from a magnetized surface. If the magnetization is reversed, the change in polarization (the Kerr angle) is reversed too. The magnetized areas, that is, pits, cannot be seen in regular light, but only in polarized light. The change in direction of magnetization could be associated with numbers 0 or 1, making this technique useful for binary data storage.

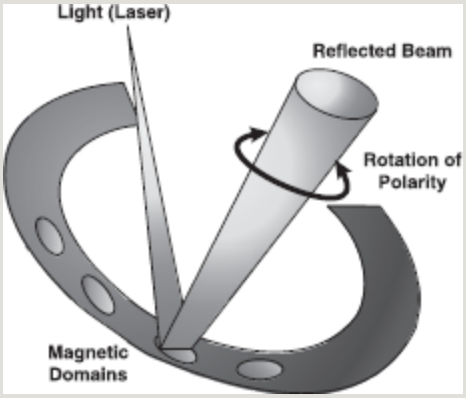


Figure 3.23 Reading from Magneto-optical Disk

Various other rewritable disks include CD-RW, DVD-RW, DVD-RAM and DVD+RW. The data layer for these disks uses a phase-changing metal alloy film. This film can be melted by the laser's heat to level out the marks made by the laser and then lasered again to record new data. In theory, one can erase and write on these disks as many as 1000 times, for CD-RW, and even 100,000 times for the DVD-RW types. DVD-RAM was the first erasable version of DVD to come to market and has

subsequently found competition in the rival DVD-RW and DVD+RW format.

3.12.2 Basics of MO Recording

After reaching the *Curie temperature* (around 200°C), every magnetic material loses magnetization due to a complete disordering of their magnetic domains, and as a result, lose the data stored on them. Moreover, the material's *coercivity*, that is, the measure of the material's resistance to magnetization by the applied Lens magnetic field decreases as the temperature approaches the Curie point, and reaches zero when this temperature is exceeded. Therefore, it is imperative that the only change to the material, when it is heated and cooled, is the change in magnetization with no damage to the material itself.

The fact that the material's coercivity drops at higher temperatures allows thermally assisted magnetic recording with relatively weak magnetic fields. Even a relatively weak laser can generate high local temperatures when focused at a small spot (about 1 micron in case of MO systems). When the material is heated, and its coercivity is low, the magnetization of the media can be changed by applying a magnetic field from the magnet. When the material is cooled to room temperature, its coercivity rises back to such a high level that the magnetic data cannot be easily affected by the magnetic fields. The basic schematic of this recording process is illustrated in Figure 3.24.

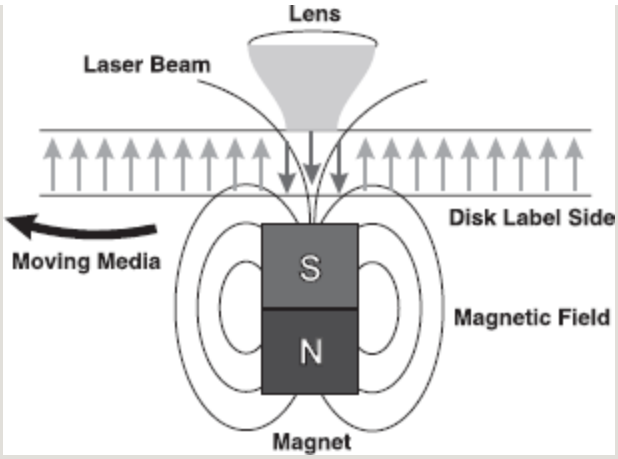


Figure 3.24 Writing to Magneto-Optical Disk

The recording layer is heated by the laser to a point where its magnetic orientation is dissipated. As this spot on the disk cools, the new magnetic orientation—corresponding to the new information—is set by the magnetic head to correspond to "0" and "1" of digital signals. Normally, to rewrite data on a MO disk, it is necessary to erase all previous signals before new data can be recorded. This means either two lasers are used to record (one to erase and one to record) or a single laser must first erase the data in a first rotation and then record the data the second time. This requires either a recorder with two lasers or longer recording times for a single laser system.

3.13 UNIVERSAL SERIAL BUS

Universal serial bus (USB), developed by Intel, is a set of connectivity specifications that establishes communication between personal computers and devices such as mouse, keyboard, pen drive, external hard disk drives, etc. Nowadays, almost every computer or laptop is equipped with one or more USB ports. The USB

connector is a narrow socket around 1 cm wide and 0.5 cm high.

All USB devices come with a USB connector that is plugged into the USB port on the computer. As you plug in a USB device, it is detected by the computer and the required software is configured automatically; there is no need to restart the computer. Another advantage of USB is that it determines (and provides) the electrical power required by the device connected to it.

3.13.1 Pen Drive

A pen/flash drive is a removable storage device that is frequently used nowadays to transfer audio, video, and data files from one computer to another. A pen drive consists of a small printed circuit board, which is fitted inside a plastic, metal, or rubber casing to protect it. The USB connector which is present at one end of pen drive is protected by either a removable cap or pulling it back in the casing. Figure 3.25 shows a pen drive.

Figure 3.25 Pen Drive

The pen drive is a high-storage–capacity (ranging from 1 GB to 32 GB) device and is physically small enough to fit into a pocket. In addition, it is fast, robust and reliable and requires very less power to operate, which it gets through a USB port, and hence no battery is required.

Using a pen drive is easy. Following are the steps to use pen drive.

1. Insert the pen drive in the USB port of a computer. The computer detects the device, configures the necessary software and displays a drive corresponding to the pen drive in **My Computer**.
2. Perform any of the following actions:
    1. Double-click the drive icon in **My Computer** to access the files stored in the pen drive.
    2. Right-click the file or folder you want to send to the pen drive, which displays a menu. Select **Send To** and select the drive corresponding to the pen drive from the sub menu that appears.
3. After performing all your actions, left-click the **Safely Remove Hardware** icon in the system tray to display a menu containing **Safely remove USB Mass Storage Device** option.
4. Select this option to display a notification icon indicating that it is now safe to remove the hardware and plug-out the pen drive.

3.13.2 External Hard Disk Drive

External hard disk drive is a type of hard drive that resides in its own enclosure (called *hard drive cage*) outside the computer case and is connected to the system through interfaces like USB (Figure 3.26). With the addition of USB interfaces in the personal computer, external hard drive has become quite popular in the computer market. The storage capacity of external hard disk ranges from 20 GB to several TBs and it is compatible with operating systems that support interface standards such as USB MSC (mass storage class) and IEEE1394.

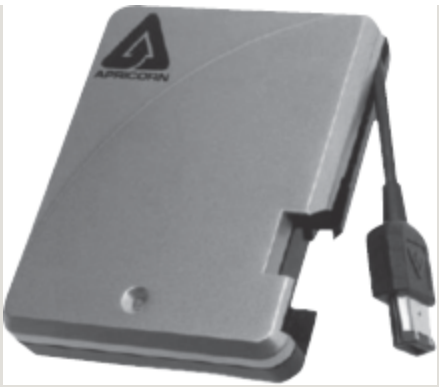


Figure 3.26 An External Hard Disk Drive with USB Connector

The internal structure and functioning of an external hard disk drive is similar to the internal hard disk drive. Hence, the external hard disk drive is a reliable and high capacity storage media. In addition, it is portable and provides the plug and play feature. It means just plug in the connector of external hard drive into the USB interface and access the stored data as well as transfer data to/from the external hard disk drive.

These features of external hard disk make it suitable to be used for taking backup. Using an external hard disk drive as a means of backup for important (or sensitive) data is advantageous. This is because you can unplug the external hard drive when not in use to protect your data from being compromised by online or offline activities.

3.14 MEMORY STICK

Memory Stick, also known as *memory card*, is a digital storage device, which is designed to be used with portable electronic devices such as mobile phones, digital cameras, PDAs, iPod, etc. (Figure 3.27). It was launched in 1998 by Sony and immediately gained popularity due to its support for fast data transfer speed and large storage capacity. Though the original Memory Stick

provided storage capacity of up to 128 MB, nowadays Memory Sticks with storage capacity of up to 32 GB are available in the market. Even the Memory Stick with the smallest storage capacity (4 MB) can store up to 80 images, which is much more than a standard 3½-inch floppy disk. The Memory Stick can be removed from the portable device and accessed by a personal computer using Memory Stick-capable memory card reader.

Figure 3.27 Memory Stick

Over the years and with advancement in technology, the Memory Stick is getting smaller in physical size and larger in logical size. Nowadays, several different standards or formats of Memory Stick are available in the market, which are as follows:

- **Memory Stick PRO:** It was introduced in the year 2003 jointly by two big companies—Sony and SanDisk. It supports marginally higher data transfer speed than the original one and provides theoretical storage capacity of up to 32 GB. It provides a 4-bit parallel interface with theoretical transfer rates of up to 480 Mb/s. It is widely used in high megapixel digital cameras and camcorders.
- **Memory Stick Duo:** It is a small size Memory Stick for small, pocket-sized devices such as mobile phones, music players, digital cameras, etc. It has reduced the use of large size Memory Stick, which had been very popular for a long time. It is equipped with MagicGate technology that is used to encrypt the data stored on the card. It provides all the features of the large standard Memory Stick and is smaller in size, but costs more.
- **Memory Stick PRO Duo:** Though Memory Stick Duo fulfils the need for pocket-sized devices, it has a slow transfer rate and limited storage capacity of 128 MB; however, Memory Stick PRO Duo supersedes it

because Memory Stick PRO Duo provides larger memory space (up to 32 GB) and high speed of data transfer to/from the card.

- **Memory Stick PRO-HG Duo:** It was introduced by Sony and SanDisk in 2006. Unlike Memory Stick PRO which has a parallel interface of 4-bit, Memory Stick PRO-HG DUO has an 8-bit parallel interface. In addition, the clock frequency has increased from 40 MHz to 60 MHz in Memory Stick Pro-HG Duo. With these improvements, it provides a higher transfer speed than the Memory Stick Pro Duo.
- **Memory Stick Micro (M2):** It is a light and compact storage media which comes in the dimension of 15 mm × 12.5 mm × 1.2 mm. It has mainly been developed to meet the demands of the mobile devices market. It offers large storage capacity ranging from 16 MB to 32 GB and transfer speed of 160 Mb/s.

3.15 MASS STORAGE DEVICES

In order to get a vast amount of storage capacity of several bytes (trillions and more) in a computer system, a different kind of storage system is used. In such type of systems, multiple units of similar kinds of storage media are associated together to form a chain of mass storage devices. These storage media may include multiple magnetic tape reels or cartridges, multiple arrays of magnetic disks or multiple CD-ROMs as a storage device. We can categorize mass storage devices into three types:

- **Redundant Array of Inexpensive Disks (RAID):** The basic idea of RAID is to combine multiple hard disks into an array of disk drives to obtain high performance, large capacity and reliability. These drives appear to the host computer as a single logical drive. The disk arrays can be made fault-tolerant by redundantly storing information in various ways. Seven types of array architectures, RAID 0 through RAID 6, have been defined; each provides disk fault-tolerance with different compromises in features and performance.
- **Automated Tape Library:** An automated tape library comprises numerous sets of magnetic tapes along with their drives and controllers mounted in a single unit (see Figure 3.28). The unit comprises one or more tape drives to perform read/write operations on the tapes in the tape library. In the multiple tape drive environment, these tapes can be simultaneously read or write, thus resulting in the speedy rate of data transfer. Multiple drives lead to the reliability of the storage unit because if one of the drives fails, then the unit can continue to operate with the other tape drives. The unit, with the help of a robotic arm, retrieves the appropriate tape from the tape library, mounts it on the tape drive for processing, and then returns to the library after the job has been finished. These tape libraries can store up to several terabytes of data in it, so they can be used for archiving data for offline storage and also as data backup devices during online storage.

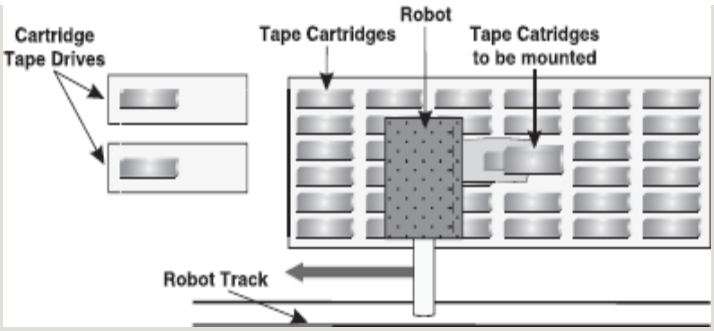


Figure 3.28 Automated Tape Library

- **CD-ROM Jukebox:** A CD-ROM jukebox comprises numerous sets of CD-ROM disks along with their drives and controllers mounted in a single unit. The unit comprises one or more CD-ROM drives to perform read/write operations on the CD-ROM in the jukebox. In the multiple CD-ROM drive environment, these CD-ROMs can be simultaneously read or write, resulting in the speedy rate of data transfer. Multiple drives lead to the reliability of the storage unit because if one of the drives fails, then the unit can continue to operate with other CD-ROM drives. The unit, with the help of a robotic arm, retrieves the appropriate CD-ROM from the CD-ROM jukebox, mounts it on the CD-ROM drive for processing, and then returns it to the appropriate slot in the jukebox after the job has been finished. These jukeboxes can store up to several terabytes of data in it, so they can be used for archiving of read-only data for offline storage and also as data backup devices during online storage.

Mass storage devices have relatively slow access time, generally in the order of seconds instead of milliseconds. Due to this, their use is limited in a number of applications. The slow access time of mass storage devices is because of the time taken by the transport mechanism to first move onto the particular storage media such as disks and then the time taken, in few milliseconds, to transfer the desired data to the main memory in the computer system. However, they have huge amounts of storage capacity and possess minimum cost per bit storage. Mass storage devices are a cost-effective option, as compared to the online tapes and disks storage, in situations where a large storage capacity is required and where prompt data access is not essential. When used as offline storage, they are referred to as *archival storage*.

LET US SUMMARIZE

1. *Computer memory* refers to the electronic holding place for instructions and data where the processor can reach quickly.
2. Various units that are used to measure computer memory are *bit, byte, kilobyte, megabyte, gigabyte and terabyte*.
3. A computer's memory can be categorized into three fundamental types: *internal processor memory, primary memory* and *secondary memory*.
4. *Internal processor memory* is placed within the CPU (processor) and usually includes cache memory and special registers, both of which can be directly accessed by the processor. This memory is used for temporary storage of data and instructions on which the CPU is currently working.
5. Primary memory consists of *RAM*, which is volatile in nature, and *ROM*, which is nonvolatile in nature.
6. *RAM*, also called *main memory*, allows the computer to store data for immediate manipulation and to keep track of what is currently being processed. It is the place in a computer where the operating system, application programs and data in current use are kept so that they can be accessed quickly by the computer's processor.
7. *SRAM* stands for static random access memory. It retains its contents as long as power is being supplied and does not require constant periodic refreshing. It is often used as cache memory due to its high speed.
8. *DRAM* stands for dynamic random access memory. The data on DRAM continues to move in and out of the memory as long as power is available, and thus DRAM must be continually refreshed in order to maintain the data. DRAM is slower and less expensive than SRAM.
9. The non-volatile memory chip, where the start-up instructions (special boot program) are stored is called *ROM*. These chips are programmed by burning appropriate fuses to form patterns of binary information.
10. *Secondary memory* is used to provide backup storage for instructions and data. It has much larger capacity than primary memory. It includes devices like hard disk, floppy disk, CD-ROM and tape drives. These devices are classified into two types according to data access, that is, *sequential access* and *direct access*.
11. *Magnetic tape* is like a plastic tape with a magnetic coating on it. The data is stored in the form of tiny segments of magnetized and demagnetized portions on the surface of the material. The tapes are an inexpensive and a reliable storage medium for organizing archives and taking backup.
12. Magnetic tape is divided into vertical columns (*frames*) and horizontal rows (*channels* or *tracks*). The data is stored in a string of frames with one character per frame and each frame spans multiple tracks. Thus, a single bit is stored in each track, that is, one byte per frame. A magnetic tape can typically have seven to nine tracks.
13. The length of tape between the *BOT* (*beginning of tape*) and *EOT* (*end of tape*) is referred to as the usable *recording (reading/writing)* surface. BOT/EOT markers are usually made up of short silver strips of reflective type.
14. *Magnetic disks* are the most widely used and popular medium for direct access secondary storage. They offer high storage

capacity and reliability, and have the capability to access the stored data directly.

15. A magnetic disk consists of a plastic/metal circular plate/platter, which is coated with a magnetic oxide layer. Data are represented as magnetized spots on a disk. A magnetized spot represents 1 and the absence of a magnetized spot represents 0.
16. *Tracks* are concentric circles where the data are stored, and are numbered from the outermost to the innermost ring, starting with zero.
17. *Sectors* refer to the number of fixed-size areas (imaginary pie slices) that can be accessed by one of the disk drive's read/write heads, in one rotation of the disk, without the head having to change its position. An intersection of a track and a disk sector is known as track sector.
18. A *floppy disk* is a round, flat piece of Mylar plastic, coated with ferric oxide, and encased in a protective plastic. It is a removable disk and is read and written by a floppy disk drive. A *floppy disk* drive is a device that performs the basic operation on a disk, including rotating the disk and reading and writing data onto it.
19. The *hard disk* is the primary storage unit of the computer. It consists of a stack of disk platters that are made up of aluminium alloy or glass substrate coated with a magnetic material and protective layers. It plays a significant role in: performance, storage capacity, software support and reliability.
20. *Zip disk* is a removable storage device having a capacity to store 250–700 MB of data. Zip disk's drive unit is measured as 18 × 13 × 4 cm and weighs about half a kilogram. It has rubber feet to stabilize the unit in either vertical or horizontal positions. The substrate for the disk is made up of plastic material on which magnetic oxide particles are coated.
21. An *optical disk* is a flat, circular, plastic disk coated with material on which bits may be stored in the form of highly reflective areas and significantly less reflective areas, from which the stored data may be read when illuminated with a narrow beam of laser diode. These disks are capable of storing enormously high amounts of data in a limited amount of space.
22. *Memory Stick*, also known as *Memory Card* is a digital storage device, which is designed to be used with portable electronic devices such as mobile phones, digital cameras, PDAs, iPod, etc.
23. *Magneto-optical* (*MO*) systems include the basic principles of both magnetic and optical storage systems. This system writes magnetically (with thermal assist) and reads optically.
24. *USB*, developed by Intel, is a set of connectivity specification that establishes communication between personal computers and devices such as mouse, keyboard, pen drive, external hard disk drives, etc.
25. In order to achieve a vast amount of storage capacity in a computer system, multiple units of similar kinds of storage media are associated together to form a chain of mass storage devices. These storage media include multiple arrays of magnetic disks or CD-ROMs as a secondary storage device.

Fill in the Blanks

1. ..... refers to the electronic holding place for instructions and data in a computer system.
2. The time period lag between a request made and the action being performed is called .....
3. The memory that requires constant refreshing to preserve the information is called .....
4. ..... memory provides backup storage for instructions and data that are not currently being used by the processor.
5. The two types of data access methods available with secondary storage devices are ..... and .....
6. An intersection of a track and a disk sector is known as .....
7. The time required to position the read/write head over the desired track is called .....
8. WORM stands for .....
9. ..... is a hybrid of magnetic and optical technologies.
10. The multiple association of a disk array box used for obtaining high performance, large capacity and reliability is called .....

Multiple-choice Questions

1. Which of the following is the largest memory size?
    1. Terabyte
    2. Gigabyte
    3. Kilobyte
    4. Megabyte
2. The access time refers to
    1. Time required to locate and retrieve stored data
    2. Time required to locate the lost data
    3. Time required to delete specific data on a certain memory location
    4. None of these
3. The type of memory that can be erased by simply exposing the device to a strong source of ultraviolet light for a certain amount of time is
    1. PROM
    2. Flash Memory
    3. EPROM
    4. EEPROM
4. The storage device that has high cost per bit of storage is
    1. SRAM
    2. Cache Memory
    3. ROM
    4. Hard Disk
5. The secondary storage device that follows the sequential mode of access is
    1. Optical Disk
    2. Magnetic Disk
    3. Magnetic Tape
    4. None of these
6. RAID stands for

1. Reproduce Array of Intelligent Disks
2. Reproduce Array of Inexpensive Disks
3. Redundant Array of Inexpensive Drives
4. Redundant Array of Inexpensive Disks

7. DVD stands for

    1. Decoded Video Disk
    2. Digital Versatile Disk
    3. Digital Virtual Disk
    4. None of these

8. In which kind of disk does the read/write head physically touch the surface?

    1. Hard Disk
    2. Compact Disk
    3. Floppy Disk
    4. None of these

9. A floppy can be made write-protected by

    1. Breaking the slider
    2. Positioning the slider to cover the hole
    3. Positioning the slider away from the hole
    4. Cannot be write-protected

10. A spiral shape track formatting is present in

    1. Floppy Disk
    2. Optical
    3. Hard Disk
    4. Half-inch Tape Cartridge

State True or False

1. A terabyte comprises 1024 megabytes.
2. The ROM chip, which can be rewritten several times and requires the action of ultraviolet radiations to erase its contents, is EPROM.
3. The memory which is utmost accessible to the processor is hard disk.
4. CD-ROM is a kind of magneto-optical disk.
5. Rotational delay time is also known as latency.
6. Auxiliary memory stores information about BIOS and start-up routines.
7. The memory that resides inside the processor is called processor memory.
8. DVD-ROM and CD-ROM has the same storage capacity.
9. The speed of memory depends upon the access time of data that is stored.
10. Floppy disk can be made write-protected.

Descriptive Questions

1. What are the types of memories available in the computer system? How are they organized in a hierarchy?
2. What is ROM? Differentiate between EPROM and EEPROM.
3. With the help of diagrams, explain how RAM, ROM and CPU interact with each other.
4. Explain the working of a magnetic tape. What are its advantages and disadvantages?
5. What are mass storage devices? How are they useful in storing large amounts of data?

6. Explain how read/write operation is carried out in an optical disk.
7. Discuss advantages and disadvantages of
    1. Magnetic disk
    2. Optical disks
8. Distinguish between the following:
    1. Sequential mode and Random mode
    2. Static RAM and Dynamic RAM
    3. Magnetic tape and Magnetic disk

ANSWERS

Fill in the Blanks

1. Memory
2. Latency
3. DRAM
4. Auxiliary
5. Sequential access, Direct access
6. Track sector
7. Seek time
8. Write once-read many
9. Magneto-optical disk
10. RAID

Multiple-choice Questions

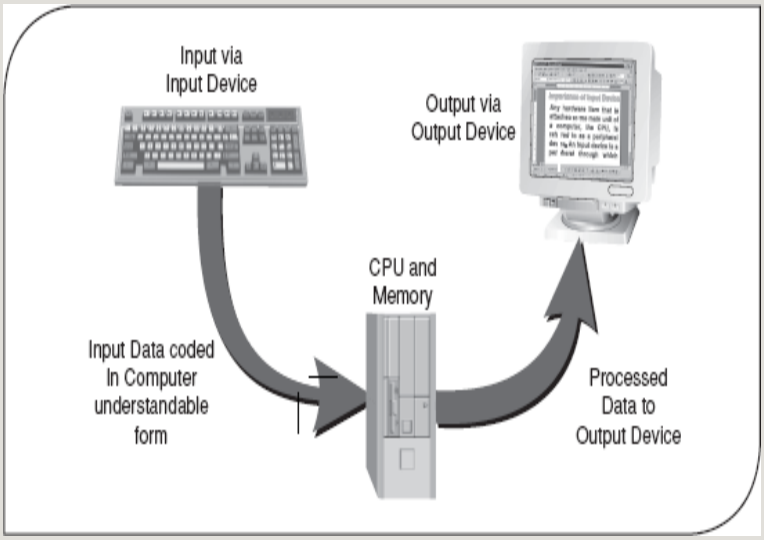
1. (a)
2. (a)
3. (c)
4. (b)
5. (c)
6. (d)
7. (b)
8. (c)
9. (c)
10. (b)

State True or False

1. False
2. True
3. False
4. False
5. True
6. False
7. True
8. False
9. True
10. True

Chapter 4

Input Output Media



This chapter deals with two basic computer units—input and output devices. Input devices are used to enter the data and instructions into the computer system before any processing can be performed. Output devices are used for the conversion of machine-readable information into human-readable form. The chapter discusses some commonly used input devices including keyboard, mouse, joystick, scanner, and optical scanners. The functions of all these input devices along with the hard copy and soft copy output devices under which, various types of printers and plotters have been explored. Next display devices such as CRT monitors, LCD monitors and speech synthesisers are explained. The chapter concludes by giving an overview of computer terminals.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The basic concepts of input and output devices
- Various types of input devices and how they function
- Various types of output devices and how they function
- Computer terminals which are special units capable of performing both input and output of data

4.1 INTRODUCTION

Previously, we discussed that a computer accepts input and processes it to get a desired output according to the sequence of instructions. Essentially, a computer system consists of four components: *input devices, CPU, output devices* and *memory*. Input devices are used to provide data to the CPU for processing. After processing, the input data is converted into meaningful information and this output is presented to the user with the help of output devices. In computer terminology, a device can be referred to as a unit of hardware, which is capable of providing input to the computer or receiving output, or both.

An *input device* is an electromechanical device that allows the user to feed information into the computer for analysis, storage and to give commands to the computer. Data and instructions are entered into the computer's memory through an input device. It captures information and translates it into a form that can be processed and used by the other parts of the computer. After processing the input data, the computer provides the results with the help of output devices. An output device converts machine-readable information into human-readable form. The basic functioning of the output device is just

the opposite of the input device, that is, the data is “fed into” the computer system through the input device while the output is “taken out” from the computer through the output device. However, the output, which comes out from the CPU, is in the form of digital signals (see Figure 4.1). The output devices display the processed information by converting them into graphical, alphanumeric, or audio-visual form.

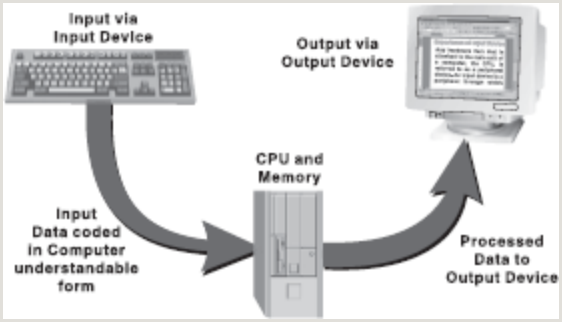


Figure 4.1 Data Processing

4.1.1 Importance of Input/Output Devices

As we know, the processing of the data by the computer system can be viewed as a three-step process:

Step 1	Data input via an input device.
Step 2	Processing of data.
Step 3	Data output via an output device.

Input devices play a major role in the processing of any data via the computer system because the output of the computer is always based on the given input. Generally, data that is given to the input devices is raw. Therefore, it is the function of the input devices to manipulate the raw data and then send them for further processing. The preparation of the computerized input is the initial step in the creation of useful output. This output must be supplied to the outside world, which is done through output devices.

4.2 TYPES OF INPUT DEVICES

Computer accepts input in two ways, either manually or directly. In case of manual data entry, the user enters the data into computer by hand, for example, by using keyboard and mouse. A user can also enter data directly by transferring information automatically from a source document (like from a cheque using MICR) into the computer. The user does not need to enter information manually. Direct data entry is accomplished by using special direct data entry devices like a barcode reader. Some of the commonly used input devices are keyboard, pointing devices like mouse and joystick, speech recognition, digital camera and scanners.

4.2.1 Keyboard

FACT FILE

Qwerty Query

The layout of a keyboard comes in various styles, such as QWERTY, AZERTY and DVORAK. QWERTY is the most common layout in English language computer keyboards. It takes its name from the first six letters shown on the keyboard's top row of letters. Similarly, French language keyboards

use A and Z in place of Q and W and are known as *AZERTY* keyboards.

A *keyboard* is the most common data entry device. Using a keyboard, the user can type text and commands. The keyboard is designed to resemble a regular typewriter with a few additional keys (see Figure 4.2). Data is entered into the computer by simply pressing keys. The layout of the keyboard has changed very little since it was introduced. In fact, the most common change in its technology has simply been the natural evolution of adding more keys that provide additional functionality. The number of keys on a typical keyboard varies from 84 to 104.

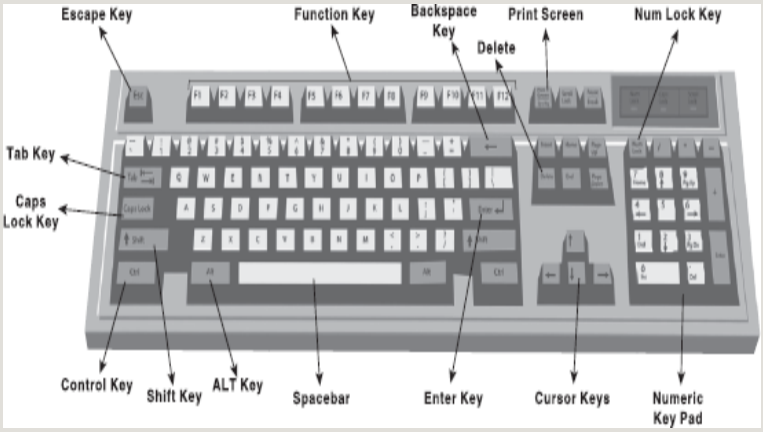


Figure 4.2 Keyboard

Portable computers such as laptops quite often have custom keyboards that have slightly different key arrangements than a standard keyboard. In addition, many system manufacturers add special buttons to the standard layout. A keyboard is the easiest input device,

as it does not require any special skill. Usually, it is supplied with a computer so no additional cost is incurred. The maintenance and operational cost of a keyboard is also less. However, using a keyboard for data entry may be a slow process because the user has to manually type all the text. In addition, it can be difficult for people suffering from muscular disorders.

How does the keyboard work? A keyboard is a series of switches connected to a small keyboard microprocessor that monitors the state of each switch and initiates a specific response to a change in state. When the user presses a key, it causes a change in the amount of current flowing through the circuit associated specifically with that key. The keyboard microprocessor detects this change in current flow. By doing this, the processor can tell when a key has been pressed and when it is being released. Depending upon which key's Figure 4.3 Working of a Keyboard circuit carries a signal to the microprocessor, the processor generates the associative code, known as *scan code*, of the key and sends it to the operating system (see Figure 4.3). A copy of this code is also stored in the keyboard's memory. When the operating system reads the scan code, it informs the same to the keyboard and the scan code stored in keyboard's memory is then erased.

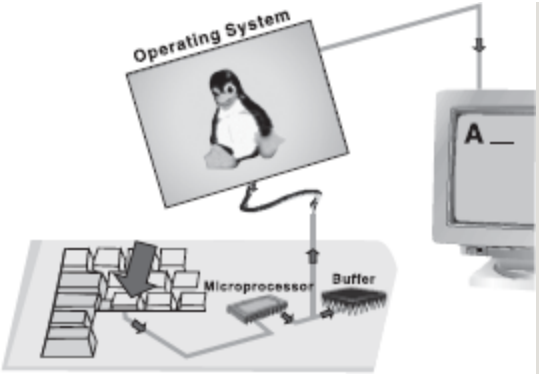


Figure 4.3 Working of a Keyboard

Initially, the processor filters all the tiny current fluctuations out of the signal and treats it as a single key press. If the user continues to hold down a key, the processor determines that the user wishes to send that character repeatedly to the computer. In this process, the delay between each instance of character can normally be set in the operating system, typically ranging from 2–30 characters/second (cps).

4.2.2 Pointing Devices

Most computers come with an alphanumeric keyboard but in some applications, the keyboard is not convenient. For example, if the user wants to select an item from a list, the user can identify that item's position by selecting it through the keyboard. However, this action could be performed quickly by pointing at the correct position. A pointing device is used to communicate with the computer by pointing to locations on the monitor screen. Such devices do not require keying of characters; instead the user can move a cursor on the screen and perform move, click, or drag operations. Some of the commonly used pointing devices are mouse, trackball, joystick, light pen, touch screen and trackpad.

Mouse: A *Mouse* is a small handheld pointing device with a rubber ball embedded at its lower side and buttons on the top. Usually, a mouse contains two or three buttons, which can be used to input commands or information (see Figure 4.4). It may be classified as a *mechanical mouse* or an *optical mouse*, based on the technology it uses. A mechanical mouse uses a rubber ball at the bottom surface, which rotates as the mouse is moved along a flat surface, to move the cursor. It is the most common and least expensive pointing device. An optical mouse uses a light beam instead of a rotating ball to detect movement across a specially patterned mouse pad. As the user rolls the mouse on a flat surface, the cursor on the screen also moves in the direction of the mouse's movement. It is pricier than their mechanical counterparts but are accurate and often do not need a mouse pad.

Figure 4.4 Mouse

FACT FILE

Wireless Keyboard and Mouse

With the increasing use of wireless technology, the wireless versions of keyboard and mouse have also been developed. Rather than connecting through wires, they connect with the computer using one of

the three technologies, namely, Bluetooth, infrared, or radio frequency. The wireless keyboard requires three AA batteries and the wireless mouse requires two AA lithium batteries to operate. They also have a power switch on the bottom to turn them ON or OFF. The use of wireless devices helps in eliminating the wiring tangles around the PC and provides mobility and flexibility to the user to position him/her self relative to the computer.

A mouse allows us to create graphic elements on the screen such as lines, curves and freehand shapes. Since it is an intuitive device, it is easier and convenient to work as compared to the keyboard. Like a keyboard, it is also supplied with a computer; therefore, no additional cost is incurred. However, it needs a flat space close to the computer. The mouse cannot easily be used with laptop (notebook) or palmtop computers. These types of computers need a trackball or a touch sensitive pad called a *touchpad*.

How does a mouse work? As shown in Figure 4.5, a mechanical mouse has a rubber ball at the bottom. When the user moves the mouse along the flat surface, the ball rolls. The distance, direction and speed of the ball's motion is tracked. This information is used by the computer to position the mouse pointer on the screen. Inside the mouse are three rollers. One of them, which is mounted at a 45° angle to the other two, is spring loaded. This roller is usually the smallest of the three. It is there simply to hold the ball against the other two rollers. The other two rollers are usually larger, and of different colour. These rollers are mounted at a 90° angle to one another, one roller measures how fast the ball is turning horizontally, and the other measures how fast it is turning vertically. When the ball rolls, it turns these two rollers. The rollers are connected to axles, and the axles

are connected to a small sensor that measures how fast the axle is turning. Both sets of information are passed to the electronics inside the mouse. This little processor, usually consisting of little more than a single chip, uses the information to determine how fast the mouse itself is moving, and in what direction. This information is passed to the computer via a mouse cord, where the operating system then moves the pointer accordingly.

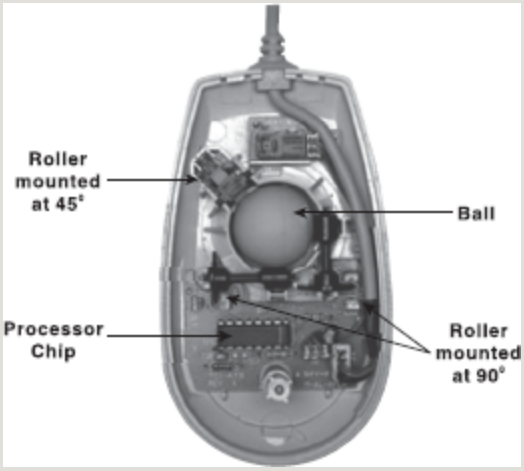


Figure 4.5 Inside a Mechanical Mouse

The optical mouse uses an infrared light and special mouse pads with fine grid lines to measure the rotation of the axle. The axle in optical mouse is connected to a little photo-interrupter wheel with a number of tiny holes in it. In front of this wheel is a light source and on the other side of the wheel is a light metre. As the wheel turns, the light flashes through the holes in the wheel. By measuring how often these flashes occur, the light sensor can measure how fast the wheel is turning and sends the corresponding coordinates to the computer. The computer moves the cursor on the screen based on the

coordinates received from the mouse. This happens hundreds of times each second, making the cursor appear to move very smoothly.

Trackball: A *Trackball* is another pointing device that resembles a ball nestled in a square cradle and serves as an alternative to a mouse. In general, a trackball is as if a mouse is turned upside down (see Figure 4.6). It has a ball, which can be rotated by fingers in any direction, the cursor moves accordingly. The size of the ball in the trackball varies from as large as a cue ball, to as small as a marble. Since it is a static device, instead of rolling the mouse on the top of the table the ball on the top is moved by using fingers, thumbs and palms. This pointing device comes in various shapes and forms but with the same functions. The three shapes, which are commonly used are a ball, button and square.

Figure 4.6 Trackball

Like the mouse, a trackball is also used to control cursor movements and the actions on a computer screen. The cursor is activated when buttons on the device are pressed. However, the trackball remains stationary on the surface, only the ball is moved with the fingers or palm of the hand. By moving just the fingers and not the entire arm, the user can get more precision and accuracy,

which is why many graphic designers and gamers choose to use trackball instead of mouse. In addition, since the whole device is not moved for moving the graphic cursor, a trackball requires less space than a mouse for operation. A trackball, generally, tends to have more buttons. A lot of computer game enthusiasts and graphic designers also tend to choose to have more buttons to cut down on keyboard use. These extra buttons can also be re-programmed to suit whatever functions they require. Trackballs are not supplied normally so an additional cost is always charged. Moreover, before using them, a user has to learn how to use them.

How does a trackball work? A trackball works in the same way as a mouse, with the ball turning rollers, the rollers turning axles, which are in turn connected to either mechanical or optical sensors that measure their rotation. As shown in Figure 4.7, a trackball consists of a number of components. As one moves the trackball, it starts a chain of events inside the box that results in the pointer moving on the computer screen. In a normal trackball, on one side of each encoding wheel is a pair of LEDs (light emitting diode) that emits infrared light. On the opposite side of each pair of LEDs is a light sensor. Every time light from the LEDs shines through a hole in the encoding wheel, a pulse of electricity is sent from the light sensor to the microprocessor. When the trackball rolls side-to-side, the horizontal (x-axis) shaft rotates, spinning the attached encoder wheel. Similarly, when the trackball is rolled up and down, the vertical (y-axis) shaft rotates, spinning the attached encoder wheel. Due to this spinning, the light blinks, which can be detected by the light sensor. The microprocessor counts how many times the light sensors detect light each second and sends this information to the computer along the cord.

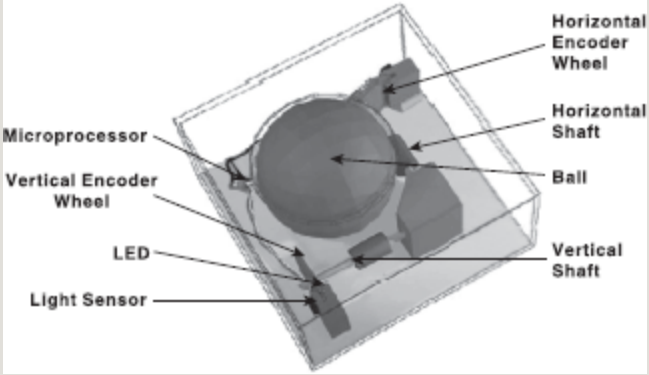


Figure 4.7 Inside a Trackball

Joystick: A *joystick* is a device that moves in all directions and controls the movement of the cursor. The basic design of a joystick consists of a stick that is attached to a plastic base with a flexible rubber sheath. This plastic base houses a circuit board that sits beneath the stick. The electronic circuitry measures the movement of the stick from its central position and sends the information for processing. A joystick also consists of buttons which can be programmed to indicate certain actions once a position on the screen has been selected using stick (see Figure 4.8). It offers three Figure 4.8 Joystick types of control: *digital, glide and direct*. Digital control allows movement in a limited number of directions such as up, down, left and right. Glide and direct control allow movements in all directions (360°). Direct control joysticks have the added ability to respond to the distance and speed with which the user moves the stick.

Figure 4.8 Joystick

A joystick is generally used to control the velocity of the screen cursor movement rather than its absolute position. It is used for computer games. The other applications in which it is used are flight simulators, training simulators, CAD/CAM systems and for controlling industrial robots.

How does a joystick work? Various joystick technologies are available and they differ mainly in how much information they pass on. All joysticks are designed to inform the computer about the positioning of the stick at any given time. This is done by providing the x – y coordinates of the stick. The x-axis represents the side-to-side position and the y-axis represents the forward block position (see Figure 4.9). The circuit board that sits directly underneath the stick carries electricity from one contact point to another. When the joystick is in neutral position, all but one of the individual circuits is broken. Each broken section is covered with a simple plastic button containing a tiny metal disc. When the stick is moved in any direction, it pushes down on one of these buttons, pressing the conductive metal disc against the circuit board. This closes the circuit, that is, it completes the connection between the two wire sections. When the circuit is closed, electricity can flow down a wire from the computer and to another wire leading back

to the computer. When the computer picks up a charge on a particular wire, it knows that the joystick is in the right position to complete that particular circuit. The joystick buttons work exactly the same way. When a button is pressed, it completes a circuit and the computer recognizes a command.

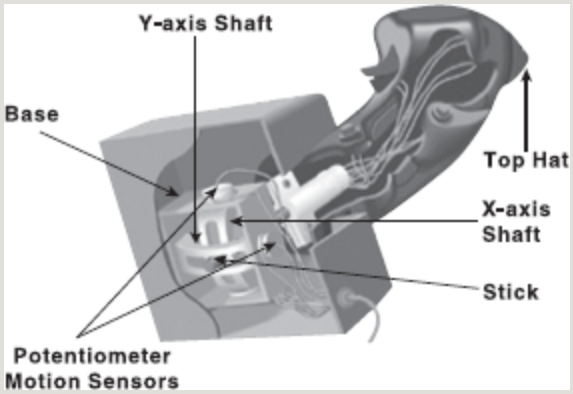


Figure 4.9 Inside a Joystick

Light Pen: A *light pen* (sometimes called *mouse pen*) is a hand-held electro-optical pointing device which a when touched to or aimed closely at a connected computer monitor, will allow the computer to determine where on that screen the pen is aimed. It facilitates drawing images and selects objects on the display screen by directly pointing to the objects. It is a pen-like device, which is connected to the machine by a cable (see Figure 4.10). Although named light pen, it actually does not emit light but its light-sensitive diode would sense the light coming from the screen. The light coming from the screen causes the photocell Figure 4.10 Light Pen to respond by generating a pulse. This electric response is transmitted to the processor that identifies the position

to which the light pen is pointing. With the movement of light pen over the screen, the lines or images are drawn.

Figure 4.10 Light Pen

Light pens give the user the full range of mouse capabilities without the use of a pad or any horizontal surface. Using light pens, users can interact more easily with applications, in such modes as drag and drop, or highlighting. It is used directly on the monitor screen and it does not require any special hand–eye coordinating skills. Pushing the light pen tip against the screen activates a switch, which allows the user to make menu selections, draw and perform other input functions. Light pens are perfect for applications where desk space is limited, in harsh workplace environments, and any situation where fast accurate input is desired. It is very useful to identify a specific location on the screen. However, it does not provide any information when held over a blank part of the screen. A light pen is economically priced and requires little or no maintenance.

How does a light pen work? The light pen contains a lens and a photo detector located in its tip (see Figure

4.11). When the electron beam that sweeps the monitor strikes the phosphor within the light pen's field of view, the light emitted by the phosphor is focused through the lens and onto the photo detector. Due to this, the signal current is increased and is transmitted to the computer. The position of the beam is tracked by the horizontal and vertical counters, which relay this information to a register. This cycle is repeated for every frame produced by the electron beam. By noting when a scan goes by and measuring the interval between scan lines or entire screen refreshes, an accurate position of the photo detector on the screen is determined. The light pen software generates x – y vectors corresponding to a point on the screen, which may be used to make a selection by activating a switch on the light pen.

Figure 4.11 Inside a Light pen

Touch Screen: A *touch screen* is a special kind of input device that allows the direct selection of a menu item or the desired icon with the touch of finger (see Figure 4.12). Essentially, it registers the input when a finger or other object is touched to the screen. It is normally used when information has to be accessed with minimum

effort. However, it is not suitable for input of large amounts of data. Typically, it is used in information-providing systems like hospitals, airlines and railway reservation counters, amusement parks, and so on.

Figure 4.12 Touch Screen

How does a touch screen work? A basic touch screen has three main components: *a touch sensor, a controller* and a *software driver* (see Figure 4.13). The touch sensor/ panel is a clear glass panel with a touch responsive surface. It is placed over a display screen so that the responsive area of the panel covers the viewable area of the video screen. There are several different touch sensor technologies in the market today, each using a different method to detect touch input. These methods are *optical, acoustical* and *electrical*. In the optical method, infrared beams interlace the surface of the screen, and when a light beam is broken, that particular location is recorded. In the acoustical method, ultrasonic acoustic waves pass over the surface of the screen, and when the wave signals are interrupted by some contact with the screen, the location is recorded. In the electrical method, the panel has an electrical current running

through it and touching the screen causes a voltage change, which is used to determine the location of the touch to the screen.

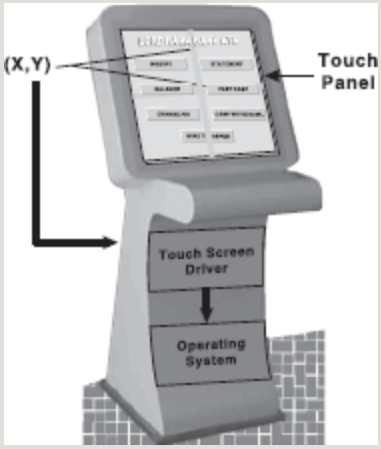


Figure 4.13 Working of Touch Screen

The controller connects the touch sensor and the computer. It takes information from the touch sensor and translates it into information that a computer can understand. The driver is a software update for the computer system that allows the touch screen and computer to work together. It tells the operating system how to interpret the touch event information that is sent from the controller.

Trackpad: Trackpad (also referred to as *touchpad*) is a stationary pointing device that works by sensing the movement of fingers across a small sensitive surface (1.5 or 2 inches) and translating them into the pointer movement on the screen (see Figure 4.14). It is generally used in laptops but can also be connected to a PC through a cord. It is also equipped in personal digital

assistants (PDAs) and media players such as the iPod. Typically, a trackpad also consists of two or three buttons which work as mouse buttons. Many trackpads are also strike sensitive, that is, the user can tap on the trackpad to perform operations like selecting an object, maximizing/ minimizing the window, etc.

Figure 4.14 Trackpad

Note: *The device manufactured by Apple is referred to as Trackpad while the device manufactured by others is known as Touchpad.*

How does a trackpad work? A trackpad consists of several layers: the top layer is the rubber layer on which you move the finger, beneath this layer are two more layers consisting of horizontal and vertical rows of electrodes. The rows of electrodes do not touch each other; rather they are separated by a non-conductive, dialectic material. The layers of electrodes are charged (one with a positive electrical charge and the other with a negative charge) by alternating current (AC), and as a

result an electric field is created between them. The strength of mutual capacitance of the electric field is sampled by the integrated circuits to which the layers of electrodes are connected.

As a finger approaches the top layer of the trackpad, its presence causes the change in capacitances where the electrodes cross over. The capacitance is most affected at the closest intersection point of electrodes under the position where the centre of the finger is touching. By reading the capacitances of closest intersections, the trackpad identifies the cursor position on the screen. These capacitances are measured about 100 times per second. As the finger is moved, the changes in measurements are translated into movement of the cursor on the screen.

4.2.3 Speech Recognition

Speech recognition is one of the most interactive systems to communicate with the computer. The user can simply instruct the computer, with the help of a microphone (along with a speech recognition software), to perform a task (Figure 4.15). It is the technology by which sounds, words or phrases spoken by humans are converted into digital signals, and these signals are transformed into computer-generated text or commands. Most speech recognition systems are speaker-dependent so they must be separately trained for each individual user. The speech recognition system “learns” the voice of the user, who speaks isolated words repeatedly. Then, these voiced words are recognizable in the future.

Figure 4.15 Speech Recognition

Speech recognition is gaining popularity in the corporate world among non-typists, people with disabilities, and business travellers who record information for later transcription. The computer-based speech-recognition systems can be used to create text documents such as letters or e-mails, to browse the Internet, and to navigate among applications by voice commands. They have relatively high accuracy rates. They allow the user to communicate with the computer directly without using a keyboard or a mouse. However, as compared to other input devices, the reliability of the speech recognizer is lesser. Sometimes, it is unable to differentiate between two similar sounding words such as see and sea. It is also not suitable for noisy places.

How does speech recognition work? A speech recognition system consists of a number of components, and together they convert spoken human words into computer text and commands. The system works like this: when a person speaks, the speech recognition software captures the sound through a microphone and converts it into a digital signal. The signals coming out from the microphone are analog waves. These analog waves are converted into digital signals by the computer's sound card. The speech recognition software

analyses the digital pattern to find matches with known sounds contained in a database, and then passes the recognized words to an application such as Microsoft Word or WordPerfect as illustrated in Figure 4.16. Part of that database consists of predefined sound patterns—a one-size-fits-all vocabulary for recognizing speech from as many different voices as possible. The rest is built when a user "trains" the software by repeating keywords so it can recognize the user's distinctive speech patterns.

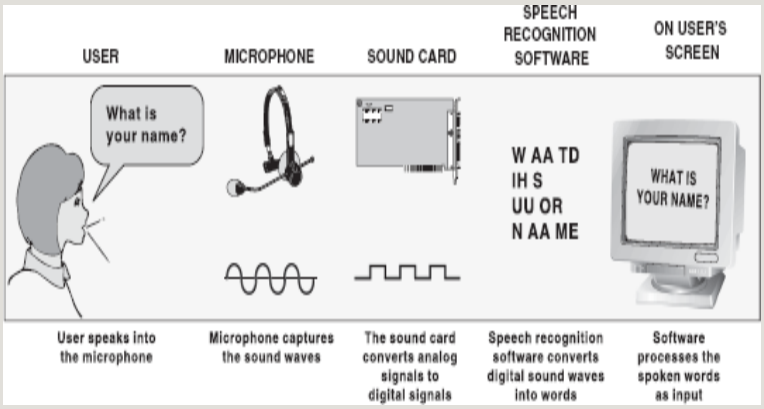


Figure 4.16 Speech Recognition System

4.2.4 Digital Camera

A *Digital camera*, as shown in Figure 4.17, stores images digitally rather than recording them on a film. Once a picture has been taken, it can be transferred to a computer system and then manipulated with an image editing software, and printed. The big advantage of digital cameras is that making photos is both inexpensive and fast because there is no film processing.

Figure 4.17 Digital Camera

How does a digital camera work? The key difference between a digital camera and a film-based camera is that the digital camera does not have a film; instead, it has a sensor that converts light into electrical charges. The image sensor employed by most digital cameras is a charge-coupled device (CCD). Some low-end cameras use complementary metal oxide semiconductor (CMOS) technology. The CCD is a collection of tiny light-sensitive diodes, which convert photons (light) into electrons (electrical charge). These diodes are called *photosites*. Concisely, each photosite is sensitive to light. The brighter the light that hits a single photosite, the greater the electrical charge that will accumulate at that site. In order to digitize the information, the signal must be passed through an analog-to-digital converter (ADC). The ADC converts that information to binary form and sends it to a digital signal processor (DSP). The DSP adjusts the image details, compresses the information and sends it to the camera's storage medium from where it is transferred to the computer's storage through a cable (Figure 4.18).

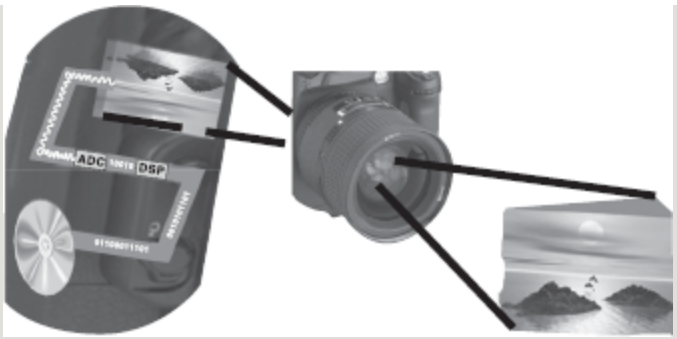


Figure 4.18 Working of a Digital Camera

4.2.5 Webcam

A *webcam* (short form of *web camera*) is a portable video camera, which captures live video or images that may be viewed in real time over a network or the Internet. It is just a small digital camera that is either built in your computer (in most laptops) or can be connected through a USB port (see Figure 4.19). It is normally placed on top of the PC monitor or laptop to capture images of the user while he/she is working on the computer.

Figure 4.19 Webcam

Nowadays, a wide variety of webcams are available, and according to their varied capabilities and features, they are classified into two categories, namely, *streaming* and *snapshot*. A streaming webcam captures moving images (about 30 images per second), thus creating a streaming video—a web video that plays on the computer immediately as its data arrive via network; the recipient need not download the video. However, a high-speed Internet connection is needed to transfer the video smoothly, and the image quality is also comparatively poor. On the other hand, a snapshot webcam captures only still images (usually, once every 30 seconds) and refreshes it continuously. It produces better quality images and is easier to configure than streaming videos.

THINGS TO REMEMBER

Webcam vs. Digital Camera

A digital camera contains a memory card to store the images while a webcam does not have any in-built

memory; it just captures the images and sends them over the network immediately. Also, a webcam is designed to capture low-resolution images (about 10 times less than a digital camera) so that they can be transmitted through the network easily and quickly.

The popularity of webcams is increasing everyday due to their unlimited uses. The most popular use of webcams is in videoconferencing to provide real-time communication where groups of people can see and interact with each other. It can be used with various messenger programs like Yahoo and Windows Live Messenger where you can share your videos while instant messaging with somebody. It is also being used in educational institutions to conduct distance-learning activities; one can attend the classes sitting at home only.

Webcams are cheap, compact and are easy to install and use. They are affordable because of their low manufacturing cost. However, a major drawback of using webcams is that they produce only real-time images and cannot be used unless attached with the PC. Some webcams also comprise advanced features such as automatic lightning controls, automatic face tracking and autofocus, which increase their cost.

4.2.6 Scanners

There are a number of situations when some information (picture or text) is available on paper and is needed on the computer for further manipulation. A *scanner* is an input device that converts a document into an electronic format that can be stored on the disk. The electronic image can be edited, manipulated, combined and printed by using the image editing software. Scanners are also called *optical scanners* as they use a light beam to scan the input data.

Note that most scanners come with a utility program that allow them to communicate with the computer and save the scanned images as a graphic files on the computer. Moreover, they can store images in both greyscale and colour mode. The two most common types of scanners are *hand-held scanners* and *flatbed scanners*.

Hand-held Scanner: A *hand-held* scanner consists of LEDs, which are placed over the document to be scanned (Figure 4.20). This scanner performs the scanning of the document very slowly from the top to the bottom with its light on. In this process, all the documents are converted and then stored as images. While working, the scanner is dragged very steadily and carefully over the document at a constant speed without stopping or jerking in order to obtain best results. Hand-held scanners are widely used where high accuracy is not of much importance. The size of the hand-held scanners is small. They come in various resolutions, up to about 800 dpi (dots per inch) and are available in either greyscale or colour. Furthermore, they are used when the volume of the documents to be scanned is low. These devices read the data on price tags, shipping labels, inventory part numbers, book ISBNs and so on.

Figure 4.20 Hand-held Scanner

How does a hand-held scanner work? When a hand-held scanner's scan button is pressed, an LED illuminates the document underneath it. An inverted angled mirror directly over the scanner's window reflects the image onto the scanner's lens, which is located at the back of the scanner. The lens focuses a single line of the image onto a Charged Coupled Device (CCD), which contains a row of light detectors. As the light shines through these detectors, each of them records the amount of light as a voltage that corresponds to white, black and grey or to a colour. These voltages are sent to a specialized analog chip, which corrects any colour detection error. After that, a single line image is passed to an Analog to Digital Converter (ADC), which converts the analog signals into binary forms that can be sent to the computer (see Figure 4.21). The converter clears itself of the data so that it can receive the next line of the image.

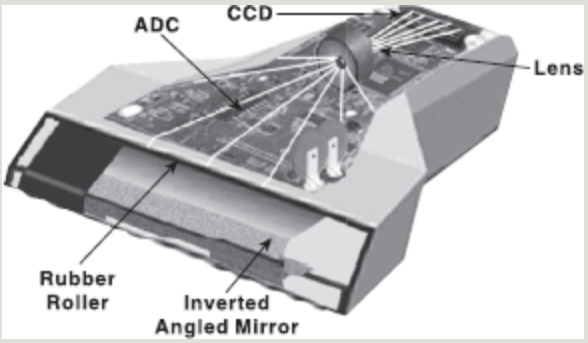


Figure 4.21 Inside a Hand-held Scanner

Flatbed Scanner: A *flatbed scanner* looks similar to a photocopier machine. It consists of a box containing a glass plate on its top and a lid that covers the glass plate (Figure 4.22). This glass plate is used for placing the

document to be scanned. The light beam is placed below the glass plate and when it is activated, it moves horizontally from left to right. After scanning one line, the light beam moves in order to scan the next line and the procedure is repeated until all the lines are scanned. Scanning an A4 size document takes about 20 seconds. These scanners can scan black and white as well as colour images. Flatbed scanners are larger in size and more expensive than hand-held scanners. However, they usually produce better quality images because they employ better scanning technology.

Figure 4.22 Flatbed Scanner

How does a flatbed scanner work? To scan a document, it is placed on the glass plate and the cover is closed. A lamp is used to illuminate the document. The scan head (mirrors, lens, filter and CCD array constitutes a scan head) is moved slowly across the document by a belt that is attached to a stepper motor. The head is attached to a stabilizer bar to ensure that there is no wobble or deviation in the pass. In scanning terms, a pass means that the scan head has completed a single complete scan of the document. The image of the document is reflected by an angled mirror to another mirror. Each mirror is slightly curved to focus the image it reflects onto a smaller surface. The last mirror reflects

the image onto a lens. The lens focuses the image through a filter on the CCD array (Figure 4.23).

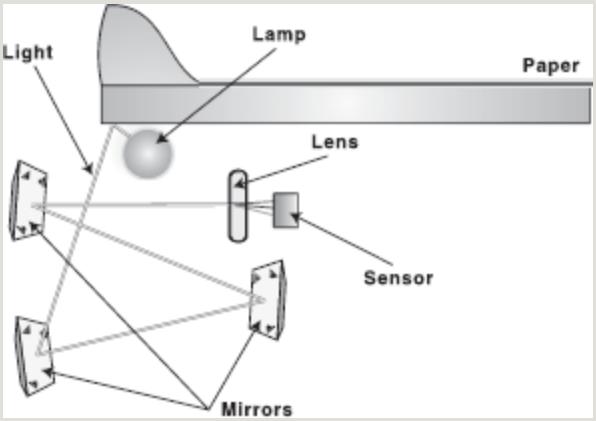


Figure 4.23 Working of a Flatbed Scanner

Some scanners use a three-pass scanning method. Each pass uses a Figure 4.23 Working of a Flatbed Scanner different colour filter (red, green or blue) between the lens and CCD array. After the three passes are completed, the scanner software assembles the three filtered images into a single full-colour image. Nowadays, most scanners use the single-pass method. The lens splits the image into three smaller versions of the original image. Each smaller version passes through a colour filter (either red, green or blue) onto a discrete section of the CCD array. The scanner combines the data from the three parts of the CCD array into a single full-colour image, which is then sent to the computer.

4.2.7 Optical Character Recognition

As stated earlier, a scanner converts an input document into an electronic format that can be stored on the disk. If the document to be scanned contains an image, it can

be manipulated using image editing software. However, if the document to be scanned contains text, you need an *optical character recognition (OCR)* software. This is because when the scanner scans a document, the scanned document is stored as a bitmap in the computer's memory. The OCR software translates the bitmap image of text to the ASCII codes that the computer can interpret as letters, numbers and special characters.

Because of OCR, data entry becomes easier, error-free and less time consuming. However, it is very expensive and if the document is not typed properly, it will become difficult for the OCR to recognize the characters. Furthermore, except for tab stops and paragraph marks, most document formatting is lost during text scanning. The output from a finished text scan will be a single column editable text file. This text file will always require spell checking and proof reading as well as re-formatting to get the desired final layout.

How does an OCR work? All OCR systems include a scanner for reading text and sophisticated software for converting the text into machine-readable form (Figure 4.24). During OCR processing, the text is analysed for light and dark areas in order to identify each alphabetic letter or numeric digit. When a character is recognized, it is converted into an ASCII code. There are two basic methods used for OCR: *matrix matching* and *feature extraction*. The matrix matching technique compares what the OCR scanner sees as a character with a library of character matrices or templates. When an image matches one of these prescribed matrices of dots within a given level of similarity, the computer labels that image as the corresponding ASCII character. Feature extraction OCR does not require strict matching to prescribed templates. This method varies depending on how much “computer intelligence” is applied by the manufacturer.

The computer looks for general features such as open areas, closed shapes, diagonal lines and line intersections. This method is much more versatile than matrix matching. At the end of the OCR processing, the final information can be saved in a number of different formats—text or rich text format (RTF).

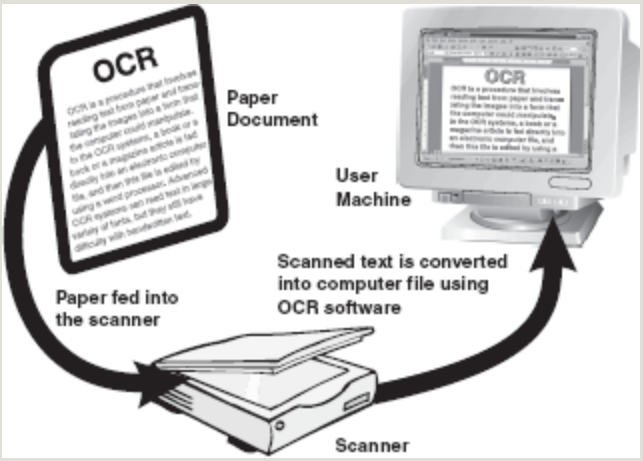


Figure 4.24 An OCR System

4.2.8 Optical Mark Recognition

Optical mark recognition (OMR) is the process of detecting the presence of intended marked responses. A mark registers significantly less light than the surrounding paper. Optical mark reading is done by a special device known as *optical mark reader*. In order to be detected by the OMR reader, a mark has to be positioned correctly on the paper and should be significantly darker than the surrounding paper. The OMR technology enables a high-speed reading of large quantities of data and transferring this data to a computer without using a keyboard. Generally, this technology is used to read answer sheets (objective type tests). In this method, special printed forms/documents

are printed with boxes, which can be marked with a dark pencil or ink (Figure 4.25). These forms are then passed under a light source and the presence of dark ink is transformed into electric pulses, which are transmitted to the computer.

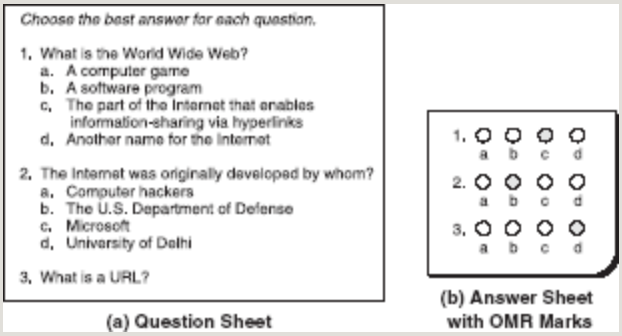


Figure 4.25 Questionnaire Using OMR Marks

OMR has a better recognition rate than OCR because fewer mistakes are made by machines to read marks than in reading handwritten characters. Large volumes of data can be collected quickly and easily without the need for specially trained staff. Usually, an OMR reader can maintain a throughput of 1500 to 10,000 forms per hour. However, the designing of documents for optical mark recognition is complicated and the OMR reader needs to be reprogrammed for each new document design. OMR readers are relatively slow because the person putting marks on the documents must follow the instructions precisely. Any folding or dirt on a form may prevent the form from being read correctly. In addition, it requires accurate alignment of printing on forms and needs a paper of good quality.

How does an OMR work? To make an OMR system work, any of the following methods of mark reading can be used.

- The first method is based on the conductivity of graphite in order to determine the presence of a pencil mark. The marks must be made only in pencil because the number of magnetic particles in the lead pencils is large.
- The second method is based on the reflection of light. In this, a thin beam of light is directed on the surface of the paper. When a lesser amount of light is transmitted through the dot, the filled box can be recognized. OMR can evaluate only those documents which are printed with the marked positions in the specified areas.

OMR is traditionally performed using the reflective light method where a beam of light is reflected on a sheet with marks to capture the reflection (presence of a mark) or absence of reflection (absence of a mark). The OMR data entry system converts the information about the presence or absence of marks into a computer data file. A simple pen or pencil mark is made on the form to indicate each selected response such as answers to survey questions. The completed forms are scanned by an optical mark reader, which detects the presence of a mark by measuring the reflected light. The OMR reader then interprets the pattern of marks into a data record and sends this to a computer for storage, analysis and reporting.

4.2.9 Magnetic-ink Character Recognition

You must have seen special magnetic encoding using characters printed on the bottom of a cheque (Figure 4.26). The characters are printed using special ink, which contains iron particles that can be magnetized. To recognize these magnetic ink characters, a *Magnetic ink character reader (MICR)* is used. It reads the characters by examining their shapes in a matrix form and the information is then passed on to the computer.

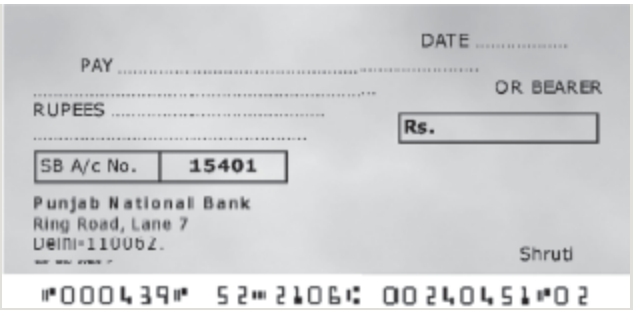


Figure 4.26 Cheque Number Written in MICR Font

The banking industry prefers MICR to OCR as MICR gives extra security against forgeries such as colour copies of payroll cheques or hand-altered characters on a cheque. If a document has been forged, say a counterfeit check produced using a colour photocopying machine, the magnetic-ink line will either not respond to magnetic fields, or will produce an incorrect code when scanned using a device designed to recover the information in the magnetic characters. The reading speed of the MICR is also higher. This method is very efficient and time saving for data processing.

4.2.10 Bar Code Reader

Bar code is a machine-readable code in the form of a pattern of parallel vertical lines of varying widths. It is commonly used for labelling goods that are available in super markets and numbering books in libraries. This code is sensed and read by a *bar code reader* using reflective light (Figure 4.27). The information recorded in the bar code reader is then fed into the computer, which recognizes the information from the thickness and spacing of bars. Bar code readers are either hand-held or fixed-mount. Hand-held scanners are used to read bar codes on stationary items. With fixed-mount scanners, items having a bar code are passed by the scanner by

hand, as in retail scanning applications or by conveyor belts in many industrial applications.

Figure 4.27 Bar Code Reader

FACT FILE

Bar Code Data

The bar code data is just reference numbers, which the computer use to look up associated record file(s), which contain descriptive information. For example, the bar codes found on food items do not contain the price or other description; instead the bar code has a product number in it. When read by a bar code reader and transmitted to the computer, the computer finds the disk record file(s) associated with that item number. This file contains the price, vendor name and other information.

Bar code data correction systems provide enormous benefits for just about every business with a bar code data-collection solution; capturing data is faster and

more accurate. A bar code scanner can record data five to seven times faster than a skilled typist. A bar code data entry has an error rate of about 1 in 3 million. Bar coding also reduces cost in terms of labour and revenue losses resulting from data collection errors. Bar code readers are widely used in supermarkets, department stores, libraries and other places. You must have seen bar code on the back cover of certain books and greeting cards. Retail and grocery stores use a bar code reader to determine the item being sold and to retrieve the item price from a computer system.

Bar code scanners are electro-optical systems that include a means of illuminating the symbol and measuring reflected light. The light waveform data are converted from analog to digital, in order to be processed by a decoder, and then transmitted to the computer software. The process begins when a device directs a light beam over a bar code. The device contains a small sensory reading element, called sensor, which detects the light being reflected back from the bar code, and converts light energy into electrical energy. The result is an electrical signal that can be converted into alphanumeric data. The pen in the bar code unit reads the information stored in the bar code and converts it into a series of ASCII characters by which the operating system gets the information stored in the bar code.

4.3 TYPES OF OUTPUT DEVICES

Output is data that have been processed into useful information. It can be displayed or viewed on a monitor, printed on a printer, or listened through speakers or a headset. Generally, there are two basic categories of output: the output which can be readily understood and used by humans, and which is stored on secondary storage devices so that the data can be used as input for further processing. The output which can be easily

understood and used by human beings are of the following two forms:

- **Hard Copy**: The physical form of output is known as hard copy. In general, it refers to the recorded information copied from a computer onto paper or some other durable surface such as microfilm. Hard copy output is permanent and a relatively stable form of output. This type of output is also highly portable. Paper is one of the most widely used hard copy output media. The principal examples are printouts, whether text or graphics from printers.
- Soft Copy: The electronic version of an output, which usually resides in computer memory and/or on disk, is known as soft copy. Unlike hard copy, soft copy is not a permanent form of output. It is transient and is usually displayed on the screen. This kind of output is not tangible, that is, it cannot be touched. Soft copy output includes audio and visual form of output, which is generated using a computer. In addition, textual or graphical information displayed on a computer monitor is also a soft copy form of output.

Based on the hard copy and soft copy outputs, the output devices are classified into hard copy and soft copy output devices. Printers, plotters and microfilms are the most commonly used hard copy output devices while monitors, voice response systems, projectors, electronic whiteboards, and headphones and headsets are some commonly used soft copy output devices.

4.3.1 Printers

Ever since the dawn of computer age, producing printed output on paper has been one of the computer's principal functions. A *printer* prints information and data from the computer onto paper. Generally, the printer prints 80 or l32 columns of characters in each line, and prints either on single sheets or on a continuous roll of paper, depending upon the printer itself. The quality of a printer is determined by the clarity of a print it can produce, that is, its resolution. Resolution is used to describe the sharpness and clarity of an image. The higher the resolution, the better the image. For printers, the resolution is measured in dpi (dots per inch). The more the dpi, the better will be the quality of image. The dots are so small and close together that they project the image as a solid one. If a printer has a resolution of 600

dpi, it means that the printer is capable of printing 360,000 dots per square inch.

Printers are divided into two basic categories: impact printers and non-impact printers. As their names specify, impact printers work by physically striking a head or needle against an ink ribbon to make a mark on the paper. This includes *dot matrix printers, daisy wheel printers and drum printers.* In contrast, *ink-jet* and *laser printers* are non-impact printers. They use techniques other than physically striking the page to transfer ink onto the page.

Dot Matrix Printer: *Dot matrix printer* (also known as the wire matrix printer) uses the oldest printing technology and it prints one character at a time (Figure 4.28). It prints characters and images as pattern of dots. The speed of dot matrix printers is measured in characters per second (cps). Most dot matrix printers offer different speeds depending on the quality of print desired. The speed can vary from about 200 to over 500 cps. The print quality is determined by the number of pins (the mechanisms that print the dots), which can vary from 9 to 24. The more pins per inch, the Figure 4.28 Dot Matrix higher the print resolution. The best dot matrix printers Printer (24 pins) can produce near letter-quality-type image. Most dot matrix printers have a resolution ranging from 72 to 360 dpi.

Figure 4.28 Dot Matirx Printer

Dot matrix printers are inexpensive and have low operating costs. These printers are able to use different types of fonts, different line densities and different types of paper. Many dot matrix printers are bi-directional, that is, they can print the characters from either direction— left or right. The major limitation of the dot matrix printer is that it prints only in black and white. In addition, as compared to printers like laser printers, they produce low to medium quality printing. The image printing ability is also very limited. These printers may not be able to print graphic objects adequately but can handle applications such as accounting, personnel and payroll very well. Dot matrix printers are commonly used in low-cost, low-quality applications like cash registers. These printers are limited to situations where carbon copies are needed and the quality is not too important.

How does a dot matrix printer work? The technology behind dot matrix printing is quite simple as shown in Figure 4.29. The paper is pressed against a rubber-coated cylinder and is pulled forward as printing progresses. The printer consists of an electro-magnetically driven print head, which is made up of numerous print wires (pins). The characters are formed by moving the electro-magnetically driven print head across the paper, which strikes the printer ribbon

situated between the paper and print head pin. As the head stamps onto the paper through the inked ribbon, a character is produced that is made up of these dots. These dots seem to be very small for the normal vision Matrix Printer and appear like solid human-readable characters.

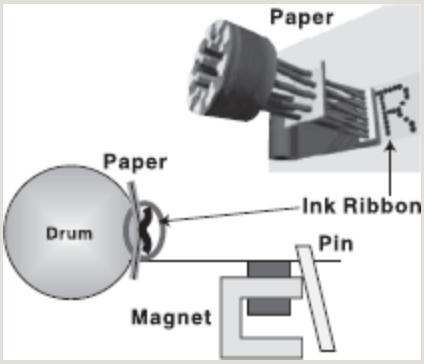


Figure 4.29 Working of Dot Matrix Printer

Daisy Wheel Printer: The major drawback of the dot matrix printer is that the pattern of dots that make up each character is visible on the print produced by it, making it look unprofessional. If you need a printer that can produce professional letter quality documents, you need a *daisy wheel printer*. The daisy wheel printer is named so because the print head of this printer resembles a daisy flower, with printing arms that appear like the petals of the flower (see Figure 4.30). These *printers* are commonly referred to as *letter quality printers* as the print quality is as good as that of a high-quality typewriter.

Figure 4.30 Daisy Wheel Printer

Daisy wheel printers produce high-resolution output and are more reliable than dot matrix printers. They can have speeds up to 90 cps. These printers are also called *smart printers* because of their bi-directional printing and built-in microprocessor control features. However, daisy wheel printers give only alphanumeric output. They cannot print graphics and cannot change fonts unless the print wheel is physically replaced. These printers are usually very slow because of the time required to rotate the print wheel for each character desired. Daisy wheel printers are slower and more expensive than dot matrix printers. However, if the appearance of the correspondence is important and you do not need graphics, a daisy wheel printer is a better choice.

How does a daisy wheel printer work? These printers have print heads composed of metallic or plastic wheels. A raised character is placed on the tip of each of the daisy wheel's “petals”. Each petal has an appearance of a letter (upper case and lower case), number or punctuation mark on it. To print, the print wheel is rotated around until the desired character is under the print hammer. The petal is then struck from behind by the print hammer, which strikes the character, pushing it against the ink ribbon, and onto the paper, creating the character (see Figure 4.31).

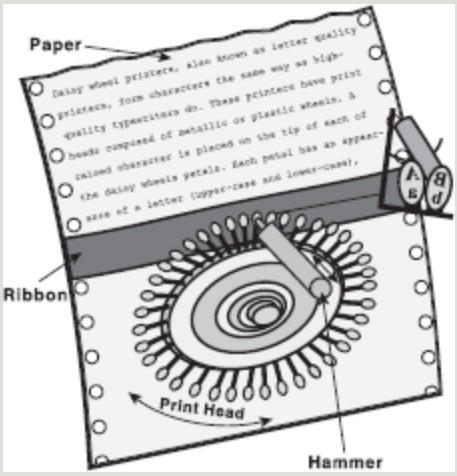


Figure 4.31 Working of Daisy Wheel Printer

Drum Printer: The dot matrix and daisy wheel printers are character or serial printers, that is, they print one character at a time. However, a *drum printer* (shown in Figure 4.32) is a line printer, that is, it can print a line in a single operation. Generally, a line printer is used because of its speed as it uses special tractor-fed paper with pre-punched holes along each side. This arrangement allows a continuous high-speed printing. Its printing speed varies from 300 lines to 2000 lines per minute with 96–160 characters on a 15-inch line. Although such printers are much faster than character printers, they tend to be quite loud, have limited multi-font capability and often produce lower print quality than most recent printing technologies. Line printers are designed for heavy printing applications. For example, in businesses where enormous amounts of materials are printed, the low-speed character printers are very slow; therefore, the users need high-speed line printers. Although, drum printers have high speed of printing, they are very expensive and their character fonts cannot

be changed. Moreover, the strike of the hammer should be precise. A single mistimed strike of the hammer may lead to wavy and slightly blurred printing.

Figure 4.32 Drum Printer

How does a drum printer works? The basics of a line printer, like the drum printer, are similar to those of a serial printer except that multiple hammers strike multiple type elements against the paper almost simultaneously, so that an entire line is printed in one operation. A typical arrangement of a drum printer involves a large rotating drum mounted horizontally and positioned in front of a very wide inked Figure 4.33 Working of Drum Printer ribbon, which, in turn, is positioned in front of the paper itself (see Figure 4.33). The drum contains characters moulded onto its surface in columns around its circumference; each column contains a complete set of characters (letters, digits, etc.) running around the circumference of the drum. The drum spins continuously at high speed when the printer is operating. In order to print a line, hammers positioned behind the paper ram the paper against the ribbon and against the drum beyond it at exactly the right instant, such that the appropriate character is printed in each column as it spins past on the drum. Once every column

has been printed, the paper is advanced upward so that the next line can be printed.

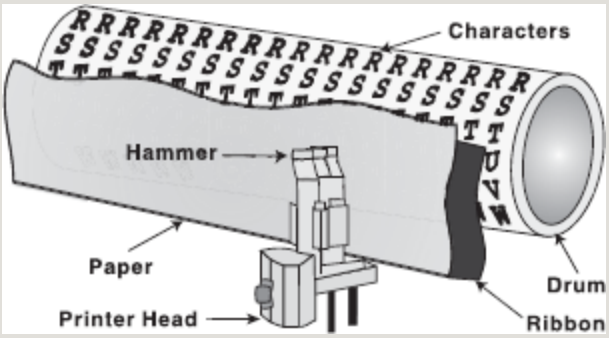


Figure 4.33 Working of Drum Printer

Ink-jet Printer: The most common type of printer found in homes today is the ink-jet printer (see Figure 4.34). An *ink-jet printer* is a printer that places extremely small droplets of ink onto paper to create an image. Being a non-impact printer, it does not touch the paper while creating an image. Instead, it uses a series of nozzles to spray drops of ink directly onto the paper. Inkjets were originally manufactured to print in monochrome (black and white) only. However, the print head has now been expanded and the nozzles increased to accommodate cyan (C), magenta (M), yellow (Y) and black (K). This combination of colours is called *CMYK*. It allows for printing images with nearly Figure 4.34 Ink-jet Printer the same quality as a photo development lab using certain types of coated paper.

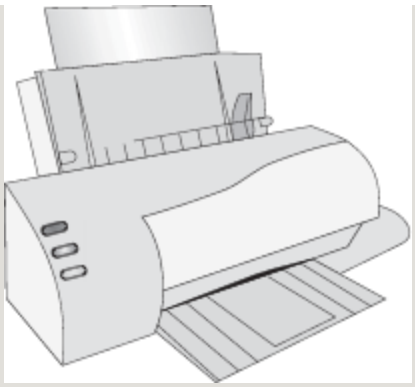


Figure 4.34 Ink-Jet Printer

Ink-jet printers are costlier than dot matrix printers and the quality is much better. These printers can print any shape of character, which a user can specify, as they produce printed output as patterns of tiny dots. This allows the printer to print many special characters, different sizes of prints, and enables it to print graphics such as charts and graphs. Ink-jet printers typically print with a resolution of 600 dpi or more. Due to the high resolution, these printers produce high quality graphics and text printouts. They are also affordable, which appeals to small businesses and home offices. These printers print documents at a medium pace but slow down if printing a document with multi-colours. These printers can print about six pages a minute and can be programmed to print symbols such as Japanese or Chinese characters.

How does an ink-jet printer work? An ink-jet printer places extremely small droplets of ink onto the paper to create a character or an image. It has a print cartridge with a small series of electrically heated chambers. These chambers are attached to the print head with a series of small nozzles that spray ink onto the surface of the paper. As the print head moves back and

forth across the page, the software gives instructions regarding the type and the quantity of colours. It also tells the position where the dots of ink should be "sprayed". There are two main ways to drop the ink droplets, namely, the *bubble-jet* and *piezoelectric* technology (see Figure 4.35). Bubble-jet printers use heat to fire ink onto the paper. There are three main stages with this method. The squirt is initiated by heating the ink to create a bubble until the pressure forces it to burst and hit the paper. The bubble then collapses as the element cools, and the resulting vacuum draws ink from the reservoir to replace the ink that was ejected. Piezoelectric technology uses a piezo crystal at the back of the ink reservoir. It flexes when an electric current flows through it. Therefore, whenever a dot is required, a current is applied to the piezo element, the element contracts and in doing so forces a drop of ink out of the nozzle.

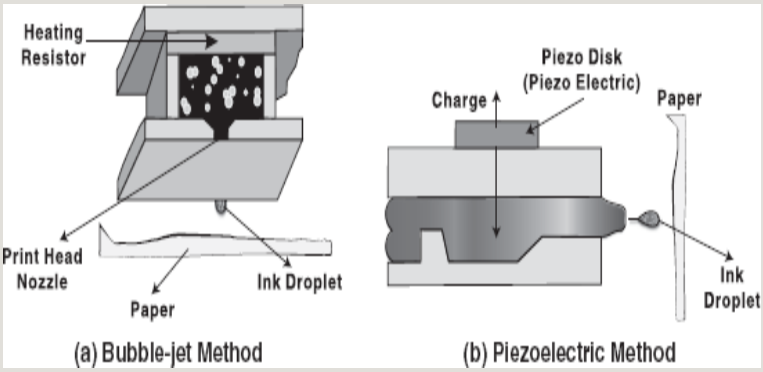


Figure 4.35 Spraying Techniques of Ink-jet Printer

Laser Printer: A *laser printer* provides the highest quality text and images for personal computers today (see Figure 4.36). It is a very fast printer, which operates

on the same principle as that of a photocopy machine. Most laser printers can print text and graphics with a very high quality resolution. They are also known as *page printers* because they process and store the entire page before they actually print it. They produce sharp, crisp images of both text and graphics, providing resolutions from 300 to 2400 dpi. Today, the resolution of most printers is 600 dpi. They are quiet and fast, are able to print 4–32 text-only pages per minute for individual microcomputers and up to 200 pages per minute for mainframes. Laser printers can print in excess of 2000 lines per minute. Furthermore, they can print in different fonts, that is, type styles and sizes. Laser printers are often faster than ink-jet printers but are more expensive to buy and maintain than the other printers. The cost of these printers depends on a combination of costs of paper, toner replacement, and drum replacement. These printers are useful for volume printing because of their speed.

Figure 4.36 Laser Printer

How does a laser printer work? The core component of laser printing system is the photoconductive drum. A rotating mirror inside the printer causes a beam of laser to sweep across the photoconductive drum. Initially, the beam of laser charges the photoconductive drum positively. When the

charged photoconductor is exposed to an optical image through a beam of light to discharge, a latent or invisible image is formed. At the point where the laser strikes the surface of the drum, it creates a dot of positive charge. These points are represented by black dots, which will be printed on the paper. After this, the printer coats the drum with a container, which contains a black powder called *toner*. This toner is negatively charged, and so it clings to the positive areas of the drum surface. When the powder pattern gets fixed, the drum is rotated and the paper is fed into the drum surface via a pressure roller. This pressure roller transfers the black toner onto the paper. Since the paper is moving at the same speed as the drum, the paper picks up the image pattern precisely. Finally, the printer passes the paper through the fuser, a pair of heated rollers. As the paper passes through these rollers, the loose toner powder gets melted and fuses with the fibres in the paper. The paper is then brought out of the printer (see Figure 4.37).

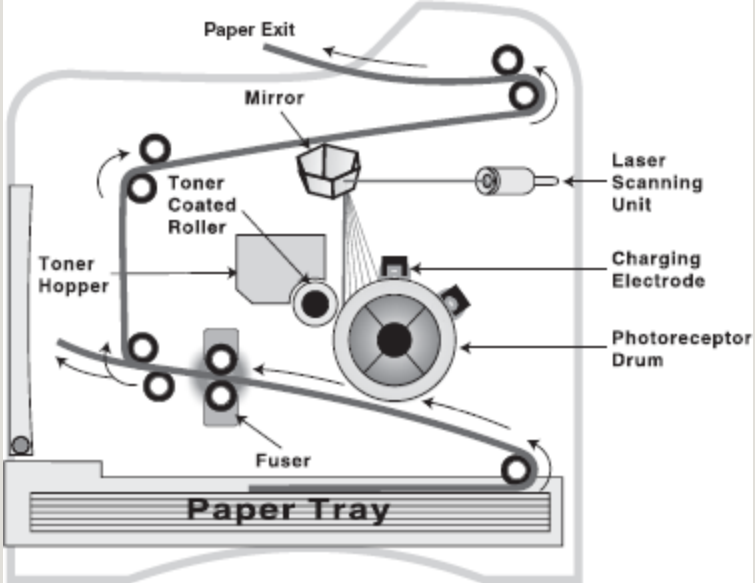


Figure 4.37 Working of a Laser Printer

Hydra Printer: *Hybrid document reproduction apparatus (HYDRA) printer*, popularly known as all-in-one printer, is a device that consolidates the capabilities of multiple devices in one machine (see Figure 4.38). It may include some or all of the devices like printer, scanner, photocopier and fax machine. Apart from these devices, some hydra printers contain memory card slots which facilitate easier printing of photos and also have the in-built wireless capabilities that make Figure 4.38 Hydra Printer sharing of this printer with other systems easier.

Figure 4.38 Hydra Printer

Following are the features that must be considered to evaluate these printers:

- Print speed
- Maximum resolution
- Memory card compatibility
- Scanner resolution
- Fax speed

Hydra printers are useful for small organizations due to their small size, less space requirement and cost effectiveness. The cost of these printers depend on the technology (inkjet or laser) being used. They save power to a great extent as only one power outlet is required for performing various operations. They are easy to install and maintain and have easy-to-use GUI (graphical user interface) that help users to understand their functions easily.

THINGS TO REMEMBER

Raster versus Vector

A grid of pixels defines raster; each pixel is a different colour to make an entire image. Vector graphics, on the other hand, are not constricted to a

grid format. Such graphics are given instructions by the computer about how the objects should be shaped and their relative size, making them resolution-independent. Hence, they are readily scaleable without distortion. Vector images are used for charts and graphs, and other graphics that must have sharp lines when scaled to various sizes. Raster images are used for creating subtle gradations of shades and colour such as in a photograph.

4.3.2 Plotters

A *plotter* is a pen-based output device that is attached to a computer for making vector graphics, that is, images created by a series of many straight lines. It is used to draw high-resolution charts, graphs, blueprints, maps, circuit diagrams and other line-based diagrams. It is similar to a printer, but it draws lines using a pen. As a result, it can produce continuous lines, whereas a printer can only simulate lines by printing a closely spaced series of dots. Multicolour plotter uses different coloured pens to draw different colours. Colour plots can be made by using four pens (cyan, magenta, yellow and black) and need no human intervention to change them.

Being vector-based, a plotter tends to draw much crisper lines and graphics. The lines drawn by these devices are continuous and very accurate. However, the plotter is considered a very slow output device because it requires excessive mechanical movement to plot. Furthermore, it is unable to produce solid fills and shading. Plotters are relatively expensive as compared to printers but can produce more printouts than standard printers. They are mainly used for Computer Aided Design (CAD) and Computer Aided Manufacturing (CAM) applications such as printing out plans for houses or car parts. These are also used with programs like AUTOCAD (computer

assisted drafting) to give graphic outputs. As shown in Figure 4.39, there are two different types of plotters: drum plotter (where the paper moves) and *flatbed plotter* (where the paper is stationary).

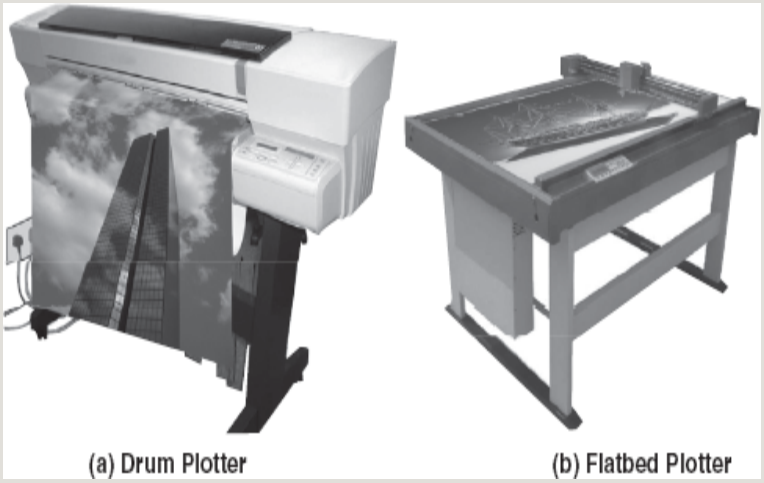


Figure 4.39 Plotters

- **Drum Plotter**: In drum plotters, the paper on which the design is to be printed is placed over a drum. These plotters consist of one or more pen(s) that are mounted on a carriage which is horizontally placed across the drum. The drum can rotate in either clockwise or anticlockwise direction under the control of plotting instructions sent by the computer. In case a horizontal line is to be drawn, the horizontal movement of the pen is combined with the vertical movement of the page via the drum. The curves can also be drawn by creating a sequence of very short straight lines. In these plotters, each pen can have an ink of a different colour to produce multicolour designs. Drum plotters are used to produce continuous output such as plotting earthquake activity or for long graphic output such as tall building structures.
- **Flatbed Plotter:** Flatbed plotters consist of a stationary, horizontal plotting surface on which paper is fixed. The pen is mounted on a carriage, which can move horizontally, vertically, leftwards or rightwards to draw lines. In flatbed plotters, the paper does not move, the pen-holding mechanism provides all the motion. These plotters are instructed by the computer on the movement of pens in the x – y coordinates on the page. These plotters are capable of working on any standard, that is, from A4 size paper to some very big beds. Depending on the size of the flatbed surface, these are used in designing of ships,

aircrafts, buildings and so on. The major disadvantage of this plotter is that it is a slow output device and can take hours to complete a complex drawing.

How does a plotter work? The heart of the plotter is the printer head assembly, consisting of a horizontal bar and the pen in use, attached to the head assembly holding (see Figure 4.40). The pen can be positioned horizontally by moving the pen assembly along the bar. Vertical positioning is achieved by either moving the bar (flatbed plotter) or the paper (drum plotter). Combinations of horizontal and vertical movement are used to draw arbitrary lines and curves in a single action, in contrast to printers, which usually scan horizontally across the page. Plotters create plots by moving a pen under computer control over a drafting paper. The instructions that a plotter receives from a computer consist of a colour and beginning and end coordinates for a line. When an image is to be drawn, a specially designed holder picks up a pen and takes it over to the start position. The pen is pushed down onto the paper and dragged over the surface to produce straight or curved lines. If the product is to be in colour, the pen is then replaced with a new pen. The process continues until the image is complete.

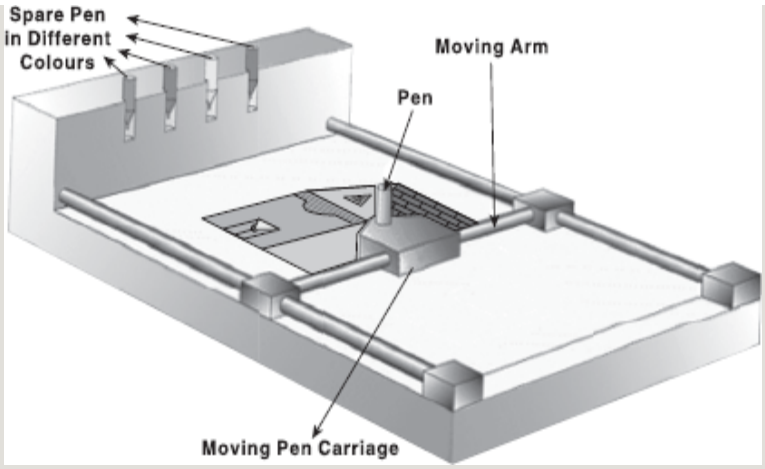


Figure 4.40 Working of Flatbed Plotter

4.3.3 Computer Output Microfilm

Computer Output Microfilm (*COM*) is an extremely high-speed, low-cost process, which records computer-generated information directly from the computer tape or cartridge to a miniaturized microfilm media. The microfilm product is in fiche or roll format, which can be duplicated rapidly and inexpensively. This process can produce data in microfilm form at a highly significant speed from that of a paper printer. The image area of the copy is dramatically reduced, up to 1/40 of its original size; yet, it retains its original clarity. Traditional roll microfilm is 16 mm wide, with a film image that is 1/24 of the size of the original document, often called 24x. Images are normally side-by-side on the film. Therefore, a letter size image is a little over 1/3 inch wide, and a ledger sheet can be well over 1/2 inch. Some cameras and computer output systems create an image 1/42 or 1/48 of the document's original size, with a few systems using even smaller or larger sizes. A jacket microfilm (microfiche), plastic sleeves with small pieces of film, represents individual document (cut from roll microfilm). All the documents for a particular context are placed in a single jacket, much like a file folder. Each 105

× 148 mm jacket typically has five horizontal rows, with an opening at the right end of each row where the film is inserted. Each row can hold 10–15 pages. Microfiche can be indexed in order to locate necessary information quickly and easily and have a life expectancy of more than 500 years. To retrieve the images, the microfilm images may be enlarged on a viewing screen with the help of a *microfilm reader* Figure 4.41 for comfortable reading.

Figure 4.41 Microfilm Reader

COM results in material, space and equipment savings along with mailing costs and information retrieval savings. In addition, hard copy prints can be made without loss of detail when compared with the original document. It facilitates indexing to access information. Furthermore, it provides an inexpensive way to preserve records from a variety of electronic sources. Such records might need to be stored for long periods and referenced only occasionally. The main disadvantage, however, is that it is expensive to install COM, and microfilms cannot be read without the assistance of a special reader device. Moreover, development of secondary-storage techniques such as the use of removable, high-capacity

hard disks had lessened their appeal as output and storage method.

Generally, a COM system is ideal for applications where there is a large amount of information to be retained, and therefore very useful for manuals, industrial catalogues and archives. Microfilm output is used for distribution of airline schedules, medical X-rays and lists of books in print. Banking and insurance companies, government agencies, public utilities and many other types of organizations are regular users of COM.

Creating Microfilms: Microfilms can be created in the following two ways:

- **Photographic Process**: It is the oldest method of creating microfilms. Essentially, it is a form of miniature photography, which uses a microfilm camera for producing microfilms. When the original documents are fed through a microfilm camera, it captures the document into a reduced-size photo. Once all the photographs are taken, the rolls of films are removed from the camera for further processing and development. The processed images can be rolled on film rolls, or cut into sections and loaded to flat microfiche holders.
- **Non-photographic Process**: In this process, the user can entirely avoid the film-developing process. This process uses a monitor and the computer output is read onto the magnetic tape. After this, data are printed on microfilm by using a microfilm recorder. Moreover, it is possible to couple a microfilm machine directly to a computer; therefore, the recorder can directly receive information from the computer. As the recorder projects the computer information onto the screen, a high-speed camera photographs the displayed information.

The process of creating microfilm from magnetic tape is illustrated in Figure 4.42.

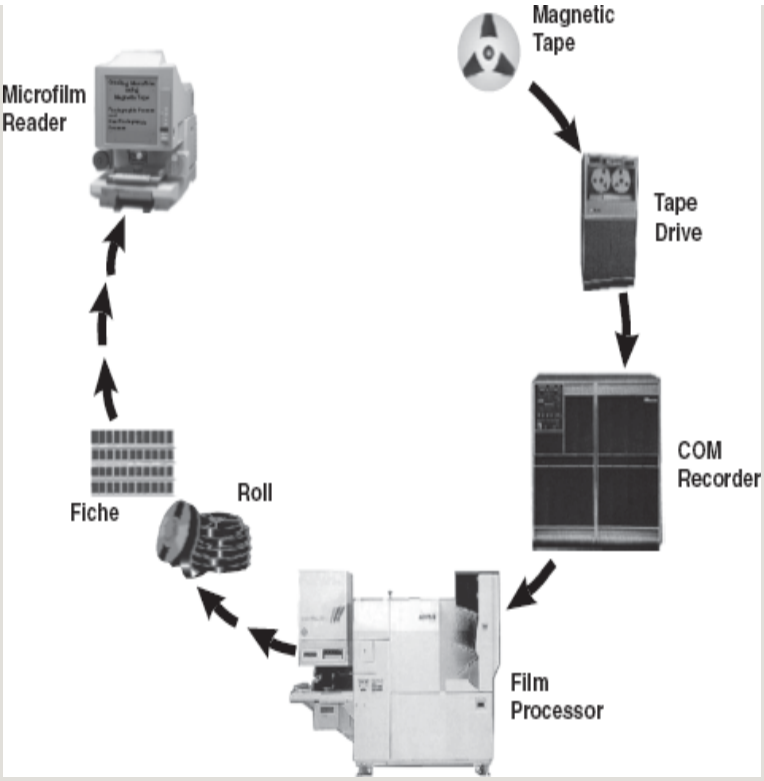


Figure 4.42 Creating Microfilm from Magnetic Tape

4.3.4 Monitor

The monitor is the most frequently used output device for producing soft-copy output. A *computer monitor* is a TV-like display attached to the computer on which the output can be displayed and viewed. The computer monitor can either be a monochrome display or a colour display. A monochrome screen uses only one colour (usually white, green, amber or black) to display text on contrasting background. Colour screens commonly display 256 colours at one time from a selection of over 256,000 choices. Monitors are available in various sizes like 14, 15, 17, 19 and 21 inches. The size of the display is described based on two parameters: aspect ratio and screen size (see Figure 4.43). Aspect ratio is the ratio of the width of the display screen to the height, that is, the ratio of vertical points to the horizontal points necessary

to produce equal-length lines in both directions on the screen. Generally, computer displays have an aspect ratio of 4:3. Like televisions, screen sizes are normally measured diagonally (in inches), the distance from one corner to the opposite corner.

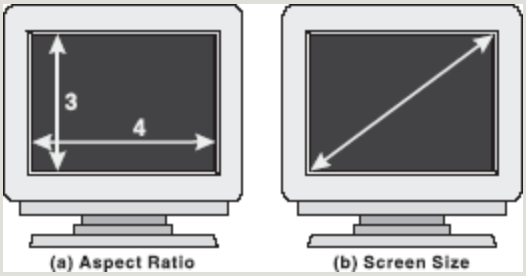


Figure 4.43 Aspect Ratio and Screen Size

Sometimes, while watching television, you may notice that the picture looks a bit blurred. The reason behind this is that the displayed image is not solid but is created by the configurations of dots. These dots are known as picture elements, pels, or simply pixels. The golden rule of a sharp image is that the more the pixels, the sharper the picture. The screen clarity depends on three basic qualities:

- **Resolution**: It refers to the number of pixels in the horizontal and vertical directions on the screen. In medium-resolution graphics, pixels are large, whereas in high-resolution graphics, pixels are small. The average CRT display is currently 800 × 600 or 1024 × 768. The more dots, or pixels, available to create the image, the sharper it will be. Therefore, a resolution of 1024 × 768 will produce sharper images (for example, smaller icons and more information) than one of 640 × 480.
- **Dot Pitch:** It is the measurement of the diagonal distance between two like-coloured (red, green or blue) pixels on a display screen. It is measured in millimetres and common dot pitches are .51 mm, .31 mm, Pixel .28 mm, .27 mm, .26 mm and .25 mm. The Figure 4.44 Dot Pitch smaller the dot pitch, the sharper will be the image when displayed on the monitor. Generally, a dot pitch of less than .31 mm provides clear

images. Multimedia and desktop-publishing users typically use .25 mm dot-pitch monitors (see Figure 4.44).

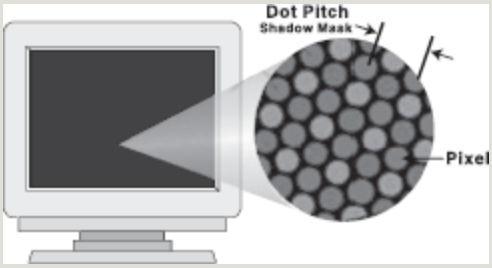


Figure 4.44 Dot Pitch

- **Refresh Rate**: It is the number of times per second the pixels are recharged so that their glow remains bright. Normally, screen pixels are made from phosphor. An electron beam strikes the phosphor and causes it to emit light, resulting in the display of the image. However, it needs to be Refreshed periodically because the phosphors hold their glow for just a fraction of a second. The Refresh rate for a monitor is measured in Hertz (Hz) and varies from 60 to 75 Hz. A Refresh rate of 60 Hz means image is redrawn 60 times a second. The higher the Refresh rate, the more solid the image looks on the screen, that is, it does not flicker.

Colour Depth: *Colour depth*, also referred to as bit depth, refers to the number of bits assigned to each pixel in the image and the number of colours that can be created from those bits. In simple words, it refers to the number of colours that a monitor can display. Different colour depths depend on the amount of display memory dedicated to each pixel. One byte is used to represent 256 colours for each pixel, 16 bits (or 2 bytes) per pixel allows up to 65535 colours, and 24-bit (or 3 bytes) colour can display 16.8 million different colours per pixel. 8-bit colour is better known as pseudo colour, 16-bit mode as high colour, and 24-bit mode is called true colour. A video display unit consists of a video card or adapter that is fitted into an expansion slot and a compatible visual display, which is compatible with the video adapter. The combination of the display modes supported by the graphics adapter and the colour capability of the monitor determine how many colours can be displayed.

Cathode Ray Tube Monitors: Nowadays, most computer monitors are based on cathode ray ube (CRT) technology. The basic operation of these tubes is similar to that in television sets. Figure 4.45 illustrates the basic components of a CRT.

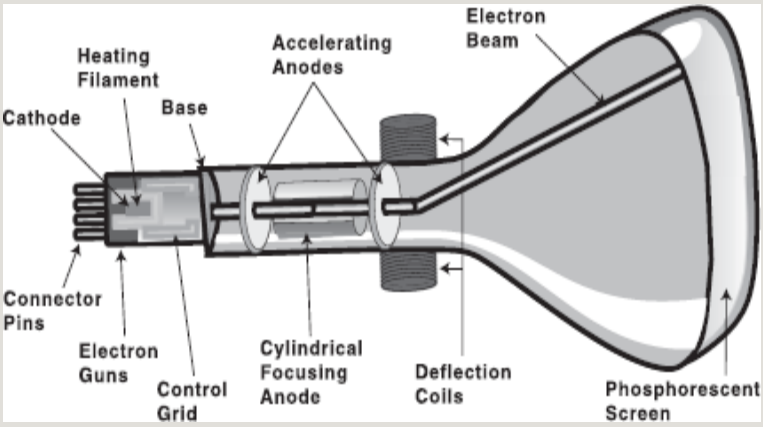


Figure 4.45 Cathode Ray Tube

A beam of electrons (cathode rays) emitted by an electron gun passes through focusing and deflection systems that direct the beam toward specified positions on the phosphor-coated screen. The phosphor then emits a small spot of light at each position contacted by the beam. When the electron beam strikes the phosphors, the light is emitted for a short period of time, this condition is known as persistence. Technically, persistence is defined as the time it takes for the emitted light from the screen to decay to 1/10 of its original intensity. Graphics monitors are usually constructed with persistence in the range of 10–60 microseconds. Since the light emitted by the phosphor fades very rapidly, some method is needed for maintaining the screen picture. One way to keep the phosphor glowing is

to redraw the picture repeatedly by quickly directing the electron beam back over the same points. This type of display is called a Refresh CRT.

The primary components of an electron gun in a CRT are the heated metal cathode and a control grid. Heat is supplied to the cathode by directing a current through a coil of wire, called the filament, inside the cylindrical cathode structure. This causes electrons to be “boiled off ” the hot cathode surface. In the vacuum inside the CRT envelope, the free, negatively charged electrons are then accelerated toward the phosphor coating by a highly positive voltage. The accelerating voltage can be generated with a positively charged metal coating on the inside of the CRT envelope near the phosphor screen, or an accelerating anode can be used, as in Figure 4.45. Note that sometimes the electron gun is built to contain the accelerating anode and focusing system within the same unit.

Before reaching the phosphor-coated screen, the electrons have to be passed through the monitor's focusing system. The focusing system is initially set up to focus the electron flow into a very thin beam and then in a specific direction. Focusing can be accomplished either by electric or by magnetic fields.

When the electrons in the beams collide with the phosphor coating, their kinetic energy is absorbed by the phosphor. Some of this energy is converted into heat while rest of the energy causes the electrons in the phosphors to move up to the higher energy levels. After this, when these electrons begin to return to the ground state, they emit light at certain frequencies. These frequencies are proportionate to the energy difference between the higher state and the ground state. As a result, the image, which we see on the screen, is the combination of all the electron light emissions.

Displaying Graphics on a CRT: Nowadays, two classes of computer graphic display are used: *raster scan* and *random scan* display.

- **Raster Scan Display**: In this system, the electron beam is swept across the screen one row at a time from top to bottom. As the electron beam moves across each row, the beam intensity is turned ON and OFF to create a pattern of illuminated spots. The picture definition is stored in a memory area called the Refresh buffer or frame buffer, which holds the set of intensity values for all the screen points. These values are then retrieved from the Refresh buffer and "painted" on the screen one row (scan line) at a time. At the end of each scan line, the electron beam returns to the left side of the screen to begin displaying the next scan line. The return to the left of the screen, after Refreshing each scan line, is called the horizontal retrace of the electron beam. At the end of each frame, the electron beam returns to the top left corner of the screen to begin the next frame. This is known as vertical retrace (see Figure 4.46). On some raster-scan systems, each frame is displayed in two passes using an interlaced Refresh procedure. In the first pass, the beam sweeps across every other scan line from top to bottom. Then after the vertical retrace, the beam sweeps out the remaining scan lines. Interlacing of the scan lines in this way allows us to see the entire screen displayed in half the time it would have taken to sweep across all the lines at once from top to bottom. This is an effective technique for avoiding flicker, provided that adjacent scan lines contain similar display information.

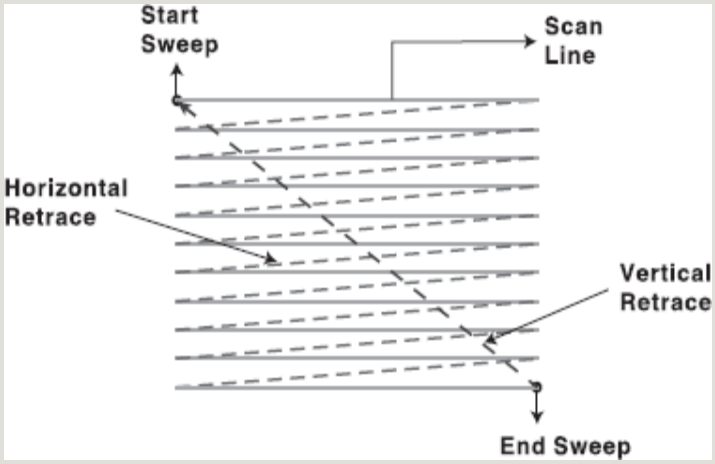


Figure 4.46 Raster Scanning

- **Random Scan Display:** In this system, a CRT has the electron beam directed only to the parts of the screen where a picture is to be drawn. Random scan monitors draw a picture one line at a time, and for this reason are referred to as *vector, stroke-writing*, or *calligraphic displays*. The component lines of a picture can be drawn and

Refreshed by a random scan system in any specified order. Refresh rate on a random scan system depends on the number of lines to be displayed. Picture definition is stored as a set of line-drawing commands in an area of memory referred to as the Refresh display file. To display a specified picture, the system cycles through the set of commands in the display file, drawing each component line in turn. After all line drawing commands have been processed, the system cycles back to the first line command in the list. Random scan systems are designed for line drawing applications. They cannot display realistic shaded scenes.

Colour display on a CRT A CRT monitor displays colour pictures by using a combination of phosphors that emit different coloured light. By combining the emitted light from the different phosphors, a range of colours can be generated. The two basic techniques for producing colour displays with a CRT are:

- **Beam Penetration**: The beam-penetration method for displaying colour pictures is commonly used with random scan displays. In this system, two layers of phosphor (usually red and green) are coated on the inner side of the CRT screen. The displayed colour depends on how far the electron beam penetrates into the phosphor layers. A beam of slow electrons excites only the outer red layer while a beam of very fast electrons penetrates through the red layer and excites the inner green layer. At intermediate beam speeds, combinations of red and green light are emitted to show two additional colours— orange and yellow. The speed of the electrons, and hence the screen colour at any point, is controlled by the beam-acceleration voltage. Beam penetration is an inexpensive way to produce colour in random scan monitors. However, with this method, only four colours are possible, and the quality of pictures is not as good as other methods.
- **Shadow Masking**: Shadow-mask method is commonly used in raster scan systems because it produces a much wider range of colours than the beam-penetration method. A shadow-mask CRT has three phosphor colour dots at each pixel position. One phosphor dot emits a red light, another emits a green light and the third emits a blue light. This type of CRT has three electron guns, one for each colour dot, and a shadow-mask grid just behind the phosphor-coated screen. Three electron beams are deflected and focused as a group onto the shadow mask, which contains a series of holes aligned with the phosphor-dot patterns. When the three beams pass through a hole in the shadow mask, they activate a dot triangle, which appears as a small colour spot on the screen. The phosphor dots in the triangles are arranged so that each electron beam can activate only its corresponding colour dot when it passes through the shadow mask (see Figure 4.47).

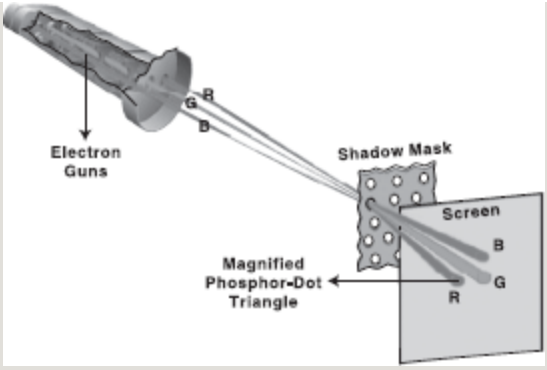


Figure 4.47 Shadow Masking

Colour variations in a shadow-mask CRT can be obtained by varying the intensity levels of the three electron beams. For example, a white (or grey) area is the result of activating all three dots with equal intensity. Yellow is produced with the green and red dots only, magenta is produced with the blue and red dots, and cyan is displayed when blue and green are activated equally.

Liquid Crystal Display Monitors: In the previous section, we discussed the most popular CRT monitors that are used as display devices. With the widespread use of smaller computers like PDAs and laptops, a new type of display, Liquid Crystal Display (LCD), has made a big impact on the computer market. LCD screens have been used for long on notebook computers but are also becoming popular as desktop monitors.

The term *liquid crystal* sounds like a contradiction. We generally conceive a crystal as a solid material like quartz and a liquid as water-like fluid. However, some substances can exist in an odd state that is semi-liquid and semi-solid. When they are in this state, their molecules tend to maintain their orientation like the molecules in a solid, but also move around to different

positions like the molecules in a liquid. Thus, liquid crystals are neither a solid nor a liquid. Manufacturers use this amazing ability of liquid crystals to display images.

An LCD screen is a collection of multiple layers as shown in Figure 4.48. A fluorescent light source, known as the backlight, makes up the rearmost layer. Light passes through the first of two polarizing filters. The polarized light then passes through a layer that contains thousands of liquid crystal blobs aligned in tiny containers called cells. These cells are aligned in rows across the Figure 4.48 Coloured Liquid Crystal Screen screen; one or more cells make up one pixel. Electric leads around the edge of the LCD create an electric field that twists the crystal molecule, which lines the light up with the second polarizing filter and allows it to pass through.

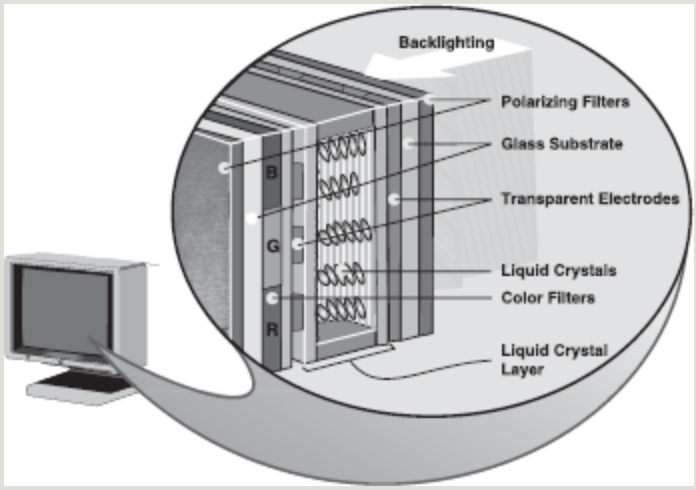


Figure 4.48 Coloured Liquid Crystal Screen

The process illustrated in Figure 4.47 is followed for a simple monochrome LCD. However, colour LCD is more

complex. In a coloured LCD panel, each pixel is made up of three liquid crystal cells. In front of each of these cells, there is a red, green or blue filter. Light passing through the filtered cells creates the colours on the LCD. Nowadays, nearly every colour LCD uses a thin-film transistor (TFT), also known as an active matrix, to activate each cell (see Figure 4.49). TFT-based LCD creates sharp, bright images as compared to previous LCD technologies. The oldest of the matrix technologies, passive-matrix, offers sharp text but leaves "ghost images" on the screen when the display changes rapidly, making it less than optimal for moving video.

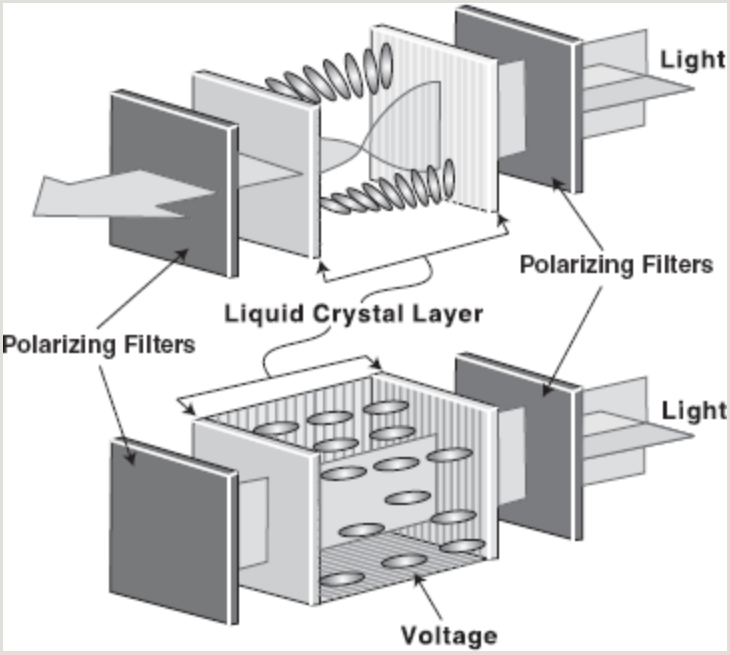


Figure 4.49 Active Matrix LCD

An LCD addresses each pixel individually. As a result, they can create sharper text than CRTs. However, LCD has only one "natural" resolution, limited by the number

of pixels physically built into the display. If you want to move up to, say, 1024 by 768 LCD on an 800 by 600 LCD, you have to emulate it with software, which will work only at certain resolutions.

Differences between LCD and CRT: Although both the CRT and LCD monitors are the most frequently used types of displays in computers, they are worlds apart in terms of what is important when making a purchase decision. Here is a brief discussion of what is different about LCD monitors.

- **Size**: LCD is lightweight and compact, which saves desktop space as compared to a CRT.
- **Resolution**: LCD is designed to work in a single resolution while CRT is designed for many resolutions.
- **Pixel Density**: Pixel density of LCD is generally not as tight as the dot pitch in CRT but for most applications, the density is acceptable.
- **Brightness**: The illuminated phosphor of a CRT is not nearly as bright as what the LCD can produce with its florescent backlight.
- **Power Consumption**: LCD consumes significantly less power than CRT and has a low emission risk. Typically, a LCD consumes approximately half of the power of a typical CRT.
- **Flicker:** With CRT monitors, the goal is to get a faster Refresh rate of at least 85 Hz,but LCD monitors are designed to run at a much slower Refresh rate (usually about 60 Hz) and flicker is never an issue.
- **Pixel Response Time:** The time taken by a pixel to change its state is called pixel response time. CRT has extremely fast pixel response time but LCD can be quite slow. As a result, the user might see “ghost images” when there is movement on the screen. The fastest LCD monitors today have a response time of 25 milliseconds, which is still only about half the speed of a CRT monitor.
- **Viewing Angle**: CRT can be viewed at almost any angle but LCD is best viewed “head on”. Even when viewing an LCD head on, narrow viewing angles can appear to have inconsistent colour and brightness.
- **Viewing Area:** The viewing area of a CRT is usually less than its advertised area. Most 19-inch CRT monitors, for example, typically have about 18 inches of viewable area. However, the LCD monitors are measured exactly, that is, if a LCD monitor is advertised as 17.4 inch, it is the same.
- **Cost:** Prices for LCD screens are quite high but they are coming down. They are still much more costlier than CRT.

Other Types of Monitors: Although CRT and LCD displays are the most commonly used, other kinds of displays also exist. Some of them are as follows:

- **Plasma Display**: A plasma display (also known as thin-panel) is constructed by filling a gas (neon or xenon) between the two glass plates. One glass plate consists of vertical conductors and the other consists of horizontal conductors. When the voltage is applied to the pair of horizontal and vertical conductors, the gas at the intersection of two conductors glows. The advantage of plasma display is that it has the capability of producing more than 16 million colours and is lighter and less bulky than CRT. However, it is heavier than LCD display and requires more power.
- **Thin-film Electroluminescent Display:** This display is similar to the plasma display except the fact that the space between the glass plates is filled with phosphorescent substance (usually zinc sulfide doped with manganese) instead of gas. When high voltage is applied to the crossing conductors, the phosphoric substance at the intersection of two conductors becomes conductor and generates the electrical energy. This energy is absorbed by the manganese atoms and then released as a spot of light. The main problem with thin-film electroluminescent displays is that they require more power.

4.3.5 Voice Response System

Previously, we discussed that voice (in speech recognition) can be taken as an input by the computer system. Similarly, the computer can also give output in the form of audio. Voice response system has an audio-response device which produces audio output. These sounds are pre-recorded in a computer system. Each sound has a unique code. Whenever an enquiry is sought from the system, the computer responds in digital form which is sent to voice devices that unscramble the digital information and produces sound messages to the requesting computer.

There are two basic approaches to get a computer to talk to the user. The first is *synthesis by analysis*, in which the device analyses the input of an actual human voice speaking words, stores and processes the spoken sounds, and reproduce them as needed. The process of storing words is similar to the digitizing process we discussed while considering voice input. In essence, synthesis by analysis uses the computer as a digital tape recorder. The second approach to synthesizing speech is synthesis by rule, in which the device applies a complex set of linguistic rules to create artificial speech. Synthesis based on the human voice has the advantage of sounding more

natural but is limited to the number of words stored in the computer.

The standard computer system can provide audio output with the addition of two components: *a speech synthesizer* that does the speaking and a screen reading software that tells the synthesizer what to say (see Figure 4.50). The synthesizer can be a computer card inserted into the computer or software that works with the computer's sound card. Screen reader is a software program that allows the user to control the synthesizer so that he can access text, which is displayed on the computer monitor, with combinations of keystrokes. Some screen review softwares use the standard computer keyboard keys, some use the numeric keypad, and some use special external keypads. A visually impaired user can use a screen reader to read anything on the screen. The screen reading software can even notify a user about various computer messages that “pop up” on the monitor from time to time.

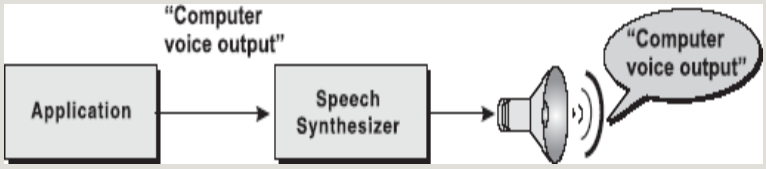


Figure 4.50 Audio Output

Voice output has become common in many places like airlines, bus terminals, banks and brokerage houses. It is typically used when an inquiry is followed by a short reply (such as a bank balance or flight time). Many businesses have found other creative uses for voice output. For example, the telephone. Automatic telephone

voices take surveys, inform customers that catalogue orders are ready to be picked up, and perhaps remind consumers that they have not paid their bills. Moreover, people with a total or partial speech handicap are faced with the problem of communicating their wishes and needs to others. Speech output systems are a valuable aid in this case.

4.3.6 Projector

Projectors have an important place in history, as its growth led to the first moving pictures of the late 19th century, and soon to television. They have allowed the world to share greater information pertaining to culture, technology, medicine and community. Screen image projector is an output device, which is used to project information from a computer on to a large screen, so that it can be simultaneously viewed by a large group of people. Projection display is commonly used for classrooms training or conference room or for multimedia presentations with a large audience. Like monitors, projectors provide a temporary, soft copy output.

Types of Projectors: Projectors are mainly of two types:

- **LCD Projectors**: LCD is the established technology used by most of the leading manufacturers. Most of today's LCD projectors contain three separate LCD glass panels, one each for red, green and blue components of the image signal being fed into the projector. As light passes through the LCD panels, individual pixels can be opened to allow light to pass, or closed to block the light. This activity modulates the light and produces the image that is projected onto the screen (see Figure 4.51).

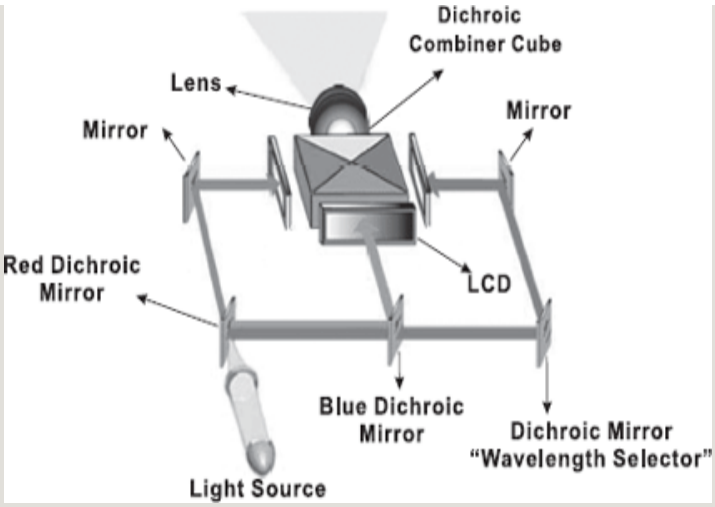


Figure 4.51 LCD Projector Display

- **DLP Projectors**: DLP is the acronym for digital light processing. It is a proprietary technology developed by Texas Instruments. DLP uses a single digital mirror device (DMD) chip that has thousands of tiny mirrors, each representing a single pixel. These mirrors tilt back and forth, directing the light either into the lens path to turn the pixel ON, or away from the lens path to turn it OFF and create the image (see Figure 4.52). DLP is a newer technology than LCD and is used on some of the smallest, lightest projectors currently available. DLP projectors handle video images extremely well.

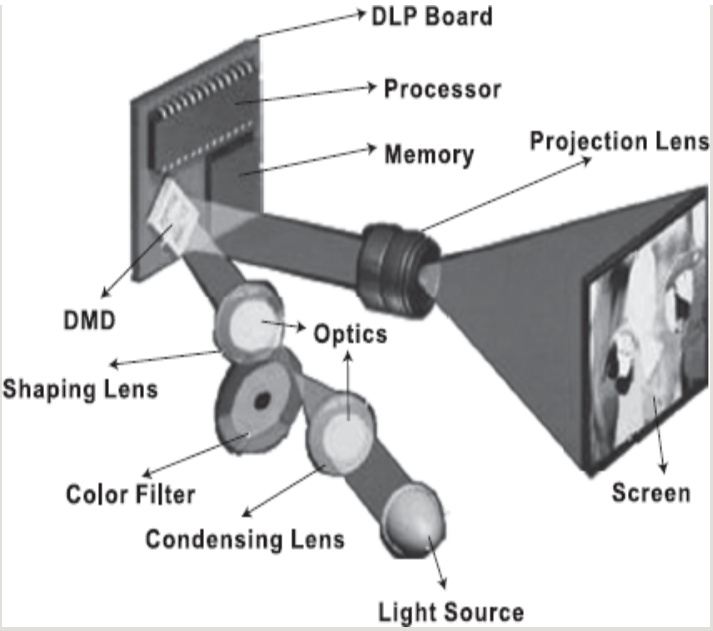


Figure 4.52 DLP Projector Display

4.3.7 Electronic Whiteboard

Electronic whiteboard is an interactive presentation device that works electronically to display digital images, drawings and text in different colours. The whiteboard is generally mounted on the wall and is connected to a computer (through USB, serial port or wireless technology like Bluetooth) and projector. The projector projects the computer's desktop on the whiteboard and user can interact with it using electromagnetic pens, fingers or other pointing devices (see Figure 4.53). The electromagnetic pens can be configured to any Figure 4.53 Electronic Whiteboard colour or width.

Figure 4.53 Electronic Whiteboard

A user can activate programs, applications and menus as well as enter text using either the on-screen keyboard or handwriting recognition utility. Nowadays, most whiteboards come with software that provides various tools and features to create virtual versions of paper flipcharts with pen, highlighter, etc., enhancing the use of electronic whiteboard to a great extent. Electronic whiteboard is widely used for various purposes such as to display presentations, for teaching in classrooms, in corporate meetings, professional sports coaching and so on.

An interactive electronic whiteboard can be classified into one of the following categories:

- **Touch-based Whiteboard**: A touch-based whiteboard allows user to use their fingers or other pointing devices to interact with it. The working on these whiteboards is easy as anything can be used as a pointer, for example, a stick, finger, etc., where the touch of a finger on the whiteboard is considered as a left click of the mouse.
- **Pen-based Whiteboard**: A pen-based whiteboard mainly uses an electromagnetic pen that alters electrical signals produced by the board and does not require batteries or power source for its operation. When the tip of the pen is pressed to the whiteboard, it results in the activation of a switch in the pen, which signifies a mouse click to the computer. These whiteboards are widely used by professional digital artists and designers as they help in providing more accurate information.

- **Wii Remote, IR Pen-based Whiteboard**: A Wii Remote, infrared (IR) pen-based whiteboard works on sound electronic and optical principles. It uses Wii remote control with IR camera placed in front of a remote control to track dots of an infrared light which is created by an IR pen. These whiteboards have become much popular due to their portability and reduced cost.

Note: *The device driver software installed on the computer enables the whiteboard to act as a human input device (HID), just like a mouse, to help the user with easy interaction.*

4.3.8 Headphone and Headset

Headphone is an audio device equipped with a pair of speakers attached to a headstrap worn by the users. Headphones are used with almost all electronic devices such as portable computers, CD/DVD players, mp3 players, iPod, etc. They comprise high-quality speakers, are light in weight and very comfortable to wear. They can be easily connected to a computer via a mini stereo plug.

The headphones allow the users to listen only; however, some applications like live chats, videoconferencing and telecalling services demand the user to also speak at the same time. For such applications, a headset is used instead of a headphone. A headset is a combination of one or two speakers and a Figure 4.54 Headphone and Headset microphone with both the speaker and the microphone attached to a headstrap. The speaker(s) allows the person to hear the conversation and the microphone allows the user to communicate with the person on the other end (see Figure 4.54).

Figure 4.54 Headphone and Headset

4.4 COMPUTER TERMINALS

A computer terminal is a special unit that can perform both input and output. A terminal is an I/O device that uses a keyboard for input and a monitor for output. Due to this reason, a terminal is also known as video display terminal (VDT). Terminals can be categorized into the following types:

- **Dumb Terminal**: It refers to a terminal that has no processing or programming capabilities. It is designed to communicate exclusively with a host computer. Usually, it consists of a screen and keyboard used to access a host computer. It has electronics circuitry enough to interpret incoming instructions from the host computer, to display characters on the screen, to interpret keystrokes on the keyboard and pass them on to the host computer. Generally, dumb terminals are used for simple data entry or retrieval tasks. An example of a dumb terminal is one used by airline clerks at airport ticket and check-in counters.
- **Smart Terminal**: A smart terminal has built-in processing capability and memory but does not have its own storage capacity. In comparison to the dumb terminal, a smart terminal can communicate, retrieve data and can perform a limited processing of its own, that is, editing or verifcation of data. However, this kind of terminal cannot be used for programming. They are often found in local area networks in offices.
- **Intelligent Terminal**: An intelligent terminal has memory and inbuilt microprocessors, thus also known as user-programmable terminal. This terminal can independently perform a certain number of jobs without even interacting with the mainframe. Although they are intelligent terminals, some workstations are designed without disk drives. Due to this, restricted data cannot be downloaded or copied.

LET US SUMMARIZE

1. An *input device* is an electromechanical device that accepts data or information from the user and translates the information into

a form which the computer can interpret.

2. *Keyboards* are the most commonly used data entry devices. By pressing down the keys of the keyboard, data are entered into the computer.
3. *Pointing devices* are the input devices by which we can point out and select items rapidly from the multiple options displayed on the screen. These devices can also be used to create graphic elements on the screen such as lines, curves and freehand shapes. The common types of pointing devices available are the mouse, trackball, joysticks, touch screens and light pens.
4. *Mouse* is a small hand-held pointing device which can be used to input commands or information. It contains two or three buttons and by pressing one of the buttons, the mouse either marks a place on the screen or makes selections from the data on the screen. Common types of mouse action available are pointing, click, double-click, right-click and drag and drop.
5. A *trackball* is an upturned mouse, with a movable ball on the top of a stationary base. It is used to control the cursor movements and the actions on a computer screen. It allows the user to perform each of these tasks separately.
6. *Joystick* is a device that moves in all directions and controls the movement of the pointer. It activates the computer with different parts and is used to control the cursor on the screen.
7. Using a *light pen*, one can select the objects on the monitor by directly pointing to the objects. It is useful for identifying a specific location. Since the light pen is a passive device with a sensor only, it provides no information when held over a blank part of the screen.
8. A *touch screen* is a type of display screen device that is placed on the computer monitor in order to allow direct selection or activation of the computer when somebody touches the screen. It registers the input when a finger or other object touches the screen. It consists of three main components: a touch sensor, a controller and a software driver.
9. *Trackpad* (also referred to as touchpad) is a stationary pointing device that works by sensing the movement of fingers across a small sensitive surface (1.5 or 2 inches) and translating them into the pointer movement on the screen.
10. *Speech recognition* is the technology by which sounds, words or phrases spoken by humans are converted into digital signals, and these signals are transformed into coding patterns where the meaning has been assigned. It is more generally called as sound *recognition*.
11. *Digital camera* records images in an electronic form, that is, the image is represented in computer's language of bits and bytes. It is a long string of 1s and 0s that represent all the tiny coloured dots or pixels that collectively make up the image.
12. A *webcam* is a portable video camera, which captures live video or images that may be viewed in real time over the network or the Internet.
13. A *scanner* is a device that scans an image and transforms the image to ASCII codes. These images can be edited, manipulated, combined and then printed. The common types of scanners are: handheld and *flatbed*.
14. The *optical character recognition* (*OCR*) software translates the bitmap image of text to the ASCII codes that the computer can interpret as letters, numbers and special characters.
15. The *optical mark recognition* (*OMR*) is the process of detecting the presence of intended marked responses. A mark registers

significantly less light than the surrounding paper. Optical mark reading is done by a special device known as optical mark reader.

16. A *magnetic ink character reader* (*MICR*) is used to recognize the magnetic ink characters; it reads the characters by examining their shapes in a matrix form.
17. *Bar codes are machine-readable codes*, represented in parallel vertical lines with varying widths. By the combination of these vertical lines (bars) and the spacing between them, bar codes represent the alphanumeric data. For reading these bar-coded data, a device is used, which is known as a bar code reader.
18. An *output device* is an electromechanical device, which converts machine-readable data or information into human-readable form. The printed form of output is referred to as *hard copy* while the form of output displayed on the screen is referred to as soft copy.
19. Based on the hard copy and soft copy outputs, the output devices are classified into: *hard copy* and *soft copy output devices*. *Printers, plotters* and *microfilms* are the most commonly used *hard copy output devices while monitors, voice response systems, projectors, electronic whiteboards*, and *headphones* and *headsets* are some commonly used soft copy output devices.
20. A *printer* is a device that prints information from the computer on to paper. The two major categories of printer technologies are *impact printers* and *non-impact printers*.
21. An *impact printer* forms characters and graphics on a piece of paper by striking a mechanism against an ink ribbon that comes into physical contact with the paper. It includes *dot matrix printers, daisy wheel printers* and *drum printers*.
22. Printers that form characters and images without making direct contact between printing mechanism and paper are called *non-impact printers*. They are categorized as *ink-jet printers* and *laser printers*.
23. *Hybrid document reproduction apparatus* (*HYDRA*) printer, popularly known as all-in-one printer, is a device that consolidates the capabilities of multiple devices in one machine. It may include some or all of the devices like printer, scanner, photocopier and fax machine. Plotters are special-purpose drawing devices, which reproduce graphic images on paper using a pen whose movements are controlled by the computer. The lines drawn by these devices are continuous and very accurate. Plotters are classified as pen *plotters* and electrostatic plotters.
24. *Computer output microfilm* is an extremely high-speed, low-cost process, which records computer-generated information directly from the computer tape or cartridge to a miniaturized microfilm media. The microfilm product is in fiche or roll format, which can be duplicated rapidly and inexpensively.
25. *Computer monitor* is used to display the keyed data on the screen and to receive messages and processed information from the computer.
26. *Cathode ray tube* (*CRT*) is the most common type of monitor for the office and the home. In a CRT, an electron gun is used which fires electrons at groups of phosphor dots coating the inside of the screen. When the electrons strike the phosphor dots, they glow to give the colours.
27. In *liquid crystal display* (*LCD*) screen, the backlight passes through the first of two polarizing filters. The polarized light

then passes through a layer that contains thousands of liquid crystal blobs aligned in tiny containers known as cells. Electric leads around the edge of the LCD create an electric field that twists the crystal molecule, which lines the light up with the second polarizing filter and allows it to pass through.

28. Although CRT and LCD displays are the most commonly used, other kinds of displays also exist. Two of them are *plasma display and thin-film* electroluminescent display.
29. The standard computer system can talk with the addition of two components: a speech *synthesizer and a screen reading software*. A speech synthesizer is a software that converts text information into spoken sentences. Screen reading software enables the user to control the synthesizer. It allows a user to access, or view text that is present on the screen.
30. *Screen image projector* is an output device, which is used to project information from a computer on to a large screen, so that it can be simultaneously viewed by a large group of people. Projectors are mainly of two types: LCD projectors and DLP projectors.
31. *Electronic whiteboard* is an interactive presentation device that works electronically to display digital images, drawings, and text in different colours. The whiteboard is generally mounted on the wall and is connected to a computer (through USB, serial port or wireless technology like Bluetooth) and projector.
32. *Headphone* is an audio device equipped with a pair of speakers which are attached via a headstrap worn by the users. A headset is a combination of one or two speakers and a microphone with both the speakers and the microphone attached to a headstrap.
33. *Computer terminals*, also called as *display or video display terminals*, enable a computer to receive or deliver data through input and output devices, respectively.

EXERCISES

Fill in the Blanks

1. Mouse and light pen are examples of ..... input devices.
2. Two basic methods used for optical character recognition (OCR) are ..... and
3. ..... is a machine-readable code in the form of parallel vertical lines.
4. MICR stands for .....
5. The bar-coded data represent the ..... and a device, which is used for reading the bar-coded data, is known as .....
6. The most used layout of a keyboard is .....
7. The printed form of output is referred to as ..... and the form of output, which is shown on a display screen, is referred to as.....
8. An impact printer, which gives only alphanumeric output, is called.....
9. ..... plotters are computer graphics devices that use drum or flatbed paper holder.
10. A non-impact printer that forms an image by spraying ink from a matrix of tiny jets is called .....
11. The mouse may be classified as ..... or .....
12. ..... is the oldest printing technology.

13. The quality of the monitor is generally judged by considering four basic terms ....., ..., ..., and .....
14. The number of bits used to describe a pixel is known as .....
15. A computer can talk to persons through a device known as.....

Multiple-choice Questions

1. ..... is the most commonly used input device.
    1. Keyboard
    2. Mouse
    3. Joystick
    4. Scanner
2. Mouse, trackball and joystick are the examples of .....
    1. Scanning devices
    2. Storing devices
    3. Pointing devices
    4. Multimedia devices
3. Which of the following is not a computer terminal?
    1. Intelligent
    2. Smart
    3. Both (a) and (b)
    4. None of these
4. Which of the following technique is best suited for bank cheques?
    1. OCR
    2. OMR
    3. BAR
    4. MICR
5. Choose the most suitable device among the following for playing computer games.
    1. Hand-held scanner
    2. Joystick
    3. Optical Mark Reader
    4. Digital Camera
6. Light pen is frequently used in .....
    1. Computer Aided Design
    2. Word Processing
    3. Presentation
    4. Spreadsheet
7. Choose the odd one out.
    1. QWERTY
    2. SULTRY
    3. AZERTY
    4. DVORAK
8. Which of the following is not a pointing device?
    1. Scanner
    2. Trackball
    3. Light Pen
    4. Mouse
9. Which of the following is not an output device?
    1. Printer
    2. Keyboard
    3. Monitor

4. Plotter

10. With the combinations of ....., ... and ..... colours, other colours can be shown.
    1. Red, Black, Blue
    2. Yellow, Blue, Green
    3. Red, Blue, Green
    4. Black, White, Yellow
11. Which of the following is used for both input and output?
    1. Computer Terminals
    2. Pen Plotter
    3. Dot Matrix Printer
    4. None of these
12. The ..... is also called as letter quality printer.
    1. Dot Matrix Printer
    2. Ink Jet Printer
    3. Daisy Wheel Printer
    4. Laser Printer
13. The ..... terminal is referred to as non-programmable terminal whereas ..... terminal is referred to as user-programmable terminal.
    1. Dumb, Intelligent
    2. Dumb, Smart
    3. Smart, Intelligent
    4. None of these
14. An individual small dot, which one sees on the computer screen is called .....
    1. Font
    2. Character
    3. Screen Point
    4. Pixel
15. LCD stands for .....
    1. Liquid Colour Display
    2. Light Colour Display
    3. Lithium Crystal Display
    4. Liquid Crystal Display

State True or False

1. Any hardware item that is attached to the CPU is referred to as peripheral device.
2. Impact printers are slower as compared to non-impact printers.
3. The speed of a printer is measured by character per minute.
4. Audio response is a soft copy output.
5. Dot pitch can be changed by the user.
6. Plotters are best suited for desktop publishing.
7. Bank cheques are an example of bar code marking.
8. QWERTY is the most popular keyboard layout.
9. A speech recognition system converts the computer text and commands into human spoken words.
10. More pixels per square inch mean better resolution.
11. A digital camera's photographs can be stored in a computer's hard disk.
12. The most common aspect ratio used in computer monitors is 4:4.
13. Raster scan monitors draw a picture one line at a time and for this reason they are referred to as vector displays, stroke-writing, or calligraphic displays.

14. Ultrasonic acoustic waves are one of the most commonly used touch screen techniques.
15. 16-bits colour allows up to 256 colours.

Descriptive Questions

1. Why are input and output devices necessary for a computer system?
2. What are pointing devices? Discuss some of the commonly used pointing devices.
3. What is a touch screen device? Explain with their advantages and disadvantages and give some typical applications for which touch screen is most suitable as an input device.
4. What is a speech recognition system? Explain its types and give logical reasons. Why do currently available speech recognition devices have limited success?
5. What do you mean by optical scanning devices? Explain any four such devices.
6. Explain the different types of printers. Make a clear difference between them in terms of speed, cost and method of operation, and suggest suitable applications.
7. Describe the application areas where plotters are useful. Discuss the various types of plotters.
8. What is speech synthesizer? How does it function? Give some of its typical applications.
9. What is LCD? Differentiate between CRT and LCD monitors.
10. Write short notes on the following:
    1. Resolution
    2. Aspect ratio
    3. Dot pitch
    4. Refresh rate
    5. Colour depth

ANSWERS

Fill in the Blanks

1. Pointing
2. Matrix matching, Feature extraction
3. Bar code
4. Magnetic ink character recognition
5. Alphanumeric data, Bar code reader
6. QWERTY
7. Hard copy, Soft copy
8. Daisy wheel
9. Pen
10. Ink-jet
11. Mechanical mouse, Optical mouse
12. Dot Matrix
13. Size, Resolution, Refresh rate, Dot pitch
14. Resolution
15. Speech synthesizer

Multiple-choice Questions

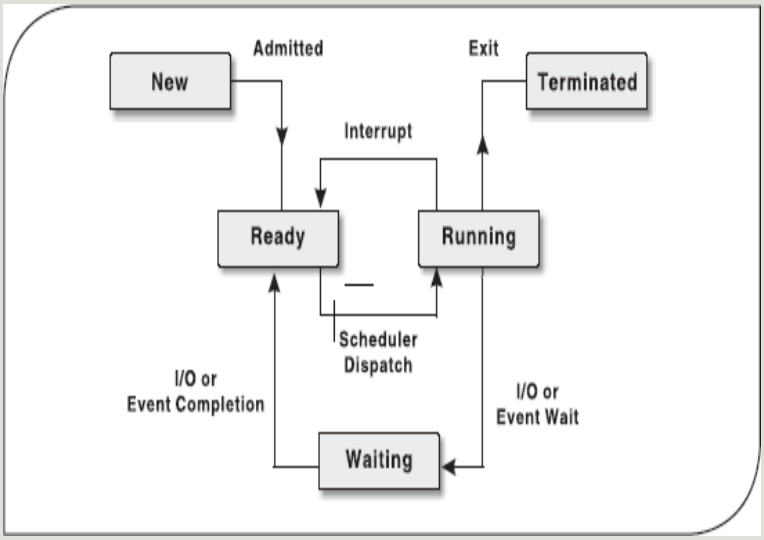
1. (a)
2. (c)
3. (c)
4. (d)
5. (b)
6. (a)
7. (b)
8. (a)
9. (b)
10. (c)
11. (a)
12. (c)
13. (a)
14. (d)
15. (d)

State True or False

1. True
2. True
3. False
4. True
5. False
6. False
7. False
8. True
9. False
10. True
11. True
12. False
13. False
14. True
15. False

Chapter 5

Operating System



This chapter provides an outlook on operating system. An operating system is a collection of system programs, which controls the operation of a computer system. The discussion begins with the evolution and types of operating system. It continues with the various functions of an operating system, namely, process management, memory management, file management, device management, security management, and user interface.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The need to develop an operating system
- What is an operating system and what are its objectives
- How operating systems evolved from single user character-based interfaces to modern day multi-user graphical user interface
- Six major types of operating systems
- Six major functions of an operating system

5.1 INTRODUCTION

In the early days of computer use, computers were huge machines, which were expensive to buy, run and maintain. The user at that time interacted directly with the hardware through machine language. A software was required which could perform basic tasks such as recognizing input from the keyboard, sending output to the display screen, keeping track of files and directories on the disk, and controlling peripheral devices such as printers and scanners. The search for such software led to the evolution of modern day *operating system* (*OS*). This software is loaded onto the top of memory and performs all the aforesaid basic tasks. Initially, the OS's interface was only character-based. This interface provides the user with a command prompt and the user has to type all the commands to perform various functions. As a result, the user had to memorize many commands. With the advancement in technology, OS became even more user friendly by providing graphical user interface (GUI). The GUI-based OS allows the user to interact with the system using visual objects such as windows, pull-down menus, mouse pointers and icons. Consequently, operating the computer became easy and intuitive.

5.2 OPERATING SYSTEM: DEFINITION

An *operating system* is a collection of system programs that together controls the operation of a computer system. The OS along with hardware, applications, other

system software and users constitute a computer system as shown in Figure 5.1. It is the most important part of any computer system. It acts as an intermediary between a user and the computer hardware. The OS has two objectives.

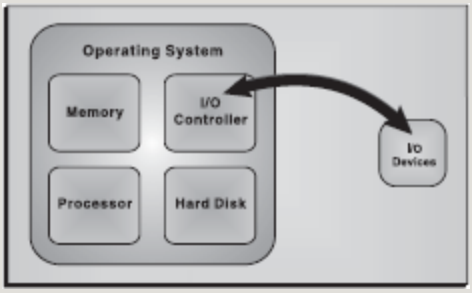


Figure 5.1 Managing Hardware

- **Managing Hardware:** The prime objective of the OS is to manage and control various hardware resources of a computer system. These hardware resources include processor, memory, disk space, I/O devices and so on. The OS supervises which input device's data is requesting for being processed and which processed data is ready to be displayed on the output device. In addition to communicating with hardware, the OS provides an error handling procedure and displays an error notification. If a device is not functioning properly, the OS tries to communicate with the device again. If it is still unable to communicate with the device, it provides an error message notifying the user about the problem. Figure 5.2 illustrates how an OS manages the hardware resources of a computer system.

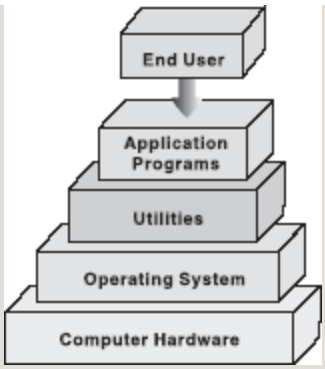


Figure 5.2 Computer System Components

- **Providing an Interface:** The OS organizes applications so that users can easily access, Figure 5.2 Computer System use and store them. When an application is Components opened, the OS assists the application to provide the major part of the user interface. It provides a stable and consistent way for applications to deal with the hardware without the user having to know all the details of the hardware. If the application program is not functioning properly, the OS again takes control, stops the application and displays an appropriate error message.

5.3 EVOLUTION OF OPERATING SYSTEM

In the early days, the computers lacked any form of OS. The user would arrive at the machine armed with his program and data, often on punched paper tape. The program would be loaded into the machine and the machine set to work. Then came machines with libraries of support code (initial OSs), which were linked to the user's program to assist in operations such as input and output. At this stage, OSs were very diverse with each vendor producing one or more OSs specific to its particular hardware. Typically, whenever new hardware architecture was introduced, there was a need of new OS compatible with the new architecture. This state of affairs continued until 1960s when IBM developed the S/360 series of machines. Although there were enormous performance differences across the range, all the machines ran essentially the same OS called *S/360*.

Then came the small 4 bit and 8 bit processors known as *microprocessors*. The development of microprocessors provided inexpensive computing for the small businesses. This led to the widespread use of interchangeable hardware components using a common interconnection, and thus creating an increasing need for standardized OS to control them. The most important among the early OSs was CP/M-80 for the 8080/8085/Z-80 microprocessors. With the development of microprocessors like 386, 486 and the Pentium series by Intel, the whole computing world got a new dimension. AT&T and Microsoft came up with character-based OSs, namely, *Unix* and *Disk OS,* respectively, which supported the prevalent hardware architectures. After the character-based OSs, Microsoft and Apple Macintosh came with their Windows 3.1 and MAC, which were GUI-based OSs and well suited for the Desktop PC market. Today, OSs such as Windows XP and Red Hat Linux have taken the driver's seat in personal desktops. These OSs with their remarkable GUI and network support features can handle diverse hardware devices.

5.4 TYPES OF OPERATING SYSTEMS

The OS has evolved immensely from its primitive days to the present digital era. From batch processing systems to the latest embedded systems, the different types of OSs can be classified into six broad categories.

- **Batch Processing OS:** This type of OS was one of the first to evolve. Batch processing OS allowed only one program to run at a time. These kinds of OSs can still be found on some mainframe computers running batches of jobs. Batch processing OS works on a series of programs that are held in a queue. The OS is responsible for scheduling the jobs according to priority and the resources required. Batch processing OSs are good at churning through large numbers of repetitive jobs on large computers. For example, this OS would be best suited for a company wishing to automate their payrolls. A list of employees will be entered, their monthly salaries will be calculated, and corresponding pay slips would be printed. Batch processing is useful for this purpose since these procedures are to be repeated for every employee and each month.

- **Multi-user or Time-sharing OS:** This system is used in computer networks which allow different users to access the same data and application programs on the same network. The multi-user OS builds a user database account, which defines the rights that users can have on a particular resource of the system.
- **Multi-programming OS:** In this system, more than one process (task) can be executed concurrently. The processor is switched rapidly between the processes. Hence, a user can have more than one process running at a time. For example, a user on his computer can have a word processor and an audio CD player running at the same time. The multi-tasking OS allows the user to switch between the running applications and even transfer data between them. For example, a user can copy a picture from an Internet opened in the browser application and paste it into an image editing application.
- **Real-time OS (RTOS):** This system is designed to respond to an event within a predetermined time. This kind of OS is primarily used in process control, telecommunications and so on. The OS monitors various inputs which affect the execution of processes, changing the computer's model of the environment, thus affecting the output, within a guaranteed time period (usually less than one second). As the real-time OSs respond quickly, they are often used in applications such as flight reservation systems, railway reservation systems, military applications, etc.
- **Multi-processor OS:** This system can incorporate more than one processor dedicated to running processes. This technique of using more than one processor is often called *parallel processing*. The main advantage of multi-processor systems is that they increase the system throughput by getting more work done in less time.
- **Embedded OS:** It refers to the OS that is self-contained in the device and resident in the ROM. Since embedded systems are usually not general-purpose systems, they are lighter or less resource intensive as compared to general-purpose OSs. Most of the embedded OSs also offer real-time OS qualities. Typical systems that use embedded OSs are household appliances, car management systems, traffic control systems and energy management systems.

5.5 FUNCTIONS OF OPERATING SYSTEMS

The main functions of a modern OS are as follows:

- **Process Management:** As a process manager, the OS handles the creation and deletion of processes, suspension and resumption of processes, and scheduling and synchronization of processes.
- **Memory Management:** As a memory manager, the OS handles the allocation and de-allocation of memory space as required by various programs.
- **File Management:** The OS is responsible for creation and deletion of files and directories. It also takes care of other file-related activities such as organizing, storing, retrieving, naming and protecting the files.
- **Device Management:** The OS provides input/output subsystem between process and device driver. It handles the device caches, buffers and interrupts. It also detects the device failures and notifies the same to the user.

- **Security Management:** The OS protects system resources and information against destruction and unauthorized use.
- **User Interface:** The OS provides the interface between the user and the hardware. The user interface is the layer that actually interacts with the computer operator. The interface consists of a set of commands or menus through which a user communicates with a program.

THINGS TO REMEMBER

States of a Process

A process changes various states during its lifetime where each state indicates the current activity of the process. These states are as follows:

- **New:** The process is being created.
- **Ready:** The process is ready to be assigned to the processor.
- **Running:** The process is being executed.
- **Waiting:** The process is waiting for signal from some other process.
- **Terminated:** The process has finished its execution.

5.5.1 Process Management

A *process* is an execution of a sequence of instructions or program by the CPU. It can also be referred to as the basic unit of a program that the OS deals with, with the help of the processor. For example, a text editor program running on a computer is a process. This program may cause several other processes to begin—like it can furnish a request for printing while editing the document. Thus, we can say that the text editor is a program that initiates two processes—one for editing the text and second for printing the document.

Hence, a process is initiated by the program to perform an action, which can be controlled by the user or the OS. A process in order to accomplish a task needs certain

resources like CPU time, memory allocation and I/O devices. Therefore, the idea of process management in an OS is to accomplish the process assigned by the system or the user in such a way that the resources are utilized in a proper and efficient manner.

Life Cycle of a Process: The OS is responsible for managing all the processes that are running on a computer and allocating each process a certain amount of time to use the processor. In addition, the OS also allocates various other resources that processes need during execution, such as computer memory or disk space. To keep track of all the processes, the OS maintains a table known as the *process* table. This table stores many pieces of information associated with a specific process, that is, program counter, allocated resources, process state, CPU-scheduling information, and so on.

Initially, a process is in the new state. When it becomes ready for execution and needs the CPU, it switches to the ready state. Once the CPU is allocated to the process, it switches to the running state. From the running state, the process goes back to the ready state if an interrupt occurs or to the waiting state if the process needs some I/O operation. In case the process has switched to ready state, it again comes to running state after the interrupt has been handled. On the other hand, if the process has switched to waiting state, then after the completion of I/O, it switches to ready state and then to running state. Thus, a process continues to switch among the ready, running and waiting states during its execution. Finally, it switches to terminated state after completing its execution as shown in Figure 5.3.

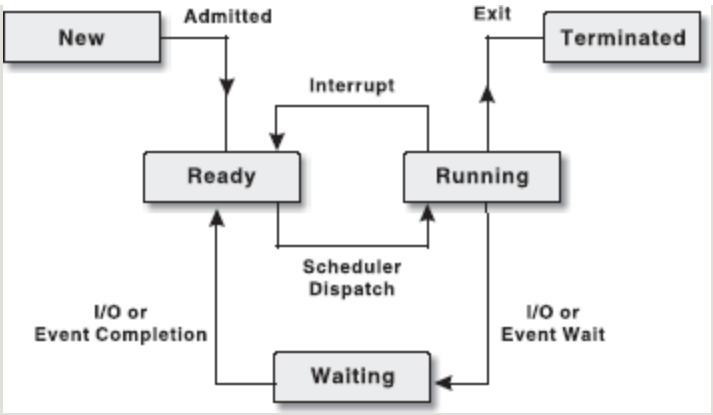


Figure 5.3 Life Cycle of a Process

Note: *The change of the state of the process from one form to another is called context change and this course of action is known as context switching*.

Let us consider the steps in an example of two processes, a text editor and a calculator, running simultaneously on a computer system.

Step 1	The OS receives a request to open a text editor.
Step 2	A new process for the text editor is initiated by the OS.
Step 3	Resources such as keyboard, memory and hard disk are made available and the process enters in to ready state.
Step 4	The scheduler then dispatches the text editor to running state.
Step 5	In the meantime, another process, a calculator, is initiated by the user.
Step 6	A new process is created by the OS for the calculator.
Step 7	Resources such as keyboard and memory are made available and this

	process enters in to ready state.
Step 8	Now as the calculator process is ready for the processor to work upon, the text editor process, if not terminated (finished), is kept in either waiting state or in the ready state.
Step 9	The calculator process then attains the running state by the scheduler and when finished, is terminated.
Step 10	Then the text editor process again enters in the running state from the ready state.

FACT FILE

Multithreading

Writing a program where a process creates multiple threads is called **multithread programming.** It is the ability by which an OS is able to run different parts of the same program simultaneously. It offers better utilization of processor and other system resources. For example, word processor makes use of multithreading—it can check spelling in the foreground as well as save a document in the background.

Threads: A *thread* is a task that runs concurrently with other tasks within the same process. Also known as *lightweight process,* a thread is the simplest unit of a process. The single thread of control allows the process to perform only one task at one time. An example of a single thread in a process is a text editor where a user can either edit the text or perform any other task like printing the document. In a multitasking OS, a process may contain several threads, all running at the same time

inside the same process. It means that one thread of a process can be editing the text while another is printing the document. Generally, when a thread finishes performing a task, it is suspended or destroyed.

Uniprogramming and Multiprogramming: As the name implies, uniprogramming means only one program at a time. In uniprogramming, users can perform only one activity at a time. In multiprogrammed systems, multiple programs can reside in the main memory at the same time. These programs can be executed concurrently, thereby requiring the system resources to be shared among them. In multiprogrammed systems, an OS must ensure that all processes get a fair share of CPU time.

Process Scheduling: In a multiprogrammed system, at any given time, several processes will be competing for the CPU's time. Thus, a choice has to be made as to which process to allocate the CPU next. This procedure of determining the next process to be executed on the CPU is called *process scheduling* and the module of OS that makes this decision is called scheduler. The prime objective of scheduling is to switch the CPU among processes so frequently that users can interact with each program while it is running.

For scheduling purposes, there exist different queues in the system: *job queue, ready queue* and *device queue*. As the processes enter the system for execution, they are kept in the job queue (or *input queue*) on a mass storage device such as hard disk. From the job queue, the processes which are ready for execution are brought into the main memory. In the main memory these processes are kept in the ready queue. In other words, the ready queue contains all those processes that are waiting for the CPU. For each I/O device attached to the system, a separate device queue is maintained. The process that

needs to perform I/O during its execution is kept into the queue of that specific I/O device and waits there until it is served by the device.

Depending on the level of scheduling decisions to be made, the following types of schedulers may coexist in a complex OS:

- **Long-term Scheduler:** Also known as *job scheduler* or *admission scheduler*, it works with the job queue. It selects the next process to be executed from the job queue and loads it into the main memory for execution. This scheduler is generally invoked only when a process exits from the system. Thus, the frequency of invocation of long-term scheduler depends on the system and workload and is much lower than other two types of schedulers.
- **Short-term Scheduler:** Also known as *CPU* scheduler, it selects a process from the ready queue and allocates the CPU to it. This scheduler is required to be invoked frequently as compared to long-term scheduler. This is because a process generally executes for a short period and then it may have to wait either for I/O or some other reason. At that time, the CPU scheduler must select some other process and allocate the CPU to it. Thus, the CPU scheduler must be fast in order to provide the least time gap between executions.
- **Medium-term Scheduler:** Also known as *swapper,* it comes into play whenever a process is to be removed from the ready queue (or from the CPU in case it is being executed) thereby reducing the degree of multiprogramming. This process is stored at some space on the hard disk and later brought into the memory to restart execution from the point where it left off. This task of temporarily switching a process in and out of main memory is known as *swapping*. The medium-term scheduler selects a process among the partially executed or unexecuted swapped-out processes and swaps it in the main memory. The medium-term scheduler is usually invoked when some space becomes free in memory by the termination of a process or if the supply of ready process reduces below a specified limit.

The various types of schedulers are illustrated in Figure 5.4.

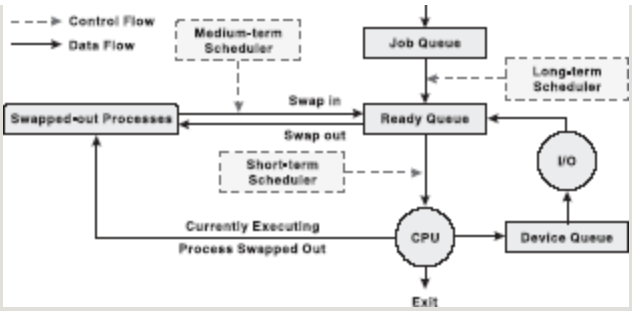


Figure 5.4 Types of Schedulers

Preemptive and non-preemptive scheduling

CPU scheduling may take place under the following four circumstances:

1. When a process switches from the running state to the waiting state.
2. When a process switches from the running state to the ready state.
3. When a process switches from the waiting state to the ready state.
4. When a process terminates.

When scheduling takes place under the “first” and “fourth” circumstances, it is said to be *non-preemptive* scheduling and if the scheduling takes place under the “second” and “third” circumstances, it is said to be *preemptive*. In the preemptive scheme, the scheduler can forcibly remove the processor from the currently running process before its completion in order to allow some other process to run. In the non-preemptive scheme, once the processor is allocated to a process, it cannot be taken back until the process voluntarily releases it (in case the process has to wait for I/O or some other event) or until the process terminates. Thus, the main difference between the two schemes is that in the preemptive scheme, the OS has the control over the process current states whereas in the case of the non-preemptive scheme, the process once entered in to the running state gets the full control of the processor.

The scheduler uses some scheduling procedure to carry out the selection of a process for execution. Two popular scheduling procedures implemented by different OSs are *first-come-first-served* and *round robin* scheduling.

- **First-come-first-served (FCFS):** As the name suggests, in FCFS, the processes are executed in the order of their arrival in the ready queue, which means the process that enters the ready queue first, gets the CPU first. To implement the FCFS scheduling procedure, the ready queue is managed as a FIFO (first-in first-out) queue. Each time the process at the start of queue is dispatched to the processor, all other processes move up one slot in the queue as illustrated in Figure 5.5. When new processes arrive, they are put at the end of the queue. FCFS falls under non-preemptive scheduling and its main drawback is that a process may take a very long time to complete, and thus holds up other waiting processes in the queue.

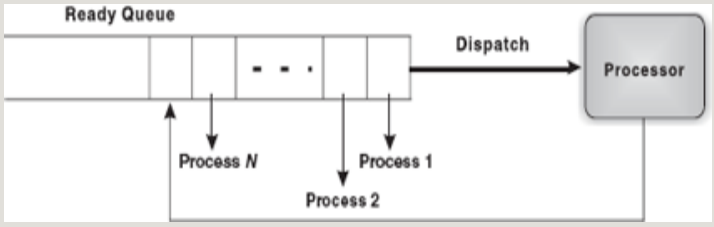


Figure 5.5 First-Come-First-Served Procedure

- **Round Robin:** Round robin scheduling was designed keeping in mind the limitations of the FCFS scheduling procedure. This procedure falls under preemptive scheduling, in which a process is selected for execution from the ready queue in FIFO sequence. However, the process is executed only for a fixed period known as *time slicing* or *quantum period* after which it is interrupted and returned to the end of the ready queue (see Figure 5.6). In the round robin procedure, processes are allocated CPU time on a turn basis.

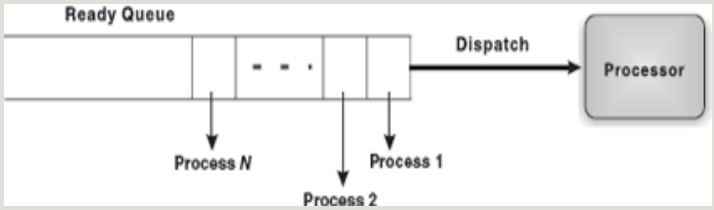


Figure 5.6 Round Robin Procedure

Nowadays, it is common in OSs for processes to be treated according to priority. This may involve a number of different queues and scheduling mechanisms using priority based on previous process activity. For example, the time required by the process for execution or how long it has been, since it was last executed by the processor.

Deadlock: In a multiprogramming environment, several processes may compete for a limited number of resources. A process requests for the required resource and if it is not available then the process enters the waiting state and remains in that state until it acquires the resource. There might be a situation when the process has to wait endlessly because the requested resource may be held by another waiting process. This type of situation is known as *deadlock*. To illustrate the deadlock situation, consider a system with two resources (say, printer and disk drive) and two processes P1 Held By and P2 running simultaneously. During execution, P1 requests for printer and P2 for disk drive. As the requested resources are available, the requests of both P1 and P2 are granted and the desired resources are allocated to them. Further, P1 requests for disk drive held by P2 and P2 requests for printer held by P1. Here, both processes will enter the waiting state. Since each process is waiting for the release of resource held by the other, they will remain in the waiting state forever, thus producing a deadlock (see Figure 5.7).

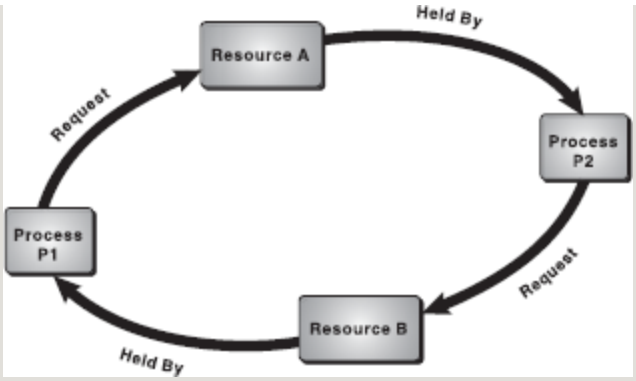


Figure 5.7 Deadlock

A deadlock situation arises if the following four conditions hold simultaneously on the system:

- **Mutual Exclusion:** Only one process can use a resource at a time. If another process requests for the resource, the requesting process has to wait until the requested resource is released.
- **Hold and Wait:** In this situation, a process might be holding some resource while waiting for additional resource, which is currently being held by another process.
- **No Preemption:** Resources cannot be preemptive, that is, resources cannot be forcibly removed from a process. A resource can only be released voluntarily by the holding process after that process has completed its task.
- **Circular Wait:** This situation may arise when a set of processes waiting for allocation of resources held by other processes forms a circular chain in which each process is waiting for the resource held by its successor process in the chain.

To ensure that deadlocks never occur, the system can use either a deadlock-prevention or a deadlock-avoidance scheme.

- **Deadlock Prevention:** Deadlock can occur only when all the four deadlock-causing conditions hold true. Hence, the system should ensure that at least one of the four deadlock-causing conditions would not hold true so that deadlock can be prevented.
- **Deadlock Avoidance:** Additional information concerning which resources a process will require and use during its lifetime should be provided to the OS beforehand. For example, in a system with one CD drive and a printer, process P might request first for the CD drive and later for the printer, before releasing both resources. On the other

hand, process Q might request first for the printer and the CD drive later. With this knowledge in advance, the OS will never allow allocation of a resource to a process if it leads to a deadlock, thereby avoiding the deadlock.

5.5.2 Memeory Management

In addition to managing processes, the OS also manages the primary memory of the computer. The part of the OS that handles this job is called *memory manager*. Since every process must have some amount of primary memory to execute, the performance of the memory manager is crucial to the performance of the entire system. As the memory is central to the operation of any modern OS, its proper use can make a huge difference. The memory manager is responsible for allocating the main memory to processes and for assisting the programmer in loading and storing the contents of the main memory. Managing the main memory, sharing, and minimizing memory access time are the basic goals of the memory manager. The major tasks accomplished by the memory manager so that all the processes function in harmony, are as follows:

- **Relocation:** Each process must have enough memory to execute.
- **Protection and Sharing:** A process should not run into another process's memory space.

Relocation: When a process is to be executed, it has to be loaded from the secondary storage (like hard disk) to the main memory (RAM). This is called *process loading*. Since, the main User Space memory is limited and other processes also need it for their execution, an OS swaps the two processes, which is called *swapping* (see Figure 5.8). Once the process is “swapped out”, it is uncertain to say when it will be “swapped in” because of the number of processes running concurrently.

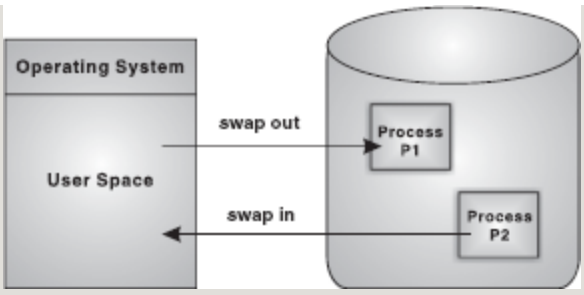


Figure 5.8 Process Swapping

Normally, when the process is swapped back into the main memory, it will be placed back to the same memory space that it occupied previously. However, in certain cases, it is not possible to place the process at the same memory location. This is not of much importance if the process is not address-sensitive. However, if the process requires some of its data or instruction to occupy the memory with a specific address, the process needs to be relocated. It is the responsibility of the memory manager to modify the addresses used in address-sensitive instructions (that use memory addresses) of the process so that it can execute correctly from the assigned area of memory.

Protection and Sharing: In multiprogrammed systems, as a number of processes may reside in the main memory at the same time, there is a possibility that a user program, during execution, may access the memory location allocated either to other user processes or to the OS. It is the responsibility of the memory manager to protect the OS from being accessed by other processes and the processes by one another. At the same time, the memory protection program should be flexible enough to allow concurrent processes to share the same proportion of the main memory. For example, consider a program that initiates different processes. If the memory

manager allocates the same portion of memory to all the processes instead of different memory allocations to different processes, a lot of memory is saved. Therefore, the memory protection routine of the OS should allow controlled sharing of the memory among different processes without letting them breach the protection criteria. If a process attempts to modify the contents of memory locations that do not belong to it, the memory protection routine intervenes and usually terminates the program.

Memory Allocation: In uniprogramming systems, where only one process runs at a time, memory management is very simple. The process to be executed is loaded into the part of memory space that is unused. Early MS-DOS systems support uniprogramming. The main challenge of efficiently managing memory comes when a system has multiple processes running at the same time. In such a case, the memory manager can allocate a portion of primary memory to each process for its own use. However, the memory manager must keep track of the running processes along with the memory locations occupied by them, and must also determine how to allocate and de-allocate available memory when new processes are created and old processes have finished their execution, respectively.

While different strategies are used to allocate space to processes competing for memory, three of the most popular are as follows:

- **Best Fit:** In this case, the memory manager places a process in the smallest block of unallocated memory in which it will fit. For example, a process requests 12 KB of memory and the memory manager currently has a list of unallocated blocks of 6 KB, 14 KB, 19 KB, 11 KB and 13 KB blocks. The best fit strategy will allocate 12 KB of the 13 KB block to the process.
- **First Fit:** The memory manager places the process in the first unallocated block that is large enough to accommodate the process. Using the same example to fulfil the 12 KB request, first fit will allocate 12 KB of the 14 KB block to the process.

- **Worst Fit:** The memory manager places a process in the largest block of unallocated memory available. To furnish the 12 KB request again, worst fit will allocate 12 KB of the 19 KB block to the process, leaving a 7 KB block for future use.

Figure 5.9 illustrates that in best fit and first fit strategies, the allocation of memory results in the creation of a tiny fragment of unallocated memory. Since the amount of memory left is small, no new processes can be loaded here. This job of splitting primary memory into segments as the memory is allocated and de-allocated to the processes is known as *fragmentation*. The worst fit strategy attempts to reduce the problem of fragmentation by allocating the largest fragments to the new processes. Thus, a larger amount of space in the form of tiny fragments is left unused. To overcome this problem, the concept of *paging* was introduced.

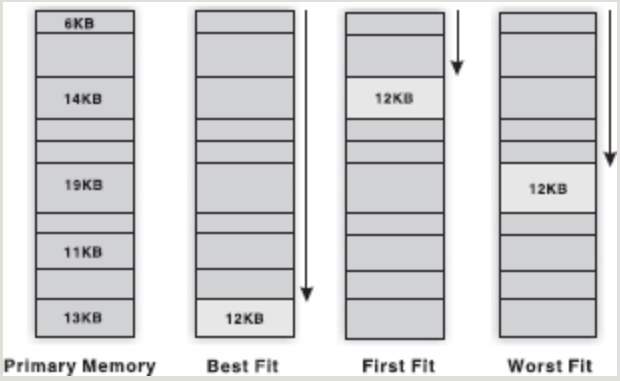


Figure 5.9 Strategies for Memory Allocation

THINGS TO REMEMBER

Physical and Logical Address

Every byte in memory has a specific address known as *physical address*. Whenever a program is brought into the main memory for execution, it occupies certain memory locations. The set of all physical addresses used by the program is known as its *physical address space*. However, before execution, a program is compiled to run starting from some fixed address and accordingly all the variables and procedures used in the program are assigned some specific address known as *logical address*. The set of all logical addresses used by the program is known as its *logical address space*.

Paging Paging is a memory management scheme that allows the processes to be stored non-contiguously in the memory. The memory is divided into fixed size chunks called *page frames*. The OS breaks the program's address space (the collection of addresses used by the program) into fixed-size chunks called *pages,* which are of the same size as that of the page frames. Generally, a page size is of 4 KB. However, some systems support even larger page sizes such as 8 KB, 4 MB, etc. When a process is to be executed, its pages are loaded into unallocated page frames (not necessarily contiguous).

Each address generated by the CPU (that is, logical address) is divided into two parts: *page number* (high-order bits) and a *page offset* or *displacement* (low-order bits). Since the logical address is a power of 2, the page size is always chosen as a power of 2 so that the logical address can be converted easily into page number and page offset. To map the logical addresses to physical

addresses in memory, a mapping table called *page table* is used. The OS maintains a page table for each process to keep track of which page frame is allocated to which page. It stores the frame number allocated to each page and the page number is used as the index to the page table. Figure 5.10 shows the logical memory, page table and physical memory.

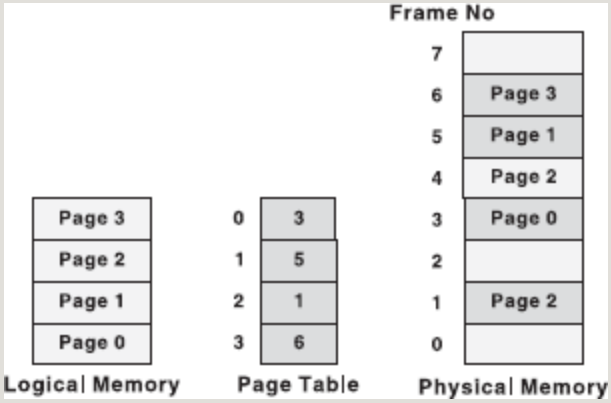


Figure 5.10 Logical Memory, Page Table and Physical Memory

Now let us see how address translation is performed in paging. To map a given logical address to the corresponding physical address, the system first extracts the page number and the offset. The system, in addition, also checks whether the page reference is valid (that is, it exists within the logical address space of the process). If the page reference is valid, the system uses the page number to find the corresponding page frame number in the page table. That page frame number is attached to the high-order end of the page offset to form the physical address in memory. The mechanism of translation of logical address into physical address is shown in Figure 5.11.

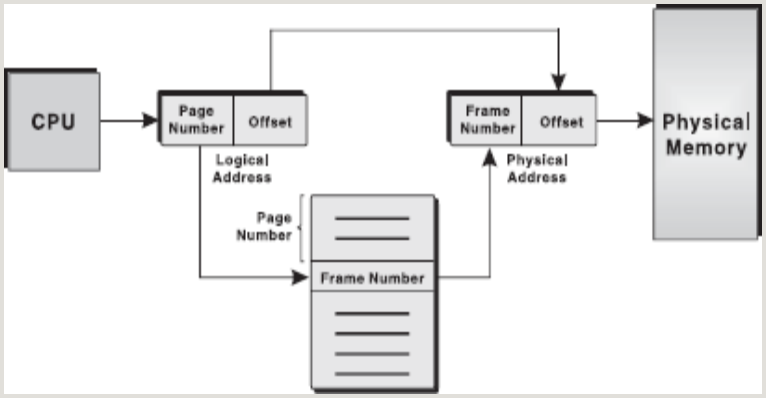


Figure 5.11 Address Translation in Paging

The main advantage of paging is that it minimizes the problem of fragmentation since the memory allocated is always in fixed units and any free frame can be allocated to a process.

Concept of Virtual Memory: A process executes only in the main memory, which is limited in size. Today, with the advent of graphic-oriented applications like 3D video games, business applications and so on, a user requires a larger memory than the main memory for running such applications. Note that it is not essential that the whole program must be loaded in the main memory for processing as only the currently referenced page needs to be present in the memory at the time of execution. Therefore, the rest of the pages of the program can reside in a hard disk portion allocated as *virtual memory* and can be loaded into the main memory whenever needed. This process of swapping the pages from the virtual memory to the main memory is called *page-in* or *swap-in*. With virtual memory, the system can run programs that are actually larger than the

primary memory of the system (see Figure 5.12). Virtual memory allows for very effective multiprogramming and relieves the user from the unnecessarily tight constraints of the main memory.

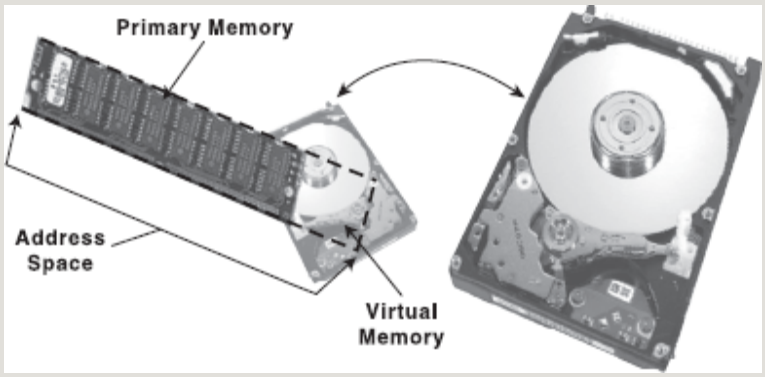


Figure 5.12 Virtual Memory

Virtual memory, in other words, is a way of showing the main memory of a computer system to appear effectively larger than it really is. The system does this by determining which parts of its memory are often sitting idle, and then makes a command decision to empty their contents onto a disk, thereby freeing up precious RAM.

Note: *In virtual memory systems, the logical address is referred to as virtual address and logical address space is referred to as virtual address space.*

Page Faults: In virtual memory systems, the page table of each process stores an additional bit to differentiate the pages in the main memory from that on the hard disk. This additional bit is set to 1 in case the page is in the main memory; otherwise 0. Whenever a page reference is made, the OS checks the page table to determine whether the page is in the main memory. If

the referenced page is not found in the main memory, a *page fault* occurs and the control is passed to the page fault routine in the OS. To handle page faults, the page fault routine checks first of all whether the virtual address for the desired page is valid (that is, it exists within the virtual address space of the process). If it is invalid, it terminates the process giving an error. Otherwise, it locates for a free page frame in memory and allocates it to the process, swaps the desired page into this allocated page frame, and updates the page table to indicate that the page is in memory.

While handling a page fault, there is a possibility that the memory is full and no free frame is available for allocation. In that case, the OS has to evict a page from the memory to make space for the desired page to be swapped in. To decide which page frame is to be replaced with the new page, the OS must track the usage information for all pages. In this way, the OS can determine which pages are being actively used and which are not (and therefore, can be removed from the main memory). Often the "least currently accessed" page (the page that has gone the longest time without being referenced) is selected.

5.5.3 File Management

File system is one of the most visible aspects of the OS. It provides a uniform logical view of the information storage, organized in terms of files, which are mapped onto the underlying physical device like the hard disk. While the memory manager is responsible for the maintenance of the primary memory, the *file manager* is responsible for the maintenance of the file system. In the simplest arrangement, the file system contains a hierarchical structure of data. This file system maintains user data and metadata (the data describing the files of the user data). The hierarchical structure usually contains the metadata in the form of directories of files

and sub-directories. Each file is a named collection of data stored on the disk. The file manager implements this abstraction and provides directories for organizing files. It also provides a spectrum of commands to read/write the contents of a file, to set the read/write position, to use the protection mechanism, to change the ownership, to list files in a directory, and to remove a file. The file manager provides a protection mechanism to allow users to administer how processes executing on behalf of different users can access the information contained in different files.

The file manager also provides a logical way for users to organize files in the secondary storage. To assist users, most file managers allow files to be grouped into a bundle called a *directory* or a *folder*. This allows a user to organize his or her files according to their purpose by placing related files in the same directory. By allowing directories to contain other directories, called *sub-directories*, a hierarchical organization can be constructed. For example, a user may create a directory called *games* that contains sub-directories called *cricket, football, golf, rugby* and *tennis* (see Figure 5.13). Within each of these sub-directories are files that fall within that particular category. A sequence of directories within directories is called a *directory path*.

Figure 5.13 File System

5.5.4 Device Management

Device management in an OS refers to the process of managing various devices connected to the computer. The device manager manages the hardware resources and provides an interface to hardware for application programs. A device communicates with the computer system by sending signals over a cable. The device communicates with the machine through a connection point called *port*. The communication using a port is done through rigidly defined protocols, like when to send the data and when to stop. These ports are consecutively connected to a bus (a set of wires) which one or more device uses to communicate with the system. The OS communicates with the hardware with the help of standard software provided by the hardware vendor called *device drivers*. A device driver works as a translator between the electrical signals from the hardware and the application programs of the OS. Drivers take data that the OS has defined as a file and translate them into streams of bits placed in specific locations on storage devices. There are differences in the way that the driver program functions, but most of them run when the device is required, and function much the same as any other process. The OS will frequently assign processes based on priority to drivers so that the hardware resources can be released and set free for further use.

Broadly, managing input and output is a matter of managing queues and buffers. A *buffer* is a temporary storage area that takes a stream of bits from a device like keyboard to a serial communication port. Buffers hold the bits and then release them to the CPU at a convenient rate so that the CPU can act on it. This task is important when a number of processes are running and taking up the processor's time. The OS instructs a buffer to continue taking the input from the device. In addition, it also instructs the buffer to stop sending data back to the

CPU if the process, using the input, is suspended. When the process, requiring input, is made active once again, the OS will command the buffer to send data again. This process allows a keyboard to deal with external users at a higher speed.

Spooling: SPOOL is an acronym for simultaneous peripheral operation on-line. *Spooling* refers to storing jobs in a buffer so that CPU can be efficiently utilized. Spooling is useful because devices access data at different rates. The buffer provides a waiting station where data can rest while the slower device catches up. The most common spooling application is print spooling. In print spooling, documents are loaded into a buffer, and then the printer pulls them off from the buffer at its own rate. Meanwhile, a user can perform other operations on the computer while the printing takes place in the background. Spooling also lets a user place a number of print jobs on a queue instead of waiting for each one to finish before specifying the next one. The OS manages all requests to read or write data from the hard disk through spooling (see Figure 5.14).

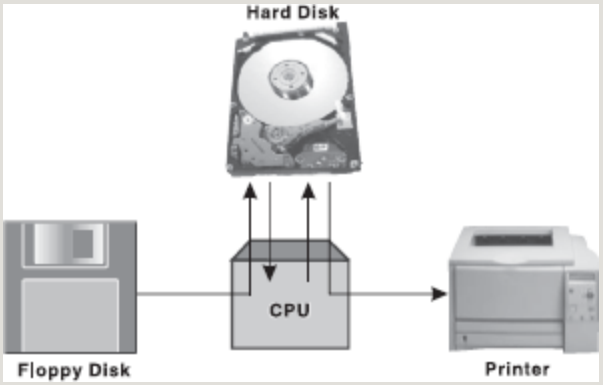


Figure 5.14 Spooling

5.5.5 Security Management

Security in terms of a computer system covers every aspect of its protection in case of a catastrophic event, corruption of data, loss of confidentiality and so on. Security requires not only ample protection within the system, but also from the external environment in which the system operates. In this section, we will be covering security in terms of internal protection, which is one of the most important functions of the OS. This involves protecting information residing in the system from unauthorized access. Various security techniques employed by the OS to secure the information are *user authentication* and *backup of data*.

User Authentication: The process of authenticating users can be based on a user's possession like a key or card, user information like username and password or user attributes like fingerprints and signature. Apart from these techniques, user information is often the first and most significant line of defence in a multi-user system. After the user identifies himself by a username, he is prompted for a password. If the password supplied by the user matches the password stored in the system, the system authenticates the user and gives him access to the system. A password can also be associated with other resources (files, directories and so on), which when requested, prompts the user for password.
Unfortunately, passwords can often be guessed, illegally transferred, or exposed. To avoid such situations, a user should keep the following points in mind:

- Password should be at least six characters in length.
- The system should keep track of any event about any attempt to break the password.
- The system should allow limited number of attempts for submitting a password on a particular system.
- Password based on dictionary words should be discouraged by the system. Alphanumeric passwords, such as *PASS011,* should be used.

Backup of Data: No matter what kind of information a system contains, backup of data is of utmost importance for its users. Backup or archiving is an important issue for a user and especially for business organizations. Typically, a computer system uses hard drives for online data storage. These drives may sometimes fail, or can be damaged in case of a catastrophic event, so care must be taken to ensure that the data is not lost. To ensure this, the OS should provide a feature of backing up of data, say from a disk to another storage device such as a floppy disk or an optical disk. The purpose of keeping backups is to restore individual files or complete file system in case of data loss. Recovery from the loss of an individual file, or of an entire disk, may be done from backup. OSs usually provide some system software that is used for taking backups of the data.

5.5.6 User Interface

OSs organize applications so that users can easily access them, use them and store application data. When an application is opened, the OS lets the application provide the majority of the user interface. The OS still has the responsibility of providing access to the hardware for whatever the application needs. If the program cannot function properly, the OS again takes control, stops the application, and displays an error message. An effective interface of an OS does not concern the user with the internal workings of the system. A good user interface should attempt to anticipate the user's requirements and assist him to gather information and use necessary tools. Common interfaces provided by different OSs can be categorized as *command line interface* (*CLI*) and *graphical user interface* (*GUI*).

Command Line Interface: In early days of computing, OSs provided the user with the facility of entering commands via an interactive terminal. Those were the only means of communication between a

program and its user, based solely on textual input and output. Commands were used to initiate programs, applications and so on. A user had to learn many commands for proper operation of the system (see Figure 5.15).

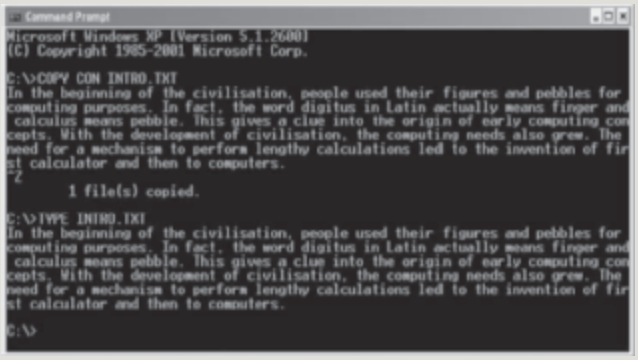


Figure 5.15 Command Line Interface

Graphical User Interface: With the development in chip designing technology, computer hardware became quicker and cheaper, which led to the birth of GUI-based OS. These OSs provide users with pictures rather than just characters to interact with the machine. The OS displays icons, buttons, dialog boxes, etc., on the screen (see Figure 5.16). The user sends instructions by moving a pointer on the screen (generally mouse) and selecting certain objects by pressing buttons on the mouse while the mouse pointer is pointing at Figure 5.16 Graphical User Interface them.

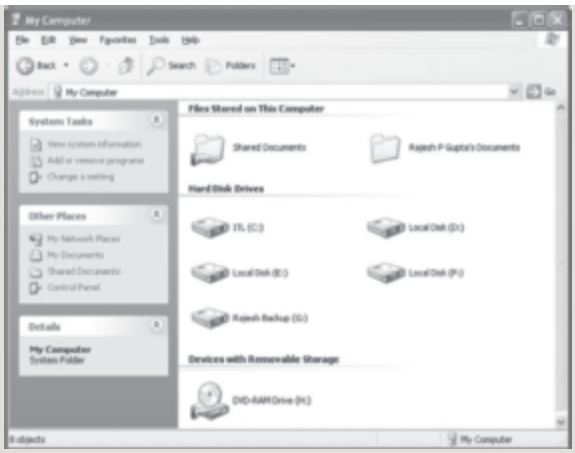


Figure 5.16 Graphical User Interface

LET US SUMMARIZE

1. *OS* is a type of software that controls and coordinates the operation of the various types of devices in a computer system. The two objectives of an OS are controlling the computer's hardware and providing an interface between the user and the machine.
2. OS has six major roles to perform: *process management, memory management, file management, device management, security management* and *providing user interface*.
3. A *process* or task is a portion of a program in some stage of execution. A program can consist of several processes, each working on its own. It may be in one of a number of different possible states, such as new, running, waiting, ready or terminated.
4. A *thread* is the simplest part of a process. To enhance efficiency, a process can consist of several threads, each of which execute separately.
5. In the *uniprogramming* system, only one process can exist at a time while in the *multiprogramming* system, multiple processes can be initiated at a time.
6. Deciding which process should run next is called *process scheduling*. Process scheduling is necessary, so that all programs are executed and run fairly.
7. *Preemptive switching* means that a running process will be interrupted (forced to give up) and the processor is given to another waiting process.
8. The process of switching from one process to another is called *context switching*. A period that a process runs for before being context switched is called a *time slice* or *quantum period*.
9. In *first-come-first-served scheduling*, the processes are executed in the order of their arrival in the ready queue, which means the process that enters the ready queue first, gets the CPU first. New processes are placed at the end of the queue.
10. *Round robin scheduling* employs a technique called time slicing. When the time slice is up, the running process is interrupted

and placed at the rear of the queue. The next process at the top of the queue is then started.

11. A process is said to be in a state of *deadlock* when it is waiting for an event, which will never occur. It can occur if four conditions prevail simultaneously; they are mutual-exclusion, circular wait, hold and wait, and no preemption.
12. The part of the OS that manages the primary memory of the computer is called the *memory manager*.
13. *Paging* is a memory management scheme that allows the processes to be stored non-contiguously in memory.
14. *Virtual memory* is a way of showing the main memory of a computer system to appear effectively larger than it really is.
15. The system that an OS uses to organize and keep track of files is known as the *file management system*.
16. A program that controls a device is called the *device driver*. OS's device manager uses this program to let a user use the specific device.
17. A *user interface* is a set of commands or menus through which a user communicates with the system.

EXERCISES

Fill in the Blanks

1. .....is a program which acts as a mediator between the user and the hardware.
2. A table where many pieces of information associated with a specific process, that is, program counter, process state, CPU-scheduling information and so on are stored, is known as.....
3. .....is also called a lightweight process.
4. .....refers to storing the jobs in the buffer so that the CPU can be efficiently utilized.
5. Two main user interfaces, which an OS has, are ..... and .....
6. Four main conditions that cause deadlock to occur are ....., ..., ... and .....
7. A process is said to be in a state of ..... when it is waiting for an event that will never occur.
8. The major tasks accomplished by the memory manager so that all the processes run in a harmonious manner are ..... and .....
9. RTOS stands for .....
10. First-come-first-served and round robin are types of .....

Multiple-choice Questions

1. The OS that is self-contained in a device and resident in the ROM is .....

    1. Batch Processing System
    2. Real-time OS
    3. Embedded OS
    4. Multi-processor OS

2. The example of non-preemptive scheduling is .....

    1. First-Come-First-Served
    2. Round Robin
    3. Last-In-First-Out

4. Shortest-Job-First

3. An interface where facility is provided for entering commands is .....
    1. Menu-driven
    2. Command-driven
    3. Graphic-driven
    4. None of these
4. The OS that allows only one program to run at a time is .....
    1. Batch Processing
    2. Embedded
    3. Real-time
    4. Multitasking
5. The substitution made by the OS between the processes to allocate space is .....
    1. Swapping
    2. Deadlock
    3. Fragmentation
    4. Paging
6. The memory management scheme that allows the processes to be stored non-contiguously in memory is .....
    1. Paging
    2. Spooling
    3. Swapping
    4. None of these
7. The fit policy of a memory manager to place a process in the largest block of unallocated memory is .....
    1. First Fit
    2. Best Fit
    3. Worst Fit
    4. Bad Fit
8. With which memory can the system run programs that are actually larger than the primary memory of the system?
    1. Cache Memory
    2. Primary Memory
    3. Virtual Memory
    4. None of these
9. What allows the user to run two or more applications on the same computer so that he/she can move from one to the other without closing the application?
    1. Virtual Storage
    2. Multi-processing
    3. Multi-tasking
    4. Multiprogramming
10. The scheduler that selects a process from the ready queue and allocates CPU to it .....
    1. Short-term
    2. Long-term
    3. Medium-term
    4. All of these

State True or False

1. OS is a hardware component.
2. Microsoft Windows XP is a GUI-based OS.
3. A document printing process uses spooling.
4. Virtual memory allows for very effective multiprogramming.

5. Another term for time-sharing is multitasking.
6. Railway reservation systems use batch-processing OSs.
7. Round robin is a non-preemptive scheduling technique.
8. Best-fit, first-fit and worst-fit are memory allocation techniques.
9. The directory system of Microsoft Windows operating is hierarchical.
10. The job of splitting the primary memory into segments as the memory is allocated and de-allocated to the processes is known as fragmentation.

Descriptive Questions

1. What is an OS? Explain various types of OS.
2. Define a process. Diagrammatically explain the life cycle of a process.
3. Discuss various types of interfaces in the OS.
4. What is a deadlock? How can it be handled?
5. Explain how memory protection and process allocation is done by an OS.
6. Write down the differences between:
    1. Uniprogramming and Multiprogramming
    2. Preemptive and Non-preemptive Scheduling
    3. Deadlock Avoidance and Deadlock Prevention

ANSWERS

Fill in the Blanks

1. OS
2. Process table
3. Thread
4. Spooling
5. GUI, CLI
6. Circular wait, Mutual exclusion, No preemption, Hold and wait
7. Deadlock
8. Relocation, Protection and sharing
9. Real-time OS
10. Process scheduling

Multiple-choice Questions

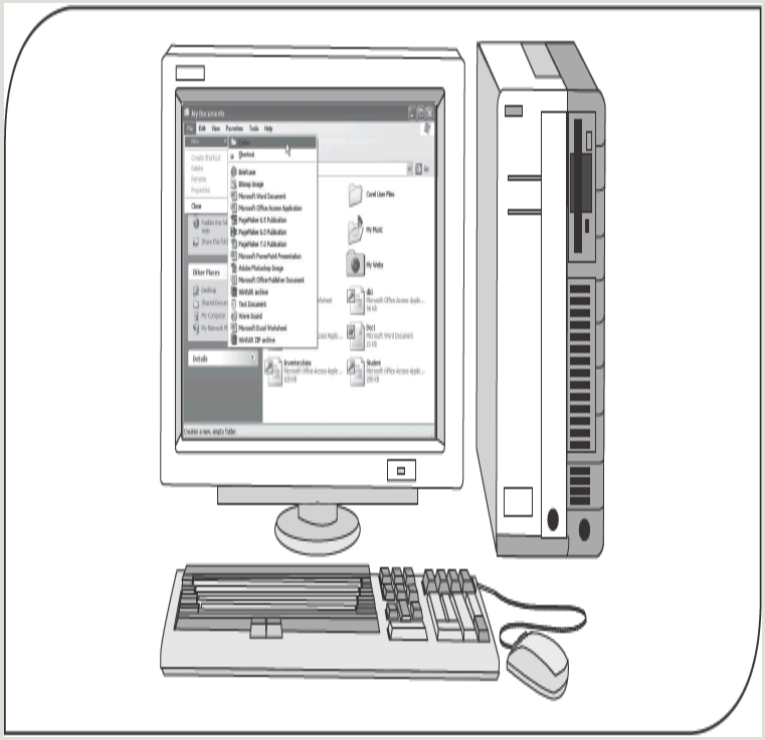
1. (c)
2. (a)
3. (b)
4. (a)
5. (a)
6. (a)
7. (c)
8. (c)
9. (c)
10. (a)

State True or False

1. False
2. True
3. True
4. True
5. False
6. False
7. False
8. True
9. True
10. True

Chapter 6

Microsoft Windows XP



This chapter provides an outlook on Microsoft's Window XP. In this hands-on chapter, the reader will learn how to work with Windows XP using its various in-built features. The chapter also covers various tools and features such as Microsoft help and support center, windows explorer, narrator, magnifier, and system restore. The discussion begins with the basics of the Desktop and My Computer. Files and folders are studied next after which the discussion continues with a focus on various system utilities such as on-screen keyboard, disk cleanup, and disk defragmenter.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The basic outlook on Microsoft Windows XP and how to boot a computer using XP professional
- The Windows XP desktop and its various parts
- The Windows XP My Computer and its various parts
- How to work with files and folders by creating, deleting, moving, renaming and copying them Windows explorer and how to use it for managing files and folders Windows help and support centre and how to use it to get information on various queries
- How to search for files and folders
- Some of the basic utilities provided by Windows XP
- The control panel and how to view and adjust system settings Some of the basic accessories provided by Windows XP

6.1 INTRODUCTION

Windows XP is an operating system introduced in 2001 by Microsoft and is a part of the Windows family of operating systems. Along with a redesigned look and feel to the user interface, this new operating system uses a more stable and reliable environment than previous versions of Windows. The operating system provides a new interface that uses graphics to simplify navigation and completing common tasks, thus making learning and working easier.

Microsoft released Windows XP in two versions: **Windows XP Home** and **Windows XP Professional**. Windows XP Home Edition is designed for home use and small businesses, while Windows XP Professional is specifically designed for advanced home computing, businesses and larger organizations. Windows XP Professional is designed to handle more

powerful computing needs by providing more reliability, security and administrative options. We will be discussing the Windows XP Professional version in this chapter.

6.2 FEATURES OF WINDOWS XP

Windows XP provides a new interface to simplify navigation and perform common tasks (like creating and opening of a file) easily, thus making learning and working easier. These are the various features that make Windows XP a powerful, fast, efficient and stable operating system.

- **Improved Interface**: Windows XP provides a completely redesigned interface Start menu that lets you access programs more quickly. Also, Windows XP keeps a record of programs and files you use more often. In addition, Windows XP displays the most used items in the Start menu while it hides the least used items.
- **More Reliable**: Windows XP is based on the same technology that is used in Microsoft Windows NT and Windows 2000. This makes Windows XP much more stable than Windows 95, Windows 98 and Windows ME, and greatly reduces the number of system crashes and system restarts.
- **Better File and Folder Management**: Windows XP makes it easier to view and work with files and folders. For example, Thumbnail view lets you preview photos and images and the new File and Folder tasks pane lets you easily copy, move, rename or delete any file or folder.
- **Better Help and Support**: Windows XP provides efficient help and support with new features such as Remote Assistance which allows you to contact computer experts and lets them fix a problem on your computer even if they are miles away geographically.
- **More Secure**: Windows XP provides new security tools that help keep your computer more secure and provide new technologies that run in the background, making your computer run more efficiently. This operating system includes the new Security Centre which allows you to check the status of the important security elements on your computer—Windows Firewall, Automatic Updates and virus protection software—making it convenient for you to understand how to keep your computer better protected against viruses and other security threats.
- **Remote Desktop**: Windows XP includes a feature known as Remote Desktop which allows you to access your Windows session from another computer just as if you were in front of your computer. For example, you can connect to your office computer from home, have access to all of your files and network resources as though you were in front of your computer at office. You can leave programs running at work and when you get home, you can see your desktop at work displayed on your home computer with the same programs running.

- **ClearType**: Windows XP provides a special feature called ClearType that makes the words on your computer screen look as smooth as the words on a page. This increases the font resolution of the words written in different programs like Notepad and WordPad.

6.3 STARTING WINDOWS XP PROFESSIONAL

When Windows XP boots up, a log on screen appears. This log on screen depends upon your network identification, membership and system set-up. If you are a member of a domain, the **Welcome to Windows** screen appears, informing you that you need to use the (**Ctrl+Alt+Delete**) key combination to begin. If you are not connected to a network, then you can log on to the computer with or without providing any username or password (Figure 6.1). In addition, Windows XP allows you to set up multiple user accounts where each user can use the operating system with his own settings.

Figure 6.1 Log on Window

6.4 THE DESKTOP

Desktop is the first screen visible after logging in to the Windows XP operating system (Figure 6.2). When Microsoft Windows XP is opened, the desktop appears. The desktop is a full-screen display where all Windows

activity takes place. It is part of a graphical user interface, where icons, menus and dialog boxes on the screen represent programs, files and options on your computer.

Figure 6.2 Desktop

The desktop is actually an invisible folder called *Desktop Folder* (you can access this folder by using windows explorer). You can move files and folders to the desktop for easy access. In addition, it allows you to run programs and use a file system without using the command prompt of the operating system. Note that icons are graphic symbols that represent a program or a utility and are one of several different types of objects on the monitor's screen. The desktop has many parts to it, which are discussed in this section.

6.4.1 Taskbar

The taskbar is used as a link to easily open and close programs. It also allows you to switch between the open application programs and files. If you have more than one instance of the same program open at the same time

then similar windows are grouped into one button and labelled with the name of the program (Figure 6.3).

Figure 6.3 Taskbar

THINGS TO REMEMBER

Start Menu

Some of the items stored in the Start menu display a submenu, while others directly start an application or program. A submenu appears when the menu item has a right-facing arrow as shown here (▸). Click this arrow to display the submenu and select the applications/program to start it. Some of the Start menu items are Shut Down, Run, Help and Support, Search, Settings, My Recent Documents, and Programs.

By default, the taskbar is placed at the bottom of the Desktop screen. The taskbar can be moved to any of the four sides of the desktop and can be resized to display more window buttons. Note that the number following the program name on the button indicates the number of open windows represented by the button. The taskbar includes *Start Menu, Quick Launch Toolbar* and *System Tray*.

Start Menu: The Start menu provides a way to open most of the software applications that are installed on the computer. The Start menu has been Customized to provide quick access methods to all the programs and Windows utilities present on the computer. Click Start menu to display a list of items stored in it (Figure 6.4). Some of the most commonly used Start menu items are listed in Table 6.1.

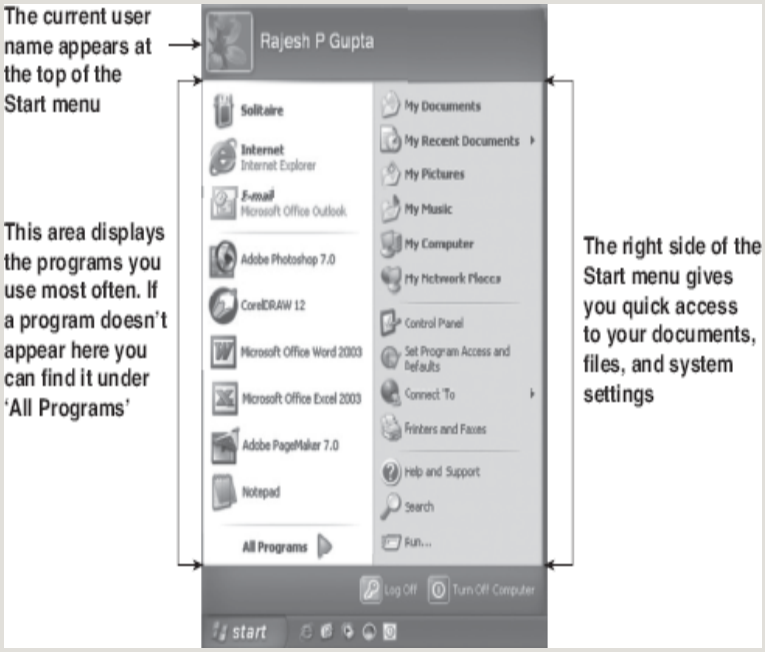


Figure 6.4 Start Menu

Table 6.1 Start Menu Items

Items	Description
Shut Down	This item has several options such as shut down and restart computer, close all programs and log on as a different user, or shut down completely.
Run	Starts a program by typing in the program's name and path. There is a browse button to search for the program to be opened in case the path is not known.
Help and Support	Gives access to the Help feature for the Windows XP environment.
Search	Searches for different types of files and in various locations. Using this feature, you can search for files or even a certain word found in the file.
Settings	Provides quick access to the Control Panel, printer options and taskbar properties.
My Recent Documents	Reopens one of the last 15 files you have opened by clicking on its name.
Programs	Provides access to a list of available programs and submenus of related programs.

Quick Launch Toolbar: The Taskbar also displays its own set of toolbars; the most frequently used is the Quick Launch toolbar. The Quick launch bar is a Customizable toolbar that displays single-click links to programs and commands. By default, the Quick Launch toolbar displays links to Internet Explorer, Media Player and Show Desktop. If this toolbar is not visible, then follow the steps given below:

1. Right-click on the taskbar and select **Properties** to display the **Taskbar and Start Menu Properties** dialog box.
2. Choose the **Taskbar** tab and select **Show Quick Launch**, and click **OK**.

Additional links can be added to the Quick Launch toolbar by dragging the link to the toolbar. To remove links from this toolbar, right-click it and select **Delete** from the short cut menu.

Notification Area: The Notification area (also known as *system tray* in older versions of Windows operating

system) is located on the right-hand side of the taskbar, and generally contains icons of various programs that are running in the background. This includes system applications such as the volume control and Anti-Virus programs (if any). In addition, this area also displays icons of clock and other features related to time and date. To change the system date and time, double-click on the displayed time in the taskbar. This displays the Date and Time Properties dialog box (Figure 6.5). You can change the time, date and the time zone using this dialog box.

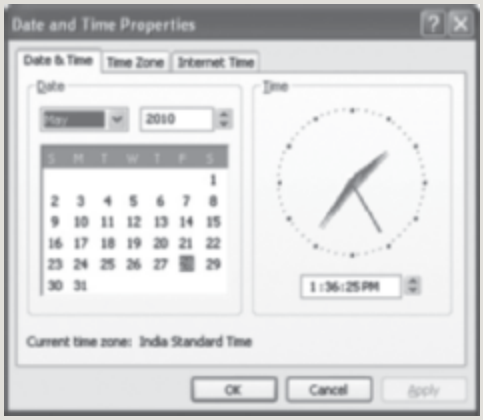


Figure 6.5 Date and Time Properties Dialog Box

6.4.2 Customizing Taskbar

The taskbar can be Customized according to user's preferences. One can Customize the settings such as locking and hiding the taskbar. To modify the taskbar's settings, right-click on an empty area on the taskbar and select **Properties** from the shortcut menu. This displays the **Taskbar and Start Menu Properties** dialog box as shown in Figure 6.6. This dialog box contains two tabs: *Taskbar* and *Start Menu*.

Figure 6.6 Taskbar and Start Menu Properties Dialog Box

The Taskbar tab contains **Taskbar appearance** and **Notification area** section. Taskbar appearance section contains a series of options each having an associated check box. These options are listed in Table 6.2. The **Notification area** section is used to show the clock and hide the inactive icons on the notification area.

Table 6.2 Taskbar Appearance Options

Options	Description
Lock the Taskbar	Locks the taskbar and toolbar on the desktop at its current position so that it cannot be moved to a new location.
Auto-hide the Taskbar	Hides the taskbar and displays it only when the mouse pointer is moved over the taskbar's location.
Keep the Taskbar on Top of other Windows	Ensures that the taskbar is always visible, even when you run a program in a maximized (full screen) window.
Group Similar Ttaskbar Buttons	Displays opened files (more than one) of a similar program as a single button.
Show Quick Launch	Displays the quick launch bar on the taskbar.

6.4.3 Desktop Icons

The desktop includes icons, each having a specific function. Desktop icons can be thought of as symbols that represent a program, shortcut to a program or a folder. When you use Windows XP for the first time, the desktop contains only Recycle Bin. The rest of the icons usually found on other Windows operating systems are not seen on the desktop. To display the required icons follow the steps given below and shown in Figure 6.7.

Figure 6.7 Showing Desktop Icons

1. Right-click anywhere on the desktop, and select **Properties** from the shortcut menu. This displays the **Display Properties** dialog box.
2. Select the Desktop tab and click the **Customize Desktop** button.
3. In the Desktop icons pane of the **General** tab, select the desired icons and click **OK** twice.

The main icons that appear by default on the desktop are listed in Table 6.3.

Table 6.3 Icons on the Desktop

Icons	Name	Description
	My Computer	Displays documents and files. It also provides access to hard drives and removable drives of the computer (both local and network).
	My Network	Places Enables you to communicate with network resources on the network for file sharing purpose.
	My Documents	It is used to store documents. It also contains data files, which can be easily organized.
	Recycle Bin	It is a temporary storage area used for storing deleted files and recover files accidentally deleted from the hard drive or floppy drive.
	Internet Explorer	Provides access to the Internet.

In addition, you can create shortcuts of commonly used applications on the desktop. For this, right-click the desired file or folder and select **Send To** from the shortcut menu as shown in Figure 6.8. This displays another popup menu, from where select **Desktop (create shortcut)** option.

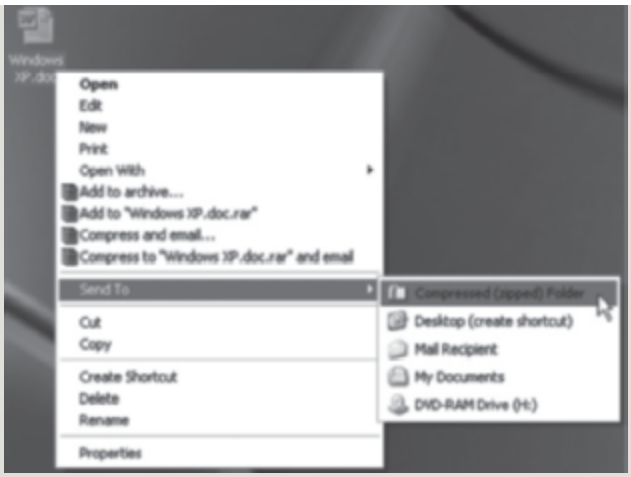


Figure 6.8 Creating Shortcut on Desktop

Organizing Desktop Icons: The above-discussed icons (see Table 6.3) are the basic icons available on the desktop. In addition to these, you can add a number of shortcut icons on the desktop. However, if icons are large in number, the desktop may look untidy and create a problem when working on the desktop. To manage the desktop, Windows XP provides different methods to organize the desktop icons in a number of ways.

The simplest way to organize the icons is to drag them to a new location with the mouse. To use Windows XP's in-built feature for organizing icons, right-click anywhere on the desktop and select the **Arrange Icons By** and then select the desired option(s) from the pop-up menu to arrange the icons (see Figure 6.9). You may arrange the icons by Name, Size, Type or Modified (date). Another way to arrange icons is to select Auto Arrange. Note that if the Auto Arrange option is checked, then you will not be able to drag the icons until it is unchecked. If Align to Grid is checked, then the icons will be arranged in line.

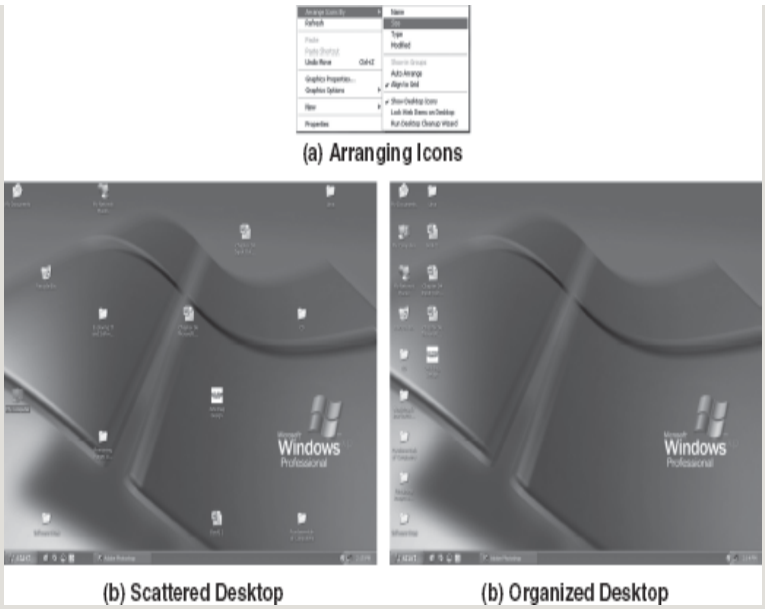


Figure 6.9 Organizing the Desktop

6.4.4 Customizing the Desktop

Windows XP allows you to Customize the desktop according to your preference. The desktop can be changed by modifying the colour scheme of Windows, wallpaper (displayed on the desktop), screensaver and much more. To Customize the desktop, right-click anywhere on the desktop and select **Properties** from the shortcut menu. This displays **Display Properties** dialog box as shown in Figure 6.10. This dialog box contains five tabs at the top of the window. These are *Themes, Desktop, Screen Saver, Appearance* and *Settings*.

Figure 6.10 Display Properties Dialog Box

Themes: Themes are collections of sounds, icons, fonts, colours and screen savers, which often come with their own background graphics, based on a specific theme. You can select from a range of in-built themes from the **Theme** drop-down box. The **Sample** window, displays a preview of the chosen desktop theme.

Desktop: This tab allows you to change the wallpaper (background) of the desktop (see Figure 6.11). Click **Customize Desktop** to show icons of hidden items on the desktop as well as change icons of the selected desktop item. Note that you can also place the wallpaper on the centre of the screen, stretch it to fill the whole screen or tile it over the full screen by using the **Position** drop-down box in the Desktop tab.

Figure 6.11 Customizing Desktop

Screen Saver: The screen saver is an inbuilt utility that appears on your screen when the mouse or keyboard has not been used for a specified period. The screen saver is primarily used to prevent damage to the phosphor element present in the monitor. The screen saver tab allows you to specify the screen saver and its settings (see Figure 6.12).

Figure 6.12 Screen Saver Tab

In this tab, you can also use the **Power** option, which allows you to change the power settings of the monitor. It provides options to save power when the computer is not in use. This is especially useful when the computer is used for a longer duration, for example, the whole day, or when some background operation is going on (like transferring of data between two computers) and the monitor is not required for any operation.

Appearance: This tab allows you to change the windows style, colour scheme and font size. You can choose either the **Windows XP style** or **Windows Classic style** from the **Windows and buttons** drop-down box. Each of these styles has pre-defined colour schemes, which can be selected from the **Color scheme** dropdown box. These schemes are used to provide colour to all the working objects (like icons, title bar). The **Font**

size dropdown box is used to select the size of the font in the folders, icons, title bar and so on (see Figure 6.13).

Figure 6.13 Appearance Tab

This tab also includes the **Advanced** and **Effects** option. The **Advanced** option lets you create a colour scheme of your own choice, while the **Effects** option provides the option to use larger icons, displays shadows under menus, provides fade effects and much more.

Settings: This tab allows you to change the number of colours or display resolution from the **Color quality** and **Screen resolution** section, respectively. If you want to increase the screen resolution from 800 × 600 to 1024 × 768 pixels, then drag the bar available in the **Screen resolution** section (see Figure 6.14). As you increase the pixels, the size of the fonts and icons also

increase but the amount of information available on the screen decreases in size. The **Color quality** section displays the current colour settings for the installed monitor.

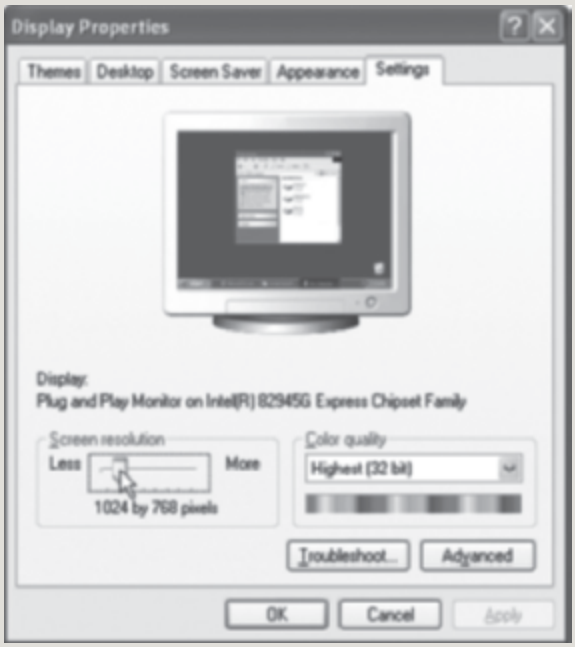


Figure 6.14 Settings Tab

Click the **Advanced** option to display the Properties dialog box from which you can change the settings for the Video adapter installed on the computer. You can also click on the Troubleshoot option to display the Windows **Help and Support Center**.

Once you have finished modifying the settings for the desktop under respective tabs, click **Apply** to implement the new settings and then click **OK** to close the **Display Properties** dialog box.

As stated earlier, **My Computer** allows users to access disk drives, view and manage files and folders, and access other areas of the operating system such as Control Panel and installed Printers. To open the My Computer window double-click the **My Computer** icon ( ) placed on the desktop. This window has many parts to it as shown in Figure 6.15, which are discussed in this section.

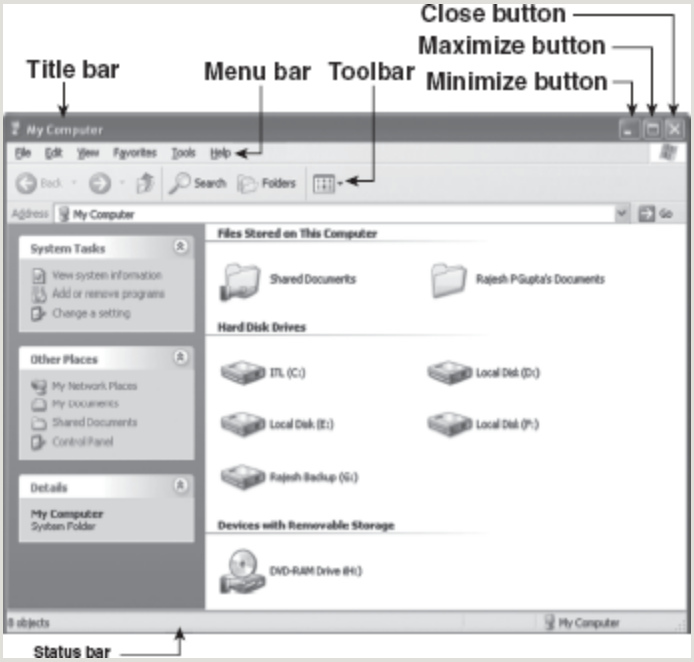


Figure 6.15 My Computer Window

Title Bar: This bar is located at the top of the window and displays the application name. At the extreme upper-left corner of the title bar, there is the **Control Menu** icon. Click this icon to display a pull-down menu, which contains options to change the size and placement of the window. In addition, like all other Microsoft

applications, the title bar also contains the **Minimize, Maximize, Restore** and **Close** buttons.

The **Minimize** button ( ) hides a window without closing it, reducing the application as a button on the taskbar. Using the minimizing function, you can run several programs at one time without cluttering the desktop with several open windows. To restore the minimized application at any time, click the application's button placed on the taskbar. The **Maximize** button ( ) returns a program to full size (maximized) screen. Note that when the window is maximized, the Maximize button is replaced by the Restore button ( ), which returns the window to its original size as it was before the maximized state. The Close button ( ) is used to close the active application.

Menu Bar: This bar contains different submenus, which control all options, functions and commands for all applications present in Microsoft Windows operating systems. Each command in the menu bar represents a menu and is activated by pressing **Alt+** the underlined letter of the menu from the keyboard. By default, My Computer window contains six menus, which have an associated pull-down menu. For example, the **File** menu contains commands to open, create and delete a new file/folder.

Toolbar: Toolbar contains shortcut buttons that are used to perform actions found within the menu bar. By default, there are six buttons present on the toolbar each having a specified function. These buttons are listed in Table 6.4.

Table 6.4 Toolbar Buttons

Command	Button	Description
Back		Moves back to the previously viewed folders/items.
Forward		Moves forward to the last viewed folders/items (if the button is active).
Up		Takes you up one level. If you are working in C:\My Documents, then the Up button displays the contents of the C drive.
Search		Initiates the search function, which allows you to locate files or folders stored on the computer or network.
Folders		Displays the folders pane on the left side of the window.
Views		Allows you to select the way to view the contents of the window.

Toolbars can be customized to include frequently used actions. To customize the toolbar, right-click anywhere on the toolbar and select **Customize** from the shortcut menu. This displays the Customize Toolbar dialog box (see Figure 6.16). In this dialog box, use the **Add** option to add new buttons on the toolbar or use the **Remove** option to delete buttons from the toolbar.

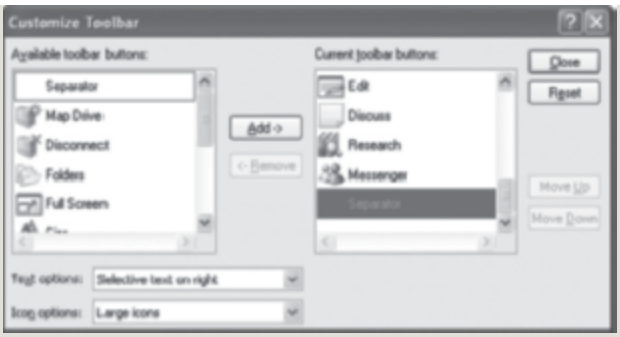


Figure 6.16 Customize Toolbar Dialog Box

Note: *To determine the function of each toolbar button, place the cursor over the button. This shows a tag that displays its function which is also the name of the button.*

Address Bar and Status Bar: The address bar is located between the toolbar and the window panes. The address bar displays the current disk drive and folder, such as: "C:\ Documents and Settings\My Documents". The status bar is located at the bottom of the window and displays the total number of objects (files and sub-folders) in the folder and the space (bytes) they occupy. If a drive or file is selected, it displays the amount of disk free space on the common tasks area as well as on the status bar.

Windows Pane Area: The Window contains two panes: *left* and *right*. The left pane also known as **common tasks area** is divided into various sections such as *System Tasks, Other Places* and *Details* section. System tasks and other places contain links to view information relating to system and access to various system utilities such as the **My Documents** folder and Control Panel. The right pane area includes some icons, which are listed in Table 6.5.

Table 6.5 My Computer Objects

Command	Button	Description
3½ Floppy Drive		Allows saving of files on an external drive, that is, floppy disk.
Local Disk		Allows access to the local drives placed within the computer. Generally, it is the "C" drive, but sometimes a computer has more than one local drive.
CD Drive		Allows access to the CD drive. It serves as a multi-purpose utility allowing you to install software, take back-ups of data, watch and listen to audio visual CDs.
Network Drive		Allows access to a drive that is stored on a network server.

Working with Views: Windows XP allows you to view files or folders with some or no details. The main difference between these views is the amount of details you want to know about a file or folder. To select a particular view, click the **Views** button ( ) on the toolbar. An alternate method is to select the desired view from the **View** menu. There are five views in Windows XP: *Thumbnails, Tiles, Icons, List* and *Details* view (see Figure 6.17). Note that all the views show the name of the files/folders. To know more about a file/folder place the mouse pointer on the desired item, this displays information such as the Author, File Type and File Size.

Figure 6.17 Different Views

- **Thumbnails View**: It displays a small (preview) version of up to four files in the folder.
- **Tile View**: It displays the items with large icons and provides information such as name and size of the application.
- **Icon View**: It displays items in a normal view with no additional information.
- **List View**: It is used to arrange the items in an order.
- **Details View**: It displays items in a tabular form, providing details such as *Type of file* and *size*, which is spread out column wise. You can also sort the item in this view. For example, to sort by type click on the *Type* (column header), to sort by date click on the *Date Modified* (column header) and so on.

A **file** is defined as an organized and structured collection of related *information* stored in secondary storage device. The information can be of different types such as text, data or images. Similar to the way you keep files or documents in file *folders* for easy reference at office or home, computers also contain folders, which hold different documents or files. Folders provide a very convenient and useful method of storing files and subfolders in an organized way on a computer's hard drive or on a removable drive (see Figure 6.18). For example, a folder could be created on the hard drive to keep all business correspondence with a particular customer. This folder can also contain files other than documents for that customer.

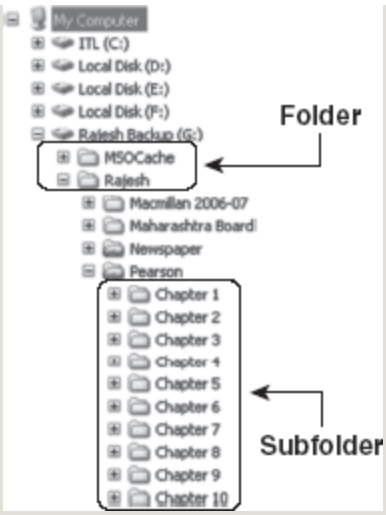


Figure 6.18 Folders and Subfolders

You can also create folders inside a folder. For example, if you have a folder named “Sports” then you can create folders such as cricket, football and hockey for holding various files in these folders. Here cricket, football and hockey are called *subfolders*.

Creating Files and Folders: Since the hard drive can store tens of thousands of files, generally a common practice is to create folders to help in keeping files organized on the hard drive. Folders can be created on **Desktop**, in the My Documents folder or at any other specified location. To create new folders, say on the desktop, right-click on the desktop and select **New** from the shortcut menu. This displays another pop-up menu, from where select the **Folder** option (see Figure 6.19).

Figure 6.19 Creating a Folder

By default, this creates a new folder named **New Folder**. Note that the name of the folder is already highlighted, so it can be assigned a *new name* easily (see Figure 6.20).

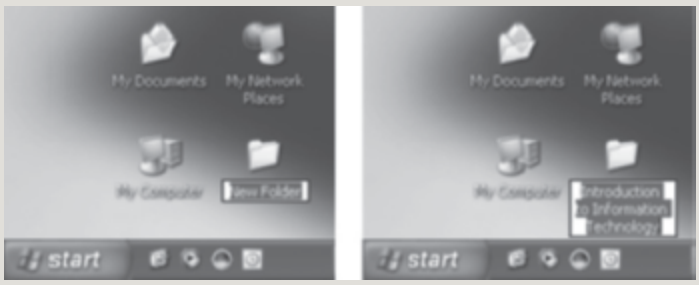


Figure 6.20 Naming a Folder

To create a new folder in the My Documents folder or in *any* other folder, follow any of the steps given below.

FACT FILE

Naming a File

Filenames can contain up to 255 characters and can contain numbers, letters (upper case or lower case), spaces and some symbols. Most files have an extension in addition to the name. The extension is assigned by the program that creates the file. For example, Microsoft Word creates a file with .doc extension.

- Right-click inside the folder and select **New** from the shortcut menu. This displays another pop-up menu, select the **Folder** option from there (see Figure 6.21).

Figure 6.21 Creating a Folder in My Documents Folder

- Select **New** from the **File** menu. This displays a pop-up menu, select the **Folder** option from there.

Moving and Copying Files and Folders: When you are working on a file or folder, sometimes you may decide to move a file or a folder or copy the same file or folder into some other folder or location. Copying a file or folder is similar to the cut operation, except that the copy command retains the file or folder at the original place, but in case of cut, the file or folder is removed from the original location. To move or copy a file or folder, select it and follow any of the steps given below.

- Right-click on the file or folder name and select **Cut** or **Copy** from the shortcut menu (see Figure 6.22). Then select the destination folder or location and select **Paste** from the shortcut menu.

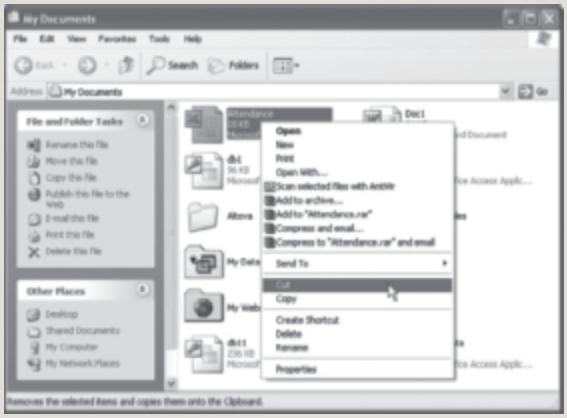


Figure 6.22 Performing Cut or Copy Operation

- Press **Ctrl+X** to cut or **Ctrl+C** to copy using the keyboard. Then select the destination folder or location and press Ctrl+V using the keyboard
- Select **Cut** or **Copy** from the **Edit** menu. Then select the destination folder or location and select **Paste** from the **Edit** menu.
- Drag the file or folder into another location. Note that this operation accounts only if the file or folder is to be moved from one location to another location of the same hard drive (see Figure 6.23).

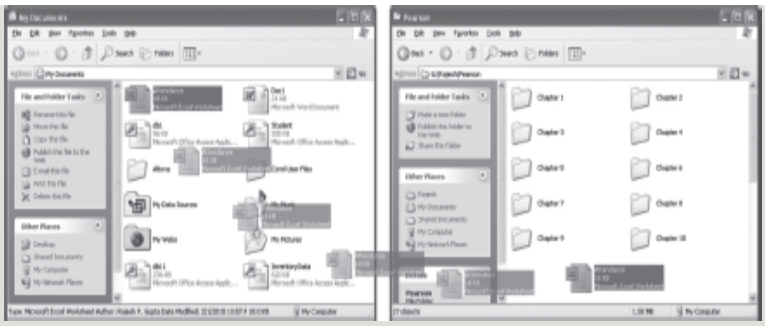


Figure 6.23 Dragging File or Folders

Renaming Files and Folders: Sometimes it may be required that a file or folder should be renamed. To rename a file or a folder select the file or folder and then follow any of the steps given below.

- Right-click on the file or folder name and select **Rename** from the shortcut menu. **Enter** the new name and press the Enter key.
- After selecting, click the file or folder a second time (do not double-click) to highlight it, then enter the new name and press the **Enter** key.
- Select **Rename this file** or **Rename this folder** option from the common task area.

Deleting Files and Folders: If you wish to delete a particular file or folder, select the file or folder and follow *any* of the steps given below.

THINGS TO REMEMBER

Deleting File/Folder Permanently

Files or folders deleted from the hard drive will be moved to the Recycle Bin. Items in the Recycle Bin take up disk space on your hard drive until they are permanently deleted. If you want to permanently delete a file or folder instead of sending it to the recycle bin, select the file or folder and press

Shift+Delete key from the keyboard. If you perform this action, file/folders are deleted permanently. Files or folders stored on a floppy disk or network drive are not sent to the Recycle Bin when deleted, instead, they are permanently deleted.

- Right-click on the file or folder and select **Delete** from the shortcut menu.
- Press the **Delete** key from the keyboard.
- Select **Delete this file** or **Delete this folder** option from the common task area.
- Drag the file or folder into the recycle bin.

When you perform any of the above-mentioned actions, Windows XP displays a message on the screen asking you to confirm deletion of file or folder. Select **Yes** to delete the file, or **No** to keep the file.

6.6.1 Recycle Bin

The Recycle Bin is a temporary storage area where you can move any files or folders that you no longer need. To view the contents of the recycle bin, double-click on the **Recycle Bin** ( ) icon placed on the desktop. If the recycle bin has items stored in it, then the recycle bin icon changes to ( ). This opens the recycle bin window, which contains a list of deleted items. Items in the recycle bin take up disk space on your hard drive until they are permanently deleted. An item remains in the recycle bin until you decide to permanently delete it or restore it.

Restoring Items: You can restore any accidentally deleted file or folder from the recycle bin. To restore an item, select the item(s) and follow any of the steps given below (Figure 6.24).

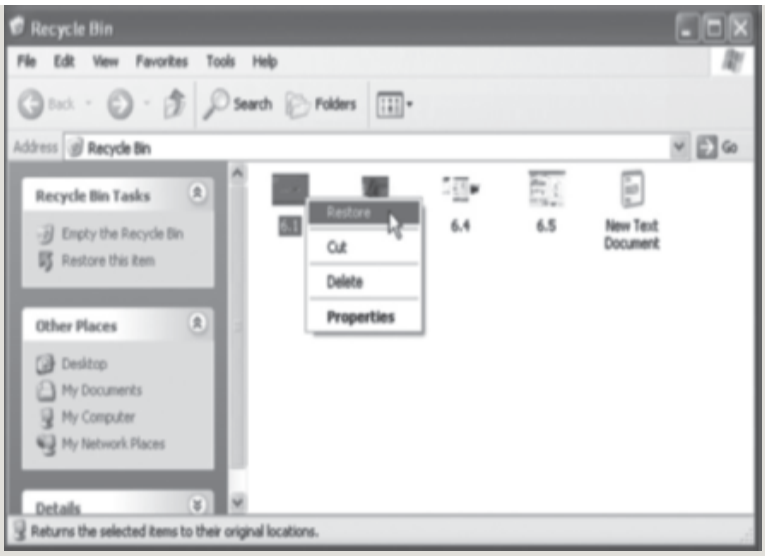


Figure 6.24 Restoring an Item

- Right-click on the item and select **Restore** option from the shortcut menu.
- Select **Restore** this item from the common task area.
- Select **Restore** from the **File** Menu.

Emptying Recycle Bin: You can permanently delete unnecessary items from the recycle bin. To empty the recycle bin follow *any* of the steps given below.

- Right-click on the **Recycle Bin** ( ) icon on the desktop and select **Empty Recycle Bin** from the shortcut menu.
- Select the **Empty Recycle Bin** option from the common task area.

When you perform any of the above-mentioned actions, Windows XP prompts you to confirm the deletion. Click **Yes** to empty the recycle bin or **No** to keep the items in the recycle bin.

To delete the files or folders manually, open the recycle bin window. Now select the item(s) to be deleted and follow any of the steps given below.

- Right-click on the item and select the **Delete** option from the shortcut menu.
- Select **Delete** from the **File** menu.

When you perform any of the above-mentioned actions, Windows XP prompts you to confirm the deletion. Click **Yes** to permanently delete the item.

Note: *Recycle bin contains files deleted from the hard drive only. Files or folders stored on a floppy disk or network drive are not sent to the recycle bin when deleted. Instead, they are permanently deleted.*

Windows XP provides a very useful feature called **Windows Explorer**, which helps you to organize files and folders stored on a computer in an efficient manner. It allows you to arrange files according to your own needs, and provides optional views for easy scanning of folder contents. In this window, you can also create new files or folders, open, move or copy, and delete existing file or folders. To open Windows Explorer, follow any of the steps given below.

- Click **Start**, point to **All Programs**, and then select **Accessories**. Now select **Windows Explorer** from the submenu that appears (see Figure 6.25).

Figure 6.25 Opening Windows Explorer

- Right-click on **My Documents** folder and select **Explore** from the shortcut menu.

When you perform any of the above-mentioned actions, Windows XP displays the Windows Explorer window. This window is divided into two panes: **Folders** pane and **Contents** pane as shown in Figure 6.26. The folders pane of Windows Explorer displays disk drives, folders and subfolders. These items can be expanded or collapsed by clicking on the plus (+) sign and minus (-) sign symbols that appear to the left of the drive letter, folder or subfolder name. A (+) means there are subfolders that are not visible. Click on the (+) to view the subfolders. Doing this the (+) changes to a (-). The contents of the selected drive or folder are displayed in the contents pane by clicking on the item's name or icon in the folders pane. The toolbar present on the Windows

Explorer window provides buttons for common actions associated with the contents present in the window. These buttons have been listed in Table 6.4.

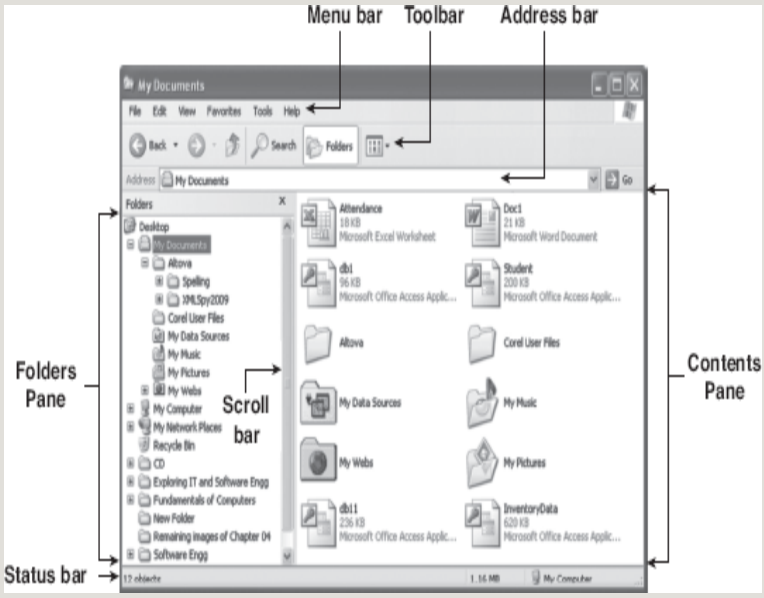


Figure 6.26 Windows Explorer Window

Note: *Both panes may contain horizontal or vertical scroll bars, which are used to display contents not visible within the current window.*

Changing Sort Order: While searching for a file, it may be easy to locate that file in the contents pane by changing the order in which the files and folders are displayed. You can sort the contents pane by clicking on the desired column heading. For example, in Figure 6.27, the contents pane is sorted by the **Type** column, in ascending order. Note that a triangle will appear next to the column heading to indicate the sort method. For ascending, the triangle points upward ( ) while for descending, the triangle point downward ( ).

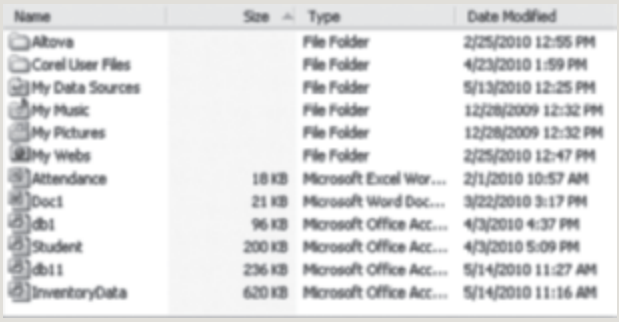


Figure 6.27 Changing Sorting Order I), while for descending, the triangle points

6.8 WINDOWS HELP AND SUPPORT CENTRE

Microsoft Windows XP provides an exhaustive **Help and Support Centre**, which includes help topics, tutorials, troubleshooting and other support services. Windows help is of great use when you are experiencing difficulty using any application or if you want to know more details about some application. To get help, follow any one of steps given below.

- Press the **F1** key.
- Select **Help and Support Centre** option from the **Start** menu.

When you perform any of the above actions, Windows XP opens Help and Support Centre. This centre allows you to find support information by selecting from a menu of topics, using **Search** or **Index** feature as shown in Figure 6.28. Search allows you to search for a specific item. While, index feature allows you to view the list of help topics in an alphabetical order.

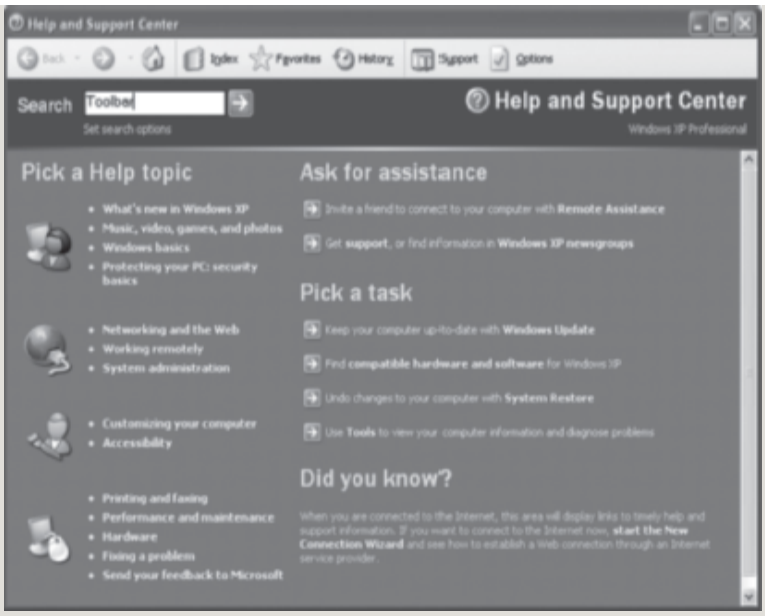


Figure 6.28 Help and Support Centre

To use the search option type the keywords in the **Search** box and click on ( ) or press the **Enter** key. This opens a new window showing the search results for the help topic. To use the **Index** feature select **Index** from the toolbar. Now type the item you want to search for in the **Type in the keyword to find** box and click **Display** to show help information.

The window pane area of Help and Support Centre groups tools and topics under four different sections: *Pick a Help topic, Ask for assistance, Pick a task and Did you know?*

- **Pick a Help Topic:** This section provides general information about help topics. For example, you can learn about what's new in XP, basic system use, networking, remote system operation, printing and security.
- **Ask for Assistance:** This section provides a link to Window's XP **Remote Assistance** feature. This feature lets a Windows XP user to connect to the Internet for assistance by browsing through your computer. You can also access the Microsoft Product Support Services

(PSS) Web site and XP-related newsgroups from this section. Note that to perform this, your computer should be connected to the Internet.

- **Pick a Task:** This section provides a link to the most commonly used items. For example, you can access the Windows Update feature (with which you can update the OS with patches, updates and much more), find Windows XP compatible hardware for the computer, restore the computer back to a previously working state, and run diagnostic tools such as system maintenance and information tools to test your computer.
- **Did You Know?:** This section contains links to common support topics and Windows XP related FAQs (frequently asked questions). If the Internet connection is available then Windows XP automatically updates this section with data from the PSS website.

6.9 SEARCHING IN WINDOWS XP

It is quite common to forget where you have saved a document or any other file, and it may take a long time to manually search all your folders. The Search command is an invaluable tool in helping you locate a document, file or folder. You can choose to search by various criteria including file name, a specific word present in the file or the last Modified date. Windows XP allows you to search for media files, such as pictures, music and video; documents, such as those created with word processing and spreadsheet programs; or all files and folders.

Carrying out a Basic Search: By default, Windows XP searches all folders and subfolders in the drive selected. If you want to search a particular folder, you can use the **Browse** option in the **Look in** list to select the folder. Once the criteria have been entered, you can perform the search. To perform a search, follow the steps given below.

1. Click **Start** and select **Search;** this displays the Search Results window (see Figure 6.29).

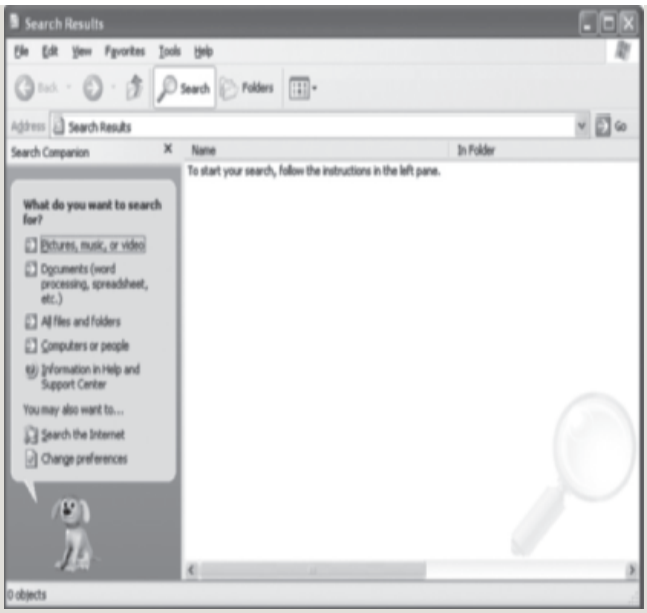


Figure 6.29 Search Results Window

2. A search companion appears on the left-hand side of the window, which asks you what type of an item do you want to search for. Search options are context-sensitive. The Search Companion displays different search fields depending upon the type of item you are searching for.
3. Click on the desired option. The Search companion now asks you to enter all or part of the file name (see Figure 6.30).

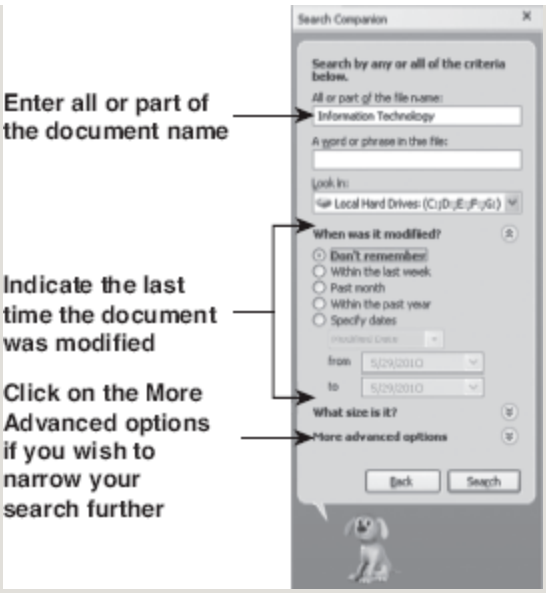


Figure 6.30 Using Basic Search

4. Click the **Search** button to start the search.

Look in the left pane of the main window to see all the search results. To open up any file or folder, simply double-click it (see Figure 6.31).

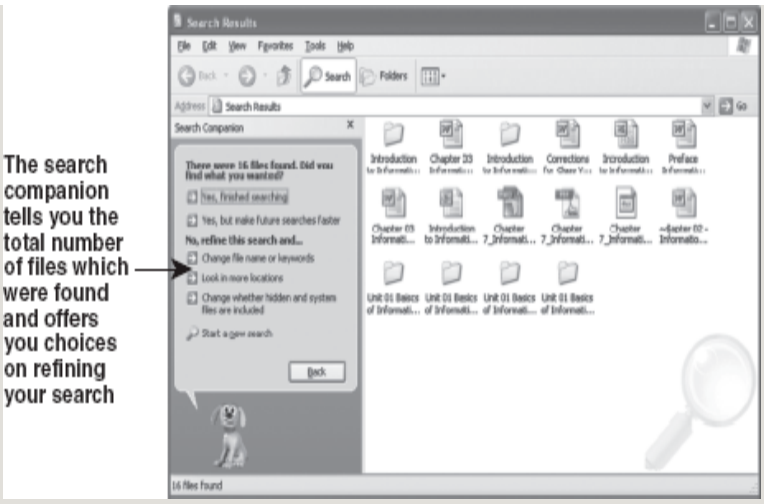


Figure 6.31 Search Results

After the files appear in the Search Results window, you can select any of the following three options.

- If satisfied with the results, select **Yes, finished searching.** This closes the Search Companion, but leaves the Search Results window open with the found files listed.
- If not satisfied with the results, you can use the links under **No, refine this search** and to change your search criteria.
- To carry out another search, you can select **Start a new search.**

Advanced Search with More Criteria Options: If you wish to narrow your search further, you can use the advanced search option. The search companion provides you with many options to consider numerous searching criteria, which helps to narrow down the search, thus making search criteria more accurate and rich. To use advanced search in the Search Results window, follow the steps given below.

1. Select the option to be searched for in the left pane (see Figure 6.32).

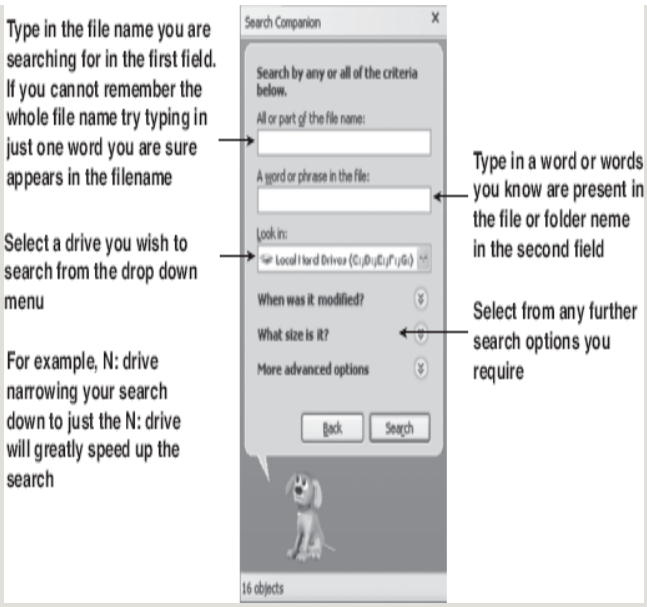


Figure 6.32 Searching Using Advanced Search Option

2. The search companion provides you with options, which can be used to carry out advanced search using options such as selecting a specific hard drive or selecting a specific folder.
3. Click the **Search** button to search for files or folders.

With the help of links, you can search by date, file size or other advanced options. These links include **"When was it Modified?", "What size is it?"** and **More advanced options.**

The **"When was it Modified?"** link allows you to search by the modification date. Date options include searching for files Modified last week, last month, within the past year or you can specify a date range. If you are not sure of the date, you can select **Do not remember,** which is the default.

You can limit the search to files of a certain size using the **"What size is it?"** link. You can specify that files be of a maximum or minimum size, or a specific size. Then, only the files that fall within the size parameters you specify will be found.

The **More advanced options** link includes options for searching system folders, hidden files and folders, subfolders, and tape backups. By default, the **Search system folders** and **Search subfolders** options are selected. You can also perform a **Case sensitive** search. When this option is enabled, the files found must contain the text entered in the **A word or phrase in the file** box, as well as match the case of the text.

6.10 SYSTEM UTILITIES IN WINDOWS XP

All versions of the Windows operating system include built-in tools that help you maintain your system. With these tools, you can ensure good performance from both the operating system and your hard drive. The basic utilities such as Disk Cleanup, Check Disk and Disk Defragmenter help maintain your hard drive to ensure maximum efficiency. More advanced tools enable you to change to a different file system and compress your hard disk. The main utilities are *Narrator, Magnifier,* On-*Screen keyboard, System Restore, Disk Cleanup, Disk Defragmenter* and *Backup*.

Narrator: The Narrator is a text-to-speech utility for users who are visually impaired. Narrator operates by reading the information available on the screen. This includes contents of the active window, menu options or the text you have typed. The Narrator is designed specifically to work with Notepad, WordPad, Control Panel programs, Internet Explorer and Desktop. To open the Narrator, Click **Start,** point to **All Programs,** and then select **Accessories.** Now choose **Accessibility** and select **Narrator** from the shortcut menu that appears as shown in Figure 6.33 (a).

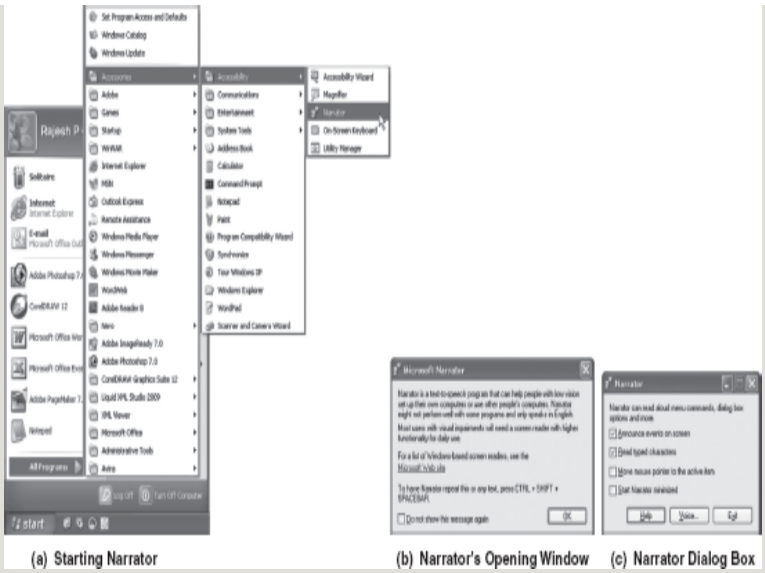


Figure 6.33 Using Narrator

When Narrator is opened, it shows a message, displaying a brief description about Narrator as shown in Figure 6.33 (b). Click **OK** to proceed to the Narrator dialog box. This dialog box contains a list of options, each having an associated check box as shown in Figure 6.33 (c). Select the appropriate options and then select **Voice** to adjust the voice settings.

Note: *Narrator is not available for all languages and is only supported on the English version of Windows XP.*

Magnifier: Magnifier is an accessibility utility that makes the screen more readable for users who have impaired vision. Magnifier operates by creating a separate window that displays a magnified portion of the screen. You can easily change the magnification level, move or resize the Magnifier window, or drag it to the edge of the screen and lock it into place. To open Magnifier click **Start,** point to **All Programs,** and then select **Accessories.** Now choose **Accessibility** and

then select **Magnifier** from the shortcut menu that appears (see Figure 6.34).

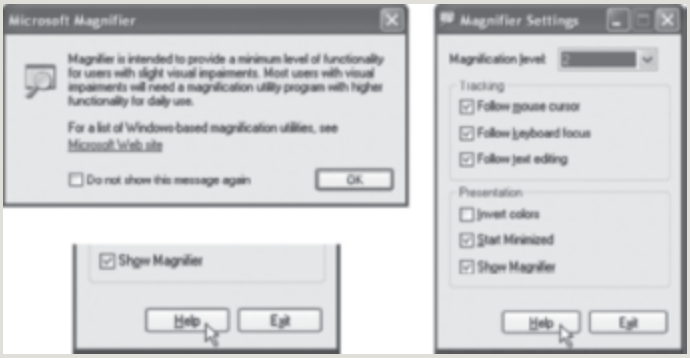


Figure 6.34 Using Magnifier Window

When Magnifier is opened, it displays three overlapped windows. The first window shows a message, displaying brief description about Magnifier. Click **OK** to display the second window, which displays various settings for the Magnifier. The third window shows magnified view of the screen. Note that in Figure 6.34 **Help** and **Exit** buttons appear magnified.

On-screen Keyboard: On-screen keyboard is a utility that displays a virtual keyboard on the screen and allows users with mobility impairments to type data using mouse pointer or a joystick. To open On-Screen Keyboard, click **Start,** point to **All Programs,** and then select **Accessories.** Now choose **Accessibility** and then select **On-Screen Keyboard** from the shortcut menu that appears (see Figure 6.35).

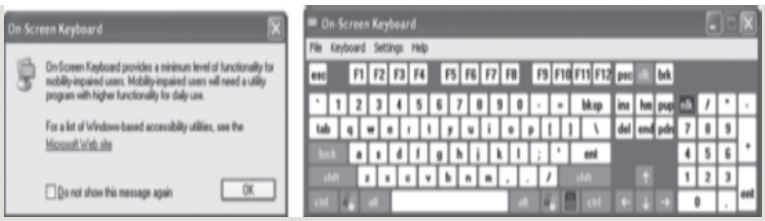


Figure 6.35 Using On-screen Keyboard

When On-Screen Keyboard is opened, it shows a message, displaying a brief description about the program. Click **OK** to access the On-Screen Keyboard. This keyboard includes three typing modes: *Click to select, Hover to select,* and *Joystick or key to select*. To access these modes, select **Typing mode** from the **Settings** menu to display the Typing Mode dialog box (see Figure 6.36).

Figure 6.36 Typing Mode Dialog Box

- **Click to Select**: Specifies that you want to type by clicking keyboard characters with a mouse or joystick.
- **Hover to Select**: Specifies that you want to type by pointing at keyboard characters with a mouse or joystick.
- **Joystick or Key to Select**: Specifies that you want to type by selecting highlighted keyboard characters with designated keyboard keys, mouse or joystick.

Note: *The program in which you want to type characters must be active while you are using On-screen Keyboard.*

System Restore: Microsoft Windows XP has included system utilities that help you recover from systems that become unstable or unreliable. System restore is one such system utility whose purpose is to restore computer to a previously known working state in the event of a problem. This is done without loss of personal files or data such as word processing documents, spreadsheets, music files, and images.

System Restore protects your computer by creating backups of vital system configurations and files. These backups are known as *restore points* and are used to revert the system to a previous state. This utility runs in the background and automatically creates a restore point when a trigger event occurs. Trigger events include application installations, AutoUpdate installations, Microsoft Backup Utility recoveries, unsigned-driver installations, and manual creations of restore points. To open system restore, click **Start,** point to **All Programs,** and then select **Accessories.** Now choose **System Tools** then select **System Restore** from the shortcut menu that appears. When System Restore is opened, it displays three options: **Restore my computer to an earlier time, Create a restore point,** and **Undo my last restoration** (see Figure 6.37).

Figure 6.37 Using System Restore

Creating a Restore Point: By default, Windows XP is configured to automatically create restore points at scheduled times or before some programs are changed. However, you can also manually create restore points. To do this, follow the steps given below:

1. Select **Create a restore point** option and click **Next.**
2. Select the restore point and click **Next.**
3. Confirm the restore point.

Using a Restore Point: To restore Windows XP to a previously created restore point, follow the steps given below:

1. Select **Restore my computer to an earlier time** option and click **Next.**
2. The window displays in bold a list of dates available. Select the date (restoration point) and then click **Next** (see Figure 6.38).

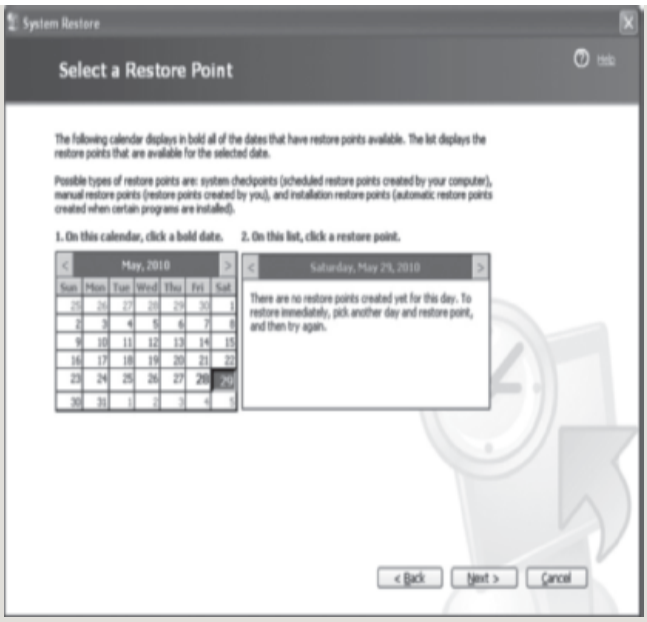


Figure 6.38 Selecting a Date

3. Now Click **Next** to confirm the restoration point.

At this point, you have to confirm, whether you want to continue. Click **Next** to continue. Now, System restore shuts down all open applications and reboots the computer. After the computer reboots, a message appears on the screen, which confirms that the restoration to the restore point is complete.

Undo the Last Restoration: If the computer fails after reverting to a certain date, you can undo the restoration. Simply select **Undo my last restoration** option and then follow the instructions on the screen.

Note: *If you restore your computer to a point before a program was installed, that program will no longer work after restoration.*

Disk Cleanup: Disk Cleanup helps to create free space on your hard drive. Disk Cleanup searches your drive, and then displays temporary files, Internet cache files, and unnecessary program files that you can safely

delete. To use Disk Cleanup, follow the steps given below:

1. Click **Start,** point to **All Programs,** and then select **Accessories.** Now choose **System Tools** then select **Disk cleanup** from the shortcut menu that appears (see Figure 6.39). This displays **Select Drive** dialog box.

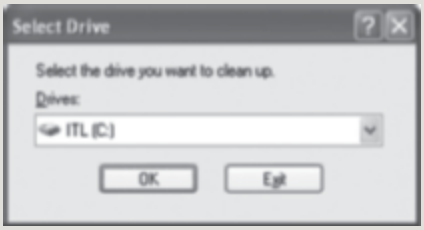


Figure 6.39 Select Drive Dialog Box

2. Select the drive from the drop down box and click **OK.** This displays the Disk Cleanup dialog box (see Figure 6.40).

Figure 6.40 Disk Cleanup Dialog Box

3. On the Disk Cleanup tab, scroll through the Files to delete list and see which kind of files Windows has selected for removal. Select any additional file types you want to delete, then click **OK.**
4. Windows XP displays a message asking you to confirm deletion. Click Yes to delete.

Disk Defragmenter: The Disk Defragmenter utility searches your hard disk for fragmented files that have been stored in non-contiguous sectors of the disk. It then consolidates fragmented files and folders on your computer's hard disk, so that files and folders are stored in contiguous sectors. As a result, you can access files and folders faster and save new ones more efficiently. To use Disk Defragmenter, follow the steps given below:

1. Click **Start,** point to **All Programs,** and then select **Accessories.** Now choose System Tools then select **Disk Defragmenter** from the shortcut menu that appears (see Figure 6.41). This displays the **Disk Defragmenter** window.

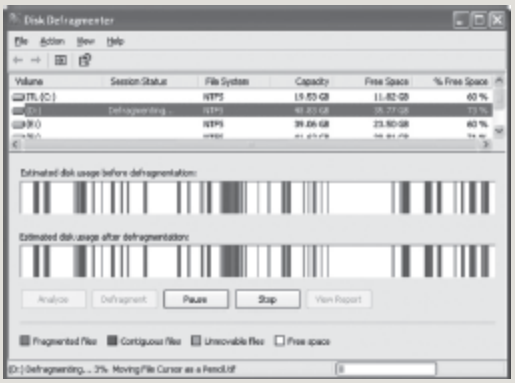


Figure 6.41 Disk Defragmenter Window

2. Select the drive you want to defragment and click **Defragment** to start defragmenting.

6.11 CONTROL PANEL

Control Panel is an important part of Windows XP as it provides various options to view and adjust system settings such as adding a new hardware, adding and removing programs (software), managing user accounts, changing settings of your mouse and keyboard, and so on. To open the Control Panel, follow any of the steps given below:

- Click **Start** and then click **Control Panel.**

- Click **Start** and then click **Run.** This displays the Run dialog box. Type *control* or *control panel* and click **OK** or press the **Enter** key.

As you perform any of the above steps, the Control Panel window appears showing you various shortcut icons for adjusting system settings. Windows XP allows you to view Control Panel in either *Classic View* or *Category View*. In the **Classic View**, all the individual shortcut icons appear with their name while in the **Category View** icons are grouped into categories which are then displayed. Switching between the views can be performed by clicking on the option (*Switch to Classic View* or *Switch to Category View*) from the left side of the Control Panel window. Figure 6.42 displays Control Panel in both views.

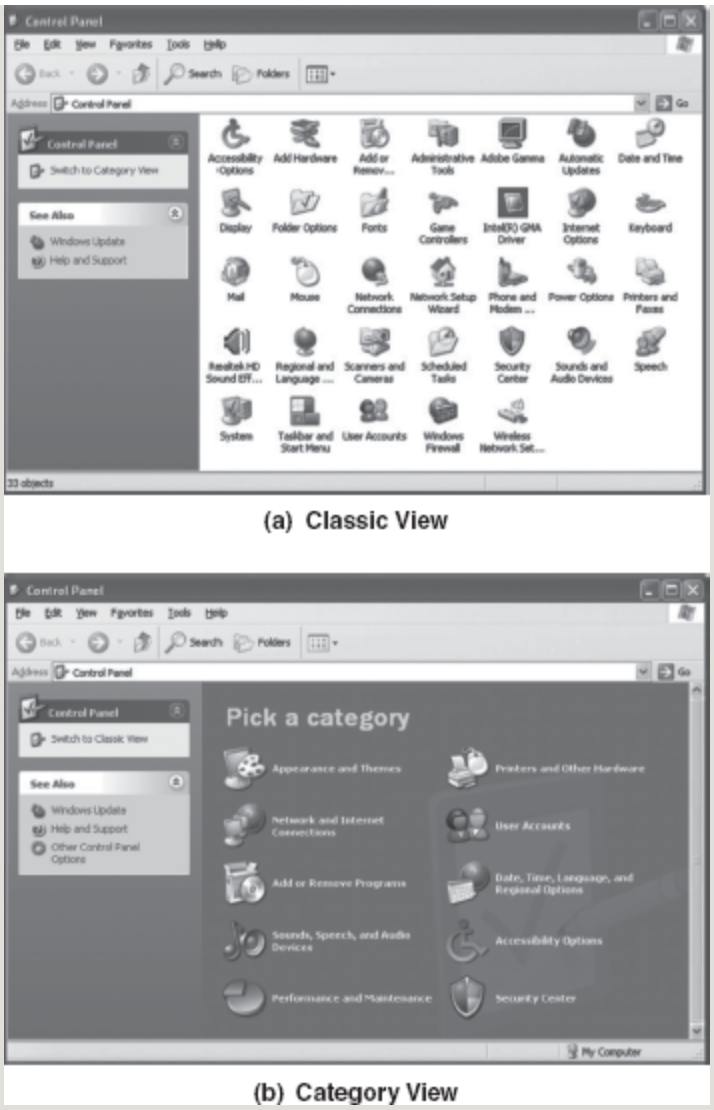


Figure 6.42 Control Panel

The description of most commonly used system settings from Control Panel is listed in (b) Category View Table 6.6.

Table 6.6 Description of Common Settings from Control Panel

Shortcut Name	Icon	Description
Accessibility Options		Comprises settings aimed for disabled.
Add Hardware		Allows you to add a new hardware device to the system.
Add Remove Programs		Allows you to add/remove a program or Windows component. It also indicates the size an installed program is occupying and the frequency of its usage.
Date and Time		Allows you to set (change) the system's date, time, and time zone.
Display		Allows you to change the display characteristics of your computer by changing desktop background, changing or disabling screen saver, specifying color styles of windows elements like menu, adjusting the screen resolution, and so on.
Folder Options		Allows you specifying the settings for files or folders like should folders open in a new window or existing one, should Windows hide system files or folders as well as hidden files, etc.
Fonts		Allows you to view the fonts installed in the computer as well as install new fonts and remove the existing fonts.
Keyboard		Allows you to change the keyboard settings like cursor blink rate.
Mouse		Allows you Customizing the mouse settings like configuring double-click and scroll speed, changing shape of pointer, swapping left and right mouse buttons, adding pointer trail, and many more.
Printers and Faxes		Allows you view the installed printers and faxes. You can remove the existing printer as well as install a new printer from here.
System		Allows you view and specify various settings for system including viewing the version of Windows, amount of RAM, and speed of CPU, changing the name of your computer, enabling or disabling hardware devices installed on your computer, and so on.

Shortcut Name	Icon	Description
User Accounts		Allows you creating, managing, and removing user accounts for persons who share your computer. For instance, you can specify the privileges to a user account, add, change, or remove the password, and so on.

6.12 WINDOWS ACCESSORIES

Windows XP operating system comes with some basic accessories including *Calculator, Notepad,* and *Paint*. Let us discuss them in detail.

6.12.1 Calculator

A built-in calculator is included in the Windows operating system that works just like a hand-held calculator, except you use it on-screen rather than in your hand. The calculator can be used in any one of the two views: *Standard* and *Scientific*. In the **Standard** view of calculator, you perform common mathematical operations such as addition, subtraction, multiplication, division, etc. On the other hand, in the **Scientific** view of calculator, you perform advanced mathematical operations such as trigonometric and statistical operations, calculate logarithm, exponent, factorial and power. To open the calculator, follow *any* of the steps given below:

- Click **Start,** point to **All Programs,** then point to **Accessories,** and click **Calculator** from the submenu that appears.
- Click **Start** and then click **Run.** This displays the Run dialog box. Type *calc* and click **OK** or press the **Enter** key.

As you perform any one of the above steps, Calculator appears on screen. Note that it opens in the same view in which it was used (and closed) the last time. You can switch between the two views from the **View** menu. Figure 6.43 shows both view of calculator.

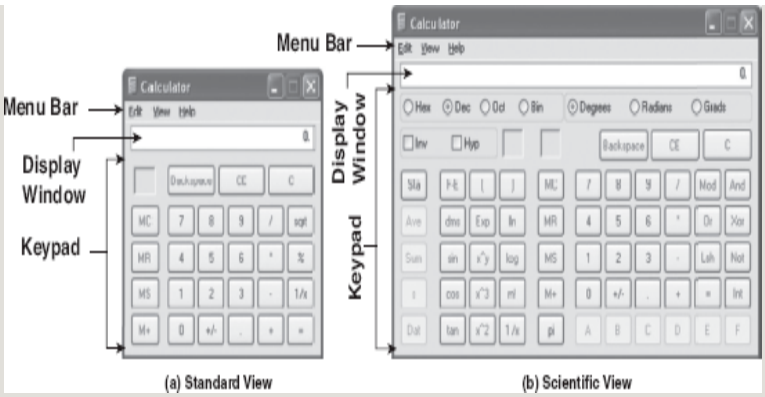


Figure 6.43 Views of Calculator

Note: *You can interact with the calculator through mouse as well as keyboard.*

6.12.2 Notepad

Windows accessories include a simple built-in text editor, Notepad, that help you create simple text documents as well as web pages. The default extension for the files created in Notepad is *.txt*. To open Notepad, follow *any* of the steps given below:

- Click **Start,** point to **All Programs,** then point to **Accessories,** and click **Notepad** from the submenu that appears.
- Click **Start** and then click **Run.** This displays the Run dialog box. Type *notepad* and click **OK** or press the **Enter** key.

As you perform any of the above steps, Notepad window appears on screen, with *Untitled – Notepad* in the title bar (see Figure 6.44). Notepad contains five menus, namely, *File, Edit, Format, View,* and *Help*. Each menu contains commands associated with it. For example, the File menu contains command to create, open, and save a file. Table 6.7 lists all the menus with their associated commands and their description.

Table 6.7 Menus and their associated Commands in Notepad

Menu	Commands	Description
File	New	opens a new Notepad Window
	Open	opens an already saved file
	Save	saves the current file
	Save As	saves the current file with some other name
	Page Setup	sets the size, orientation, and margins of the page and header and footer as well as changes the printer settings
	Print	prints the current file
	Exit	closes the Notepad window
Edit	Undo	undoes the last action
	Cut	cuts the text
	Copy	Copies the text
	Paste	pastes the cut or copied text
	Delete	deletes the selected text
	Find	finds the desired word or the phrase
	Find Next	finds the next instance of the word or the phrase
	Replace	replaces a word or a phrase
	Go To	takes the cursor to the specified line number
	Select All	selects the entire text in a file
	Time/Date	inserts the current time and date
Format	Word Wrap	wraps the text on the line without affecting the way in which text appears on the printed-paper
	Font	changes the font, font style, font size, and the script
View	Status Bar	displays/hides the status bar
Help	Help Topics	provides help on how to use Notepad and its features
	About Notepad	provides information about Notepad

Figure 6.44 Notepad Window

6.12.3 Paint

Windows provides Paint to create and edit graphics. Paint is a simple and easy to use tool and let you work with bitmap (BMP) as well as JPEG and GIF graphics. To open Paint, follow any of the steps given below:

- Click **Start,** point to **All Programs,** then point to **Accessories,** and click Paint from the submenu that appears.
- Click **Start** and then click **Run.** This displays the Run dialog box. Type *mspaint* and click **OK** or press the **Enter** key.

As you perform any of the above steps, Paint window appears on screen, with *Untitled – Paint* in the title bar (see Figure 6.45). The window contains Menu Bar, Tool Box, Vertical and Horizontal Scroll Bars, Color Box, Status Bar, and a white rectangular area (called **canvas**), where you work with graphics.

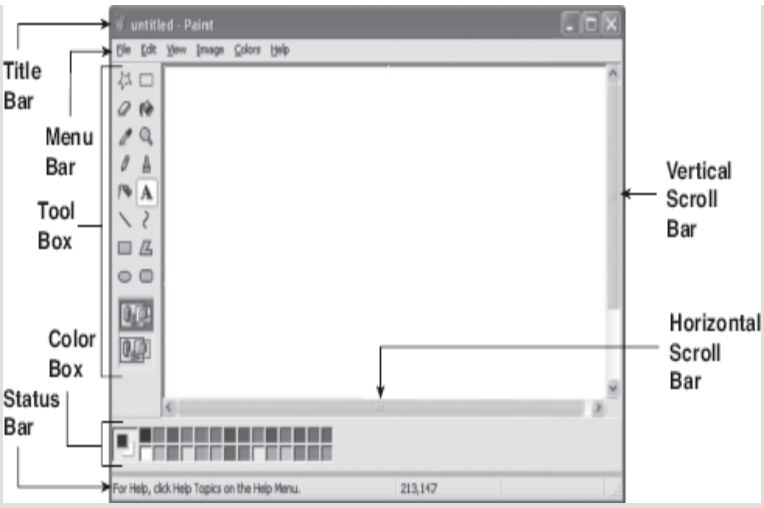


Figure 6.45 Paint Window

Menu Bar: As you can see in Figure 6.45, the Paint contains six menus, namely, File, Edit, View, Image, Colors, and Help. Table 6.8 lists all the menus with their associated commands and their description.

Table 6.8 Menus and their associated Commands in Paint

Menu	Commands	Description
File	New	opens a new Paint window
	Open	opens an already saved file
	Save	saves the current file
	Save As	saves the current file with some other name
	From Scanner or Camera	acquires an image from a scanner or a digital camera–this option is dimmed if a scanner or a digital camera is not installed
	Print Preview	displays the graphic as it will appear on the printed paper
	Page Setup	sets the size, orientation, and margins of the page as well as changes the printer settings
	Print	prints the current file
	Send	sends the graphic as an e-mail
	Set As Background (Tiled)	sets the graphic as a wallpaper in a manner that it covers the desktop with many repetitions of the graphic, if required
	Set As Background (Centered)	places the graphic in the centre of the desktop
	Exit	closes the Paint window
Edit	Undo	undoes the last action–you can go up to three steps back
	Repeat	re-applies the last change made to the file
	Cut	cuts the entire or part of the graphic
	Copy	copies the entire or part of the graphic
	Paste	pastes the cut or copied graphic
	Clear Selection	deselects the selected part
	Select All	selects the entire graphic
	Copy To	copies the graphic to another graphic file
	Paste From	pastes the contents of some other file into your active Paint file

Menu	Commands	Description
View	Tool Box	displays or hides the Tool Box
	Color Box	displays or hides the Color Box
	Status Bar	displays or hides the Status Bar
	Text Toolbar	displays or hides the Text toolbar–this option is available only when the Text tool is selected
	Zoom	allows you to zoom the graphic
	View Bitmap	displays the graphic without the Tool Box
Image	Flip/Rotate	Flip allows you to flip over the image or a selected portion of it horizontally or vertically while Rotate allows you to rotate the image or a selected portion of it–the three options for rotations are 90°, 180°, 270°
	Stretch/Skew	Stretch allows you to elongate or shrink the image in a horizontal or vertical direction while skew allows you to change the angle of the image in a horizontal or a vertical direction
	Invert Colors	displays the image similar to a photographic negative
	Attributes	changes the basic image size, that is, its height and width, which can be measured in pixels, inches, or centimeters
	Clear Image	wipes out the entire image leaving a blank area with the same image and dimension as earlier
	Draw Opaque	works with the selection–if unchecked, any background color in your selection is treated as transparent, otherwise, the background color remains opaque when the selection is moved in the image
Color	Edit Colors	allows you to edit the colors
Help	Help Topics	provides help on how to use Paint and its features
	About Paint	provides information about Paint

Tool Box: The Tool Box contains several tools for creating drawings and text, selecting and erasing part of images, filling colors in images, and so on. Table 6.9 lists all the tools available in the Tool Box along with their description.

Table 6.9 Tools available in the Tool Box

Tool	Buttons	Description
Free-Form Select		used to select the irregular part of an image for copying or cutting
Select		used to select the rectangular area of an image for copying or cutting
Eraser/Color Eraser		used to erase the parts of an image
Fill with Color		used to fill colors in an image with the currently selected color
Pick Color		used to select the color of any portion of the image you click on
Magnifier		used to magnify the image
Pencil		used to draw anything freehand
Brush		used to draw thick lines of different shapes and widths
Airbrush		used to spray color on the images
Text		used to enter text into the drawing area
Line		used to draw straight lines
Curve		used to draw curved lines
Rectangle		used to draw rectangles and squares
Polygon		used to draw polygons
Ellipse		used to draw circles and ovals

Tool	Buttons	Description
Rounded Rectangle		used to draw squares or rectangles with curved corners

Color Box: The Color Box contains several colors and lets you select a color for image. In the left-end of the Color Box is a color overlapping another color. The color on the top is the foreground color and the color in the bottom is the background color. To change the foreground color, left-click a color in the Color Box while to change the background color, right-click a color in the Color Box.

Scroll Bar: Paint window has two scroll bars: *vertical* and *horizontal*. These scroll bars let you scroll and view the portion of image currently not visible on the screen.

Status Bar: The Status bar displays messages and measurements as you work in paint. For example, it displays the information related to a tool as mouse is moved over it in the Tool Box.

LET US SUMMARIZE

1. *Windows XP* operating system is a product of Microsoft and is a part of Windows family of operating system.
2. The new interface of Windows XP provides more stable and reliable environment than the earlier versions of Windows.
3. The desktop of Windows XP contains several icons menus, dialog boxes, and shortcuts to various applications on the screen representing programs, files, and options on the computer.
4. The taskbar, which is by default placed on the bottom of the taskbar, is a link to open and close various applications and programs present in the computer.
5. The taskbar can be customized according to the user's preferences such as locking and hiding the taskbar, showing or hiding the *Quick Launch*, and so on.
6. To organize the desktop icons, right click anywhere on the desktop and select *Arrange Icons* By option to arrange the icons on the desktop according to name, size, or type.
7. *Screen Saver* is an in-built facility that appears on the computer screen, when mouse or keyboard has not been used for a specified period. This facility is used to prevent the phosphor element present in the monitor.
8. *My Computer* icon is placed on the desktop to allow users to access disk drives, view and manage files and folders and also

access *Control Panel* and installed printers.
9. There are five views in Windows XP: *Thumbnails, Tiles, Icons, List,* and *Details*. All these views show the name of files and folders.
10. A file is an organized and structured collection of related information stored in the computer's memory.
11. A folder is a collection of files and subfolders. Any file or folder can be renamed, moved or copied to any other location on the computer.
12. *Recycle Bin* is a temporary storage area where the user moves the files and folders no longer needed. In case, you want to restore the deleted items, right click on the item and click Restore to restore it back to the original location.
13. *Windows Explorer* is a useful feature in Windows XP that allows the users to organize files and folders stored on the computer in an efficient manner.
14. Windows XP provides an exhaustive *Help and Support Centre*. This includes help topics, tutorials, troubleshooting, and other support services.
15. To search the saved documents in the computer, Windows XP provides the *Basic Search* and *Advanced Search*.
16. The main utilities of Windows XP include *Narrator, Magnifier, On-screen Keyboard, Disk Cleanup, Disk Defragmenter,* and *Back up*.
17. *Narrator* is a text-to-speech utility for the virtually impaired people.
18. *Magnifier* is an accessibility utility that makes the screen more readable for users who have impaired vision. It operates by creating a separate window that displays a magnified portion of the screen.
19. *On-screen Keyboard* utility displays a virtual keyboard on the screen and allows users with mobility impairment to type data using mouse pointer and joystick.
20. *System Restore* utility includes system utilities that help in recovering from systems that become unstable and unreliable. It protects the computer by creating back-ups of vital system configurations and files.
21. *Disk Defragmenter* utility searches the hard disk for fragmented files that have been stored in non-contiguous sectors of the disk.
22. *Control Panel* is an important part of Windows XP as it provides various options to view and adjust system settings. Windows XP allows you to view Control Panel in either *Classic* View or *Category* View.
23. Windows XP operating system comes with some basic accessories including *Calculator, Notepad*, and *Paint*.

EXERCISES

Fill in the Blanks

1. By default, the taskbar in Windows XP is located at the.....of the desktop.
2. Windows XP is available in two versions: ..... and .....
3. The desktop contains icons, which are ..... representing any utility or program.

4. You can arrange the icons by name, ... , ..., and date.
5. Screen Saver is used to prevent damage to ..... element in the monitor.
6. Menu Bar contains different ..... to control, options, functions, and commands.
7. ..... view is used to arrange the items in an order.
8. ..... button returns the window to its original size as it was before the maximized state.
9. ..... utility displays a virtual keyboard on the screen.
10. For the users having impaired vision, Windows XP provides a utility called ..... to make the screen more readable.

Multiple-choice Questions

1. Which of the following is by default displayed in Quick Launch toolbar?
    1. My Computer
    2. Show Desktop
    3. Tools menu
    4. My Documents
2. To permanently delete a file or folder from the computer, which of the following keys must be pressed?
    1. Ctrl+Delete
    2. Alt+Delete
    3. Delete+Enter
    4. Shift+Delete
3. Which of the following is not an option in Taskbar and Start Menu Properties dialog box?
    1. Lock the Task Bar
    2. Auto-Hide the Taskbar
    3. Shut Down
    4. Show Quick Launch
4. Which of the following files and folders does Recycle Bin store?
    1. Hidden files and folders
    2. Deleted files and folders
    3. Files and folders in Floppy Drive
    4. None of these
5. Which of the following options is present in the Display Properties dialog box?
    1. Desktop
    2. Appearance
    3. Themes
    4. All of these
6. The functions of Windows Explorer is to .....
    1. Organize files and folders
    2. Change the settings
    3. Add new programs
    4. All of these
7. Which of the following tools is used to draw thick lines of different shapes and widths in Paint?
    1. Brush
    2. Pencil
    3. Magnifier
    4. Airbrush

8. Collections of sounds, icons, fonts, colours, and screen saver are known as .....
    1. Appearance
    2. Themes
    3. Colours
    4. Settings
9. Which of the following utilities is used to search hard disk for fragmented files?
    1. Disk Cleanup
    2. System Restore
    3. Disk Defragmenter
    4. Back up
10. Which of the following is the shortcut key to Windows Help and Support Centre?
    1. F11
    2. F1
    3. Shift+F1
    4. Shift+F11

State True or False

1. The settings of Windows XP once set cannot be changed.
2. Title Bar contains the Maximize, Restore, and Close buttons.
3. Date and Time Properties dialog box displays the system date and time.
4. After a user places a file or folder on desktop, there is no way to arrange the desktop icons.
5. The Windows Help and Support Centre provides the option to search a specific topic in it.
6. The System Restore utility and the Restore button perform the same task.
7. Details view provides the details such as type and size of files and folders in a column wise manner.
8. Network Drive allows access to a drive that is stored on the network server.
9. Files and folders once deleted by Shift+Delete keys cannot be restored from Recycle Bin.
10. User can create shortcut of their files and folders on the desktop.

Descriptive Questions

1. What are desktop icons? Describe some of the desktop icons.
2. Discuss the features of Windows XP.
3. Write short notes on the following.
    1. Address Bar and Status Bar
    2. Narrator
    3. Disk Defragmenter
    4. Themes
    5. Windows Explorer
    6. Calculator
4. What are restore points? How can you create and use restore points?
5. What are the advantages of On-Screen keyboard over physical keyboard?
6. List some of the Start menu items along with their functions.

7. What is Disk Cleanup? List the steps to use Disk Cleanup.
8. Discuss Control Panel.

ANSWERS

Fill in the Blanks

1. Bottom
2. Windows XP Home, Windows XP Professional
3. Graphic symbols
4. Type, size
5. Phosphor
6. Submenus
7. List
8. Restore
9. On-Screen keyboard
10. Magnifier

Multiple-choice Questions

1. (b)
2. (d)
3. (c)
4. (b)
5. (d)
6. (a)
7. (a)
8. (b)
9. (c)
10. (b)

State True or False

1. False
2. True
3. True
4. False
5. True
6. False
7. True
8. True
9. True
10. True

Chapter 7

Overview of Microsoft Windows 7

This chapter familiarizes the reader with Windows 7, which is the latest version of Windows operating system released by Microsoft. The chapter begins with discussion on different available editions of Windows 7, which helps the users to select the appropriate edition according to their requirements. Then discussion moves on to the new features introduced in Windows 7 such as Aero interface, gadgets, jump lists and various Aero effects. Finally, some of the new Accessories programs included in Windows 7 such as Sticky Notes, Math Input Panel and Snipping Tool are discussed.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- Different editions of Windows 7
- New features of Windows 7
- New Accessories offered by Windows 7

7.1 INTRODUCTION

The latest version of Windows operating system, that is, Windows 7 has been released by Microsoft in 2009 to be used on home and business desktops, laptops, Tablet PCs, netbooks and media Center PCs. Windows 7 provides enhanced user interface making it more user-friendly, improved performance to get the tasks done efficiently, new level of security to protect precious data from unauthorized access, and so on.

Advancements in handwriting and speech recognition, support for virtual hard disk and multi-touch, and redesigned taskbar are some of the new features in Windows 7. One more noticeable change is that some of the applications that have been included with prior releases of Windows including Windows Movie Maker, Windows Mail, Windows Calendar and Windows Photo Gallery are not included in Windows 7. The reason for removal is that these applications demand frequent updates than the operating system itself. These applications are now freely available from the Microsoft website as a separate package called *Windows Live Essential*.

7.2 DIFFERENT EDITIONS OF WINDOWS 7

Microsoft has released Windows 7 in six different editions, namely, *Windows 7 Starter, Windows 7 Home*

Basic, Windows 7 Home Premium, Windows 7 Professional, Windows 7 Enterprise and *Windows 7 Ultimate*. Among these editions, only three—Home Premium, Professional and Ultimate—are widely available for sale in most countries; other editions are targeting other markets such as the developing world or are available for enterprise use only.

FACT FILE

Upgrading to Higher Edition of Windows 7

According to Microsoft, installation of Windows 7 stores all the features of every edition on the machine (irrespective of the edition in use). However depending upon the edition, some features may be locked. Users who want to upgrade to a new edition can use *Windows Anytime Upgrade* (*WAU*) to purchase an upgrade, which contains product key only, and unlock the features of that edition by just entering the product key. Hence, upgrading to a new edition of Windows 7 does not require re-installation.

- **Windows 7 Starter:** It is the most basic edition of Windows 7 and contains the fewest features of all the Windows 7 editions. It includes some strange restrictions like the inability to change the default desktop wallpaper. The Starter edition usually comes pre-installed on computers by the computer manufacturers or dealers. It has been specifically designed for notebook computers and netbooks.
- **Windows 7 Home Basic:** It contains features that make them suitable for home users. This edition is not available in some parts of the world including North America, Australia, and Western and Central Europe.
- **Windows 7 Home Premium:** Like the Home Basic edition, this edition is also intended for home users; however, it contains some new features such as touch-screen support, Aero effects, etc., that have been excluded from the Home Basic edition.
- **Windows 7 Professional:** It is suited for small business users as it provides support for being used as a Remote Desktop server, location aware printing, Encrypting File System and so on.

- **Windows 7 Enterprise:** It targets the enterprise segment with features including Multilingual User Interface, UNIX application support and BitLocker Drive Encryption. This edition is not available through retail market; instead it is sold by Microsoft through volume licensing to the companies that have Software Assurance contract with Microsoft.
- **Windows 7 Ultimate:** It contains exactly the same features as the Windows 7 Enterprise edition. The only difference is that the Ultimate edition is available on an individual license basis. Hence, individuals can purchase the Windows 7 Ultimate edition.

Note: *Each edition of Windows 7 includes features that are available in its previous edition.*

7.3 NEW STYLE OF WINDOWS 7

As you start Windows 7, you will be greeted with a new look and style of desktop including Aero themes, gadgets, redesigned taskbar and Start menu, and Jump lists (see Figure 7.1). All the basic elements of Windows like icons, menus, dialog window, etc., have been redesigned to provide better visuals. The new Aero interface of Windows 7 takes advantage of the latest graphics hardware and provides smoother window handling, enhanced animations and glass visual effects (transparencies).

Figure 7.1 Windows 7 Desktop

7.3.1 Aero Themes

An *Aero theme* is the combination of one or more desktop backgrounds, window colour, sounds and screen saver. Some Aero themes include desktop icons and mouse pointers also. You can apply any of the themes to change the desktop background, window colour, sound and screen saver simultaneously. Moreover, you can make the desktop backgrounds to change automatically after a specified time interval using the *desktop slide show* feature of Windows 7.

To create a slide show, follow the steps given below:

1. Right-click on a blank area of the desktop and select **Personalize** from the menu that appears. A window similar to the one shown in Figure 7.2 appears.

Figure 7.2 Personalization

2. Click a theme under the **Aero Themes** section to apply it to the desktop. Note that all the Aero themes except Windows 7 include slide show.
3. Click **Desktop Background** to change the default slide show pictures or settings of the theme. This displays the Desktop Background window (see Figure 7.3).

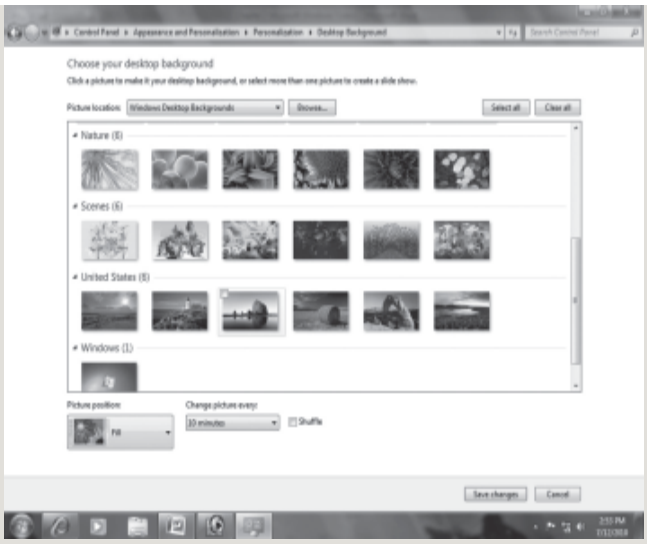


Figure 7.3 Desktop Background Window

4. Select the pictures that you want to add to the slide show. For this, follow any of the steps given below:
    1. To select the pictures provided as a part of Aero themes, place the mouse over the picture that you want to add and select the check box on the top left corner of the picture. Alternatively, click on the desired pictures while holding down the **Ctrl** key (see Figure 7.4).

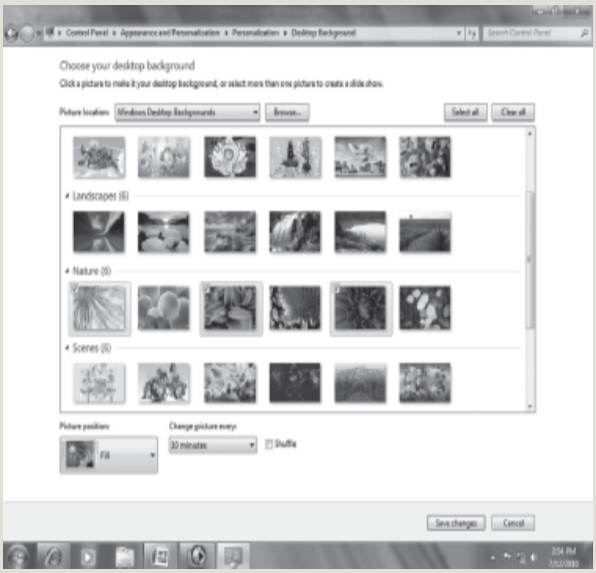


Figure 7.4 Selecting Pictures for Slide Show

    2. To select all the pictures of all the Aero themes at once, click the **Select All** button.

3. To select your own pictures, click the **Browse** button to locate the folder on the computer from which you want to add the pictures.

   Note that you must select more than one picture to create the slide show.
5. Click the down arrow under the **Picture position** and select an option from the list. For example, select **Fill** to crop the pictures to fill the screen, **Fit** to fit the pictures to the screen, **Stretch** to stretch the pictures to fit to the screen, **Tile** to tile the pictures, **or Centre** to place the pictures at the centre of the screen.
6. Click the down arrow under the **Change picture every** option and select the time from the list to specify how quickly the desktop background should change.
7. Select the **Shuffle** check box if you want the pictures to appear in random order.
8. Click the **Save changes** button.

7.3.2 Gadgets

Gadgets are mini programs that provide quick access to information and frequently used tools. For example, you can have a gadget that provides news headlines in order to keep yourself updated. Windows 7 includes many gadgets and you can download more gadgets from Microsoft's web site and install them on the machine.

You can place multiple gadgets or more than one instance of a gadget on the desktop. To add a gadget to the desktop, follow the steps given below:

1. Right-click on a blank area of the desktop and select **Gadgets** from the menu that appears. This displays a window showing all the gadgets installed on the machine (see Figure 7.5).

Figure 7.5 Gadgets in Windows 7

2. Double-click a gadget to add it to the desktop.

Once a gadget has been added to the desktop, you can move it freely (by dragging) to place it wherever you want. In case you do not want a gadget any more on the desktop, you can remove it from the desktop. For this, right-click the gadget and select **Close Gadget** from the menu that appears.

Note: *Every gadget placed on the desktop has some buttons associated with it that appear only when you place the mouse pointer over the gadget. These buttons can be used to close the gadget, drag it, change its options and resize it.*

7.3.3 Taskbar

The taskbar has undergone the major modifications in Windows 7. You may immediately notice that it is taller than the taskbar in the previous versions of Windows. One main reason for making it taller is to enable it to accommodate the touch screen input.

As shown in Figure 7.6, the taskbar has three main sections: the **Start** button ( ), the middle section and

the notification area. The **Start** button, when clicked, opens the **Start** menu (discussed in Section 7.3.4). The middle section of the taskbar is the one that you use frequently. It shows buttons corresponding to the programs or files that you have currently open. By default, it displays only three buttons corresponding to Internet Explorer 8, Media Player and Windows Explorer program. Afterwards, every time you open a program, file or folder, Windows creates a corresponding button on the taskbar. The taskbar button of the active program is always highlighted. Each taskbar button displays an icon that helps you identify the open program. You can click a button to open the respective program as well as can switch between several open programs using the taskbar.

Figure 7.6 Windows 7 Taskbar

At the right side of the taskbar is the notification area (see Figure 7.6). This area contains a clock that tells the system's date and time. In addition, it contains a group of icons that indicate the status of certain programs installed on the computer and certain computer settings. For example, whether anti-virus installed on the machine is enabled, its virus definitions are up to date, and so on. As you move the mouse pointer over an icon in the notification area, the name of the icon or the status of setting appears in a tooltip. Clicking the icon opens the program or associated settings. For example, clicking the

volume icon ( ) shows the volume controls using which you can set the volume.

Note: *Windows 7 hides some icons in the notification area to reduce clutter. These hidden icons can be displayed temporarily by clicking the* **Show hidden icons** *button ( ) in the notification area.*

Thumbnails: Recall that every open window has a corresponding button on the taskbar. When you place the mouse pointer over a button on the taskbar, a small preview of the open window(s) associated with that button appears over the taskbar (see Figure 7.7). This preview is called *thumbnail*.

Figure 7.7 Thumbnails

As you move the mouse pointer over the thumbnail of a window, that window appears on the desktop in front of all other windows (if any). You can click on the thumbnail of any window to make it the active window. In addition, you can close the window by clicking the **Close** button ( ) at the top right corner of the thumbnail.

Note that if you open multiple instances of the same program, then instead of creating a separate button for each instance, Windows 7 creates only one button on the taskbar and shows thumbnail of each instance as you place the mouse pointer over the button.

7.3.4 Start Menu

Start menu provides a common way to perform varied activities such as starting various programs, accessing commonly used folders, searching for files, folders, and programs, getting help for Windows, adjusting the computer settings, etc. You can open the **Start** menu by clicking the **Start** button ( ) on the taskbar or by pressing the **Windows logo** key ( ) on the keyboard.

As shown in Figure 7.8, **Start** menu is divided into three basic parts: *left pane, right pane and search box*. The left pane shows a short list of programs installed on the machine. You can view the complete list of programs by clicking **All Programs**. The right pane provides quick access to the commonly used folders and files as well as the computer settings. In addition, you can log off, restart or shutdown the system from the right pane. The third part of the **Start** menu is the search box. The search box lets you search files, folders or whatever you want. As you type in the search box, Windows searches instantly and provides you with the search results. This makes the searching easier and faster. You can also open programs by typing the name of their executable file in the search box. For example, type winword to open Microsoft Word.

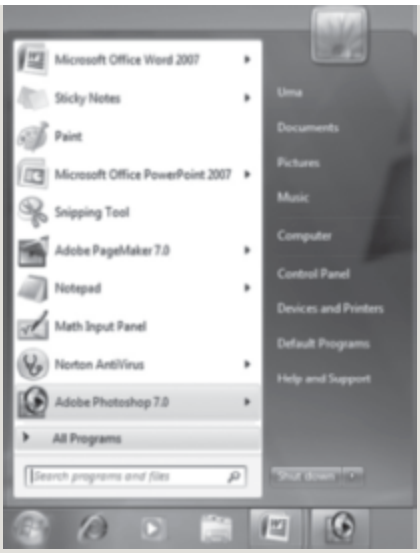


Figure 7.8 Start Menu

7.3.5 Aero Effects

Windows 7 comprises several Aero effects that not only tend to improve the visual experience but also ease the task of managing different windows. These effects include *Aero Flip 3D, Aero Peek, Aero Shake* and *Aero Snap*.

Aero Flip 3D: *Aero Flip 3D* allows you to preview all the open files and folders simultaneously without clicking on the taskbar. For this, press the **Windows logo** key (⊞)+**Tab** key. As you press these keys, Windows 7 arranges all the open windows in a three-dimensional stack (see Figure 7.9) that you can easily flip through.

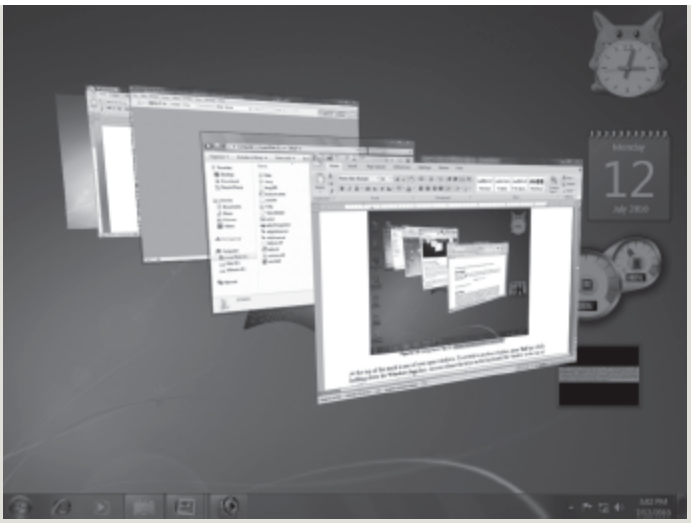


Figure 7.9 Using Aero Flip 3D

At the top of the stack is one of your open windows. To switch to another window, press the **Tab** key while holding down the **Windows logo** key. As you release the keys on the keyboard, the window at the top of stack becomes your active window.

Aero Peek: While working, there may be time when you have multiple windows open (and stacked onto one another) simultaneously. To view the desktop by minimizing all the open windows one by one is time-consuming. Alternatively, you can use Windows 7 *Aero Peek* that allows you to temporarily view the desktop. For this, just place the mouse pointer over the **Show desktop** button (a semi-transparent rectangular button) at the right-most corner of the taskbar. Doing this fades all the open windows to let you temporarily view the desktop along with its folders and gadgets and Sticky Notes placed on it (see Figure 7.10). As you move the mouse pointer away from the button, all the windows get restored automatically.

Figure 7.10 Viewing Desktop using Aero Peek

In addition to temporarily view the desktop, you can click the **Show desktop** button to minimize all the open windows to let you see the desktop. In that case, clicking the **Show desktop** button again will restore the open windows.

Aero Shake: *Aero Shake* allows you to minimize all the open windows except the desired one. This saves time when you want to focus on single window and minimize all other open windows. For using this feature, click on the title bar of the window you want to keep open and drag the mouse back and forth quickly. Note that this minimizes all the open windows except the one that you are shaking. In addition, some windows like dialog boxes cannot be minimized. Hence, these remain open on the desktop. To restore all the windows, shake the open window again.

Aero Snap: You can arrange two open windows side by side on the screen using Windows 7 *Aero Snap* feature. This is helpful while comparing contents of two files or copying files from one folder to another. To arrange two windows side by side on the screen, follow the steps given below:

1. Click the title bar of a window and drag it to one side of the screen.
2. Release the mouse when you see an outline of the window (see Figure 7.11 (a)). The window will expand to fill the outline.
3. Perform step 1 and step 2 for another window. But this time, drag to an opposite side on the screen (see Figure 7.11 (b)).

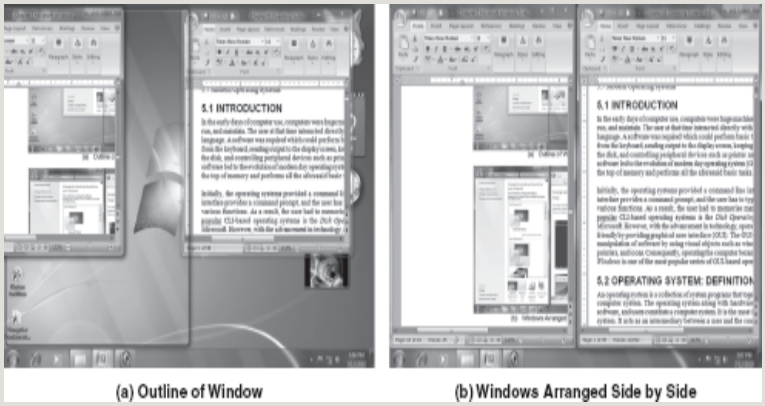


Figure 7.11 Using Aero Snap

7.3.6 Jump Lists

Jump list is a new feature of Windows 7 that provides you quick access to the recently and frequently accessed items (such as files, folders or websites) of some specific program that is used to open those items. You can open any item appearing in the Jump list by just clicking on it.

Jump lists are available for the programs that are pinned to the **Start** menu or the taskbar. In addition, Windows 7 offers Jump lists on the taskbar for the currently

running programs also. To view the Jump list for a program on the **Start** menu, point to any program in the left pane of the **Start** menu that contains a right arrow (▸) beside it. As you do this, you will see a Jump list corresponding to that program in the right pane. Figure 7.12 (a) shows the Jump list for the Microsoft Office Word 2007 program in the right pane of the **Start** menu. On the other hand, you can view the Jump list for a program on the taskbar by right-clicking on the program icon. Figure 7.12 (b) shows the Jump list for the Windows Explorer program on the taskbar.

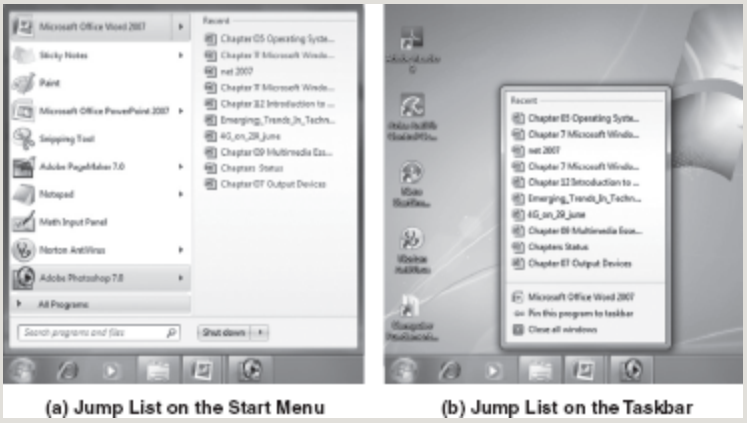


Figure 7.12 Jump Lists

7.4 WINDOWS 7 ACCESSORIES

Windows 7 includes a wide variety of accessories/programs such as Calculator, Command Prompt, Connect to a Network Projector, Math Input Panel, Snipping Tool, Sound Recorder, Sticky Notes, etc. All these programs are grouped under the **Accessories** folder available in **All Programs** on the **Start** menu (see Figure 7.13).

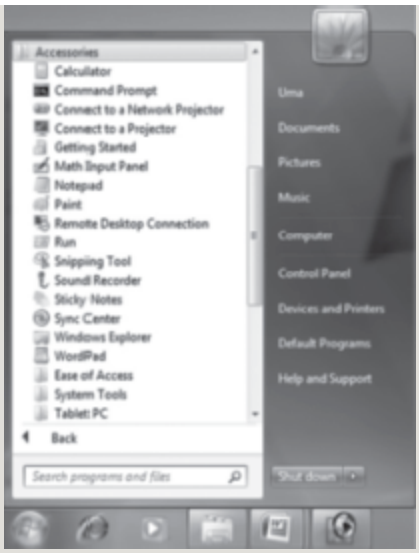


Figure 7.13 Windows 7 Accessories

Some of the Windows 7 accessories are already available in earlier versions of Windows (Windows XP) while others are entirely new. Let us discuss some of the new accessories offered by Windows 7.

Sticky Notes: *Sticky Notes* are used to write down something like a to-do list, someone's phone number, a note to remind you anything, etc. They are placed on the desktop. Figure 7.14 displays a Sticky Note with some text in it.

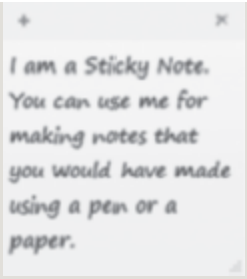


Figure 7.14 A Sticky Note

Creating a Sticky Note is simple. Just follow the steps given below:

1. Click **Sticky Notes** under the **Accessories** folder on the Start menu. A Sticky Note appears on the screen with the blinking cursor.
2. Type the text you want.

Once a Sticky Note has been added, you can add more by clicking the **New Note** button ( ) on its top-left corner. Alternatively, you can press **Ctrl+N** to add more Sticky Notes. You can also resize a Sticky Note. For this, position the mouse pointer on any side of the Sticky Note and when mouse pointer changes to a double-ended arrow, click and drag.

Note: *To close a Sticky Note, click the* **Delete Note button** *( ) on its top-right corner.*

Math Input Panel: Sometimes, you may need to use mathematical equations while working in a document. In earlier versions of Windows, it was quite difficult to specify mathematical expressions. However, Windows 7 makes it easier for you to write mathematical equations in a natural way with the help of *Math Input Panel*. This program allows you to directly write the simple or complex mathematical expressions just like on a notebook paper. As you write, the in-built math

recognizer of Windows 7 recognizes your handwritten expression. You can then transfer the recognized math into your document.

The Math Input Panel basically has been designed to be used with a tablet pen on a tablet PC but you can also use it with other input devices such as mouse, touch screen or external digitizer. To use the Math Input Panel, follow the steps given below:

1. Click **Math Input Panel** under the **Accessories** folder on the **Start** menu. Alternatively, type *mip* in the search box on the **Start** menu and press **Enter**. The Math Input Panel program basic window consisting of a menu bar and writing area appears (see Figure 7.15).

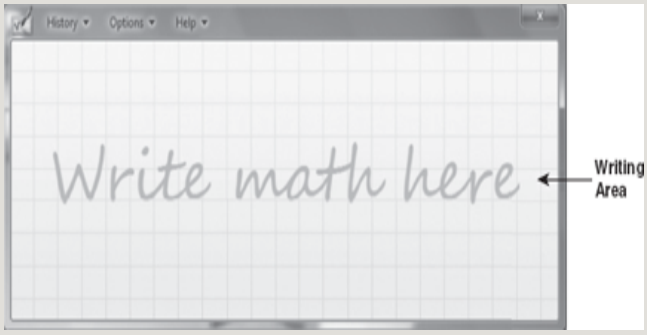


Figure 7.15 Basic Window of Math Input Panel

2. Write the mathematical expression in the writing area. As you start writing, the basic window converts to markup window that shows the preview area and editing tools (see Figure 7.16). The preview area shows the recognized math. If some part of the expression has not been recognized correctly, it can be corrected using the editing tools.

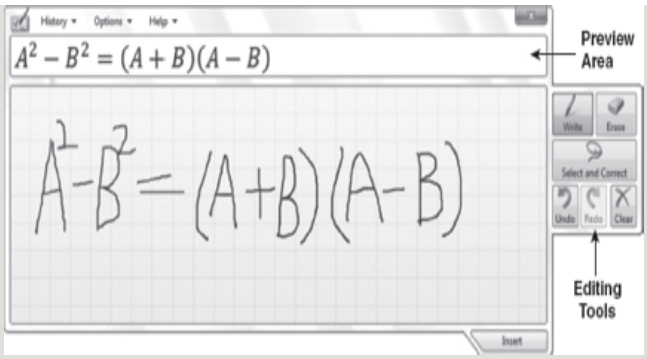


Figure 7.16 Markup Window of Math Input Panel

3. Click the **Insert** button to put the recognized math into your document.

Note: *Using the Math Input Panel, you can insert math into only those programs which support Mathematical Markup Language (MathML). Microsoft Office 2007 is an example of such program.*

Snipping Tool: *Snipping Tool* is used to capture a snapshot (called *snip*) of any object, window or a part of window on your desktop. Once you capture an image, you can annotate it with handwritten comments, save it as an HTML, PNG, GIF or JPEG file, or send it via e-mail from within the Snipping Tool program window.

Snipping Tool enables you to capture the following types of snips.

- **Free-form Snip:** Enables to capture a free-form-shaped snip by drawing irregular lines around the area you want to capture.
- **Rectangular Snip:** Enables to capture a rectangle-shaped snip by dragging the mouse pointer to form a rectangle around the area you want to capture.
- **Window Snip:** Enables to select a window like a dialog box or browser window that you want to capture.
- **Full-screen Snip:** Enables to capture the entire screen just like the **Print Screen** key on the keyboard.

To capture a rectangular snip, follow the steps given below:

1. Click **Snipping Tool** under the **Accessories** folder on the **Start** menu. The Snipping Tool window appears on the screen

(see Figure 7.17).

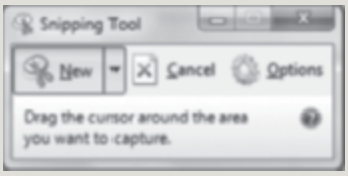


Figure 7.17 Snipping Tool

2. Click the down arrow beside the **New** button and select the **Rectangular Snip** from the list.
3. Click and drag the mouse pointer to select the area on the screen you want to capture. As you start selecting, the Snipping Tool window disappears and after you complete the selection and release the mouse button, the Snipping Tool markup window appears showing the captured image (Figure 7.18).

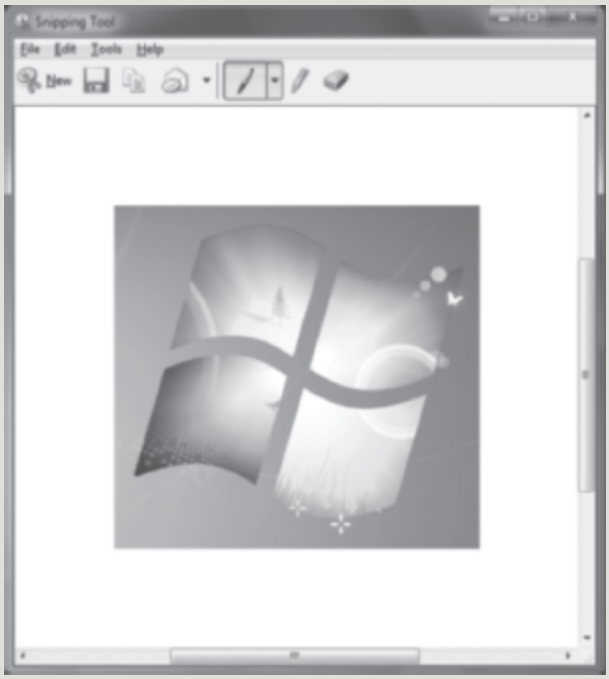


Figure 7.18 Snipping Tool Markup Window with Captured Image

4. Perform the desired changes to the captured image such as adding comments, highlighting some specific area, etc.
5. Click the **Save** button present on the toolbar to save the image.

LET US SUMMARIZE

1. The latest version of Windows operating system is Windows 7 that has been released by Microsoft in 2009 to be used on home and business desktops, laptops, Tablet PCs, netbooks and media centre PCs.
2. Microsoft has released Windows 7 in six different editions, namely, *Windows 7 Starter, Windows 7 Home Basic, Windows 7 Home Premium, Windows 7 Professional, Windows 7 Enterprise* and *Windows 7 Ultimate*.
3. In Windows 7, all the basic elements of Windows, like icons, menus, dialog window, etc., have been redesigned to provide better visuals. The new Aero interface takes advantage of the latest graphics hardware and provides smoother window handling, enhanced animations and glass visual effects (transparencies).
4. An *Aero theme* is the combination of one or more desktop backgrounds, window colour, sounds, and screen saver. Some Aero themes include desktop icons and mouse pointers also. You can apply any of the themes to change the desktop background, window colour, sound and screen saver simultaneously.
5. The *desktop slide* show feature of Windows 7 allows the desktop backgrounds to change automatically after some specified time interval.
6. *Gadgets* are the mini programs that provide quick access to information and frequently used tools. Windows 7 includes many gadgets and you can download more gadgets from Microsoft's website and install them on the machine.
7. Every time you open a program, file or folder, Windows creates a corresponding button on the taskbar. When you place the mouse pointer over a taskbar button, a small preview of the open window(s) associated with that button appears over the taskbar. This preview is called *thumbnail*.
8. Start menu is divided into three basic parts: *left pane, right pane* and *search box*. The left pane shows a short list of programs installed on the machine. The right pane provides quick access to the commonly used folders and files as well as the computer's settings. The search box lets you search for files, folders or whatever you want.
9. Windows 7 comprises several Aero effects that not only tend to improve the visual experience but also ease the task of managing different windows. These effects include *Aero Flip 3D, Aero Peek, Aero Shake* and *Aero Snap*.
10. *Aero Flip 3D* allows you to preview all the open files and folders simultaneously without clicking on the taskbar. Using this feature, you can preview all the open windows arranged in the form of a three-dimensional stack that you can easily flip through.
11. *Aero Peek* allows you to temporarily view the desktop and all its folders as well as gadgets and Sticky Notes placed on it by fading down all the open windows.
12. *Aero Shake* allows you to minimize all the open windows except the desired one. For this, click on the title bar of the window you want to keep open and drag the mouse back and forth quickly.
13. *Aero Snap* allows you to arrange two open windows side by side on the screen. This is helpful while comparing contents of two files or copying files from one folder to another.

14. *Jump list* is a new feature of Windows 7 that provides you quick access to the recently and frequently accessed items (such as files, folders or websites) of some specific program that is used to open those items. You can open any item appearing in the jump list by just clicking on it.
15. *Sticky Notes* are used to write down something like a to-do list, someone's phone number, a note to make you remind of anything, etc. They are placed on the desktop.
16. *Math Input Panel* allows you to directly write the simple or complex mathematical expressions just like on a notebook paper and the in-built math recognizer of Windows 7 recognizes your handwritten expression. You can then transfer the recognized math into your document.
17. *Snipping Tool* is used to capture a snapshot (called *snip*) of any object, window or a part of a window on your desktop. Once you capture an image, you can annotate it with handwritten comments, save it as an HTML, PNG, GIF or JPEG file, or send it via e-mail from within the Snipping Tool program window.

EXERCISES

Fill in the Blanks

1. ..... edition usually comes pre-installed on computers by the computer manufacturers or dealers.
2. The ..... interface of Windows 7 takes advantage of latest graphics hardware and provides smoother window handling, enhanced animations and glass visual effects (transparencies).
3. ..... are the mini programs that provide quick access to information and frequently used tools.
4. By default, the taskbar displays only three buttons corresponding to ....., Media Player, and ..... program.
5. ..... provides a common way to perform varied activities such as starting various programs, accessing commonly used folders, searching for files, folders and programs, etc.
6. ..... contains a group of icons that indicate the status of certain programs installed on the computer.
7. ..... feature of Windows 7 lets you temporarily view the desktop along with the folders, gadgets and Sticky Notes placed on it by fading all other open windows.
8. You can open the Start menu by clicking the ..... button on the taskbar or by pressing the ..... key on the keyboard.
9. You can open Microsoft Word program by typing ..... in the search box on the Start menu.
10. ..... provides you quick access to the recently and frequently accessed items (such as files, folders or websites) of some specific program that is used to open those items.

Multiple-choice Questions

1. Windows 7 was released in .....
    1. 2009
    2. 2008
    3. 2010
    4. None of these

2. Which of the following editions of Windows 7 is not available through retail market?

    1. Ultimate
    2. Home Professional
    3. Enterprise
    4. Starter

3. Which feature of Windows 7 allows the desktop backgrounds to change automatically after some specified time interval?

    1. Desktop slide show
    2. Gadget
    3. Aero theme
    4. Jump list

4. Which combination of keys on the keyboard arranges all the open windows in a 3D stack that you can easily flip through?

    1. Alt+Tab
    2. Windows logo key+Tab
    3. Ctrl+Tab
    4. Shift+Tab

5. Which feature allows you to arrange two open windows side by side on the screen?

    1. Aero Snap
    2. Aero Shake
    3. Aero Peek
    4. None of these

6. Using Snipping Tool, you can save your snip as .....

    1. PNG file
    2. GIF file
    3. JPEG file
    4. All of these

7. The ..... snip enables to select a window like dialog box or browser window that you want to capture.

    1. Rectangular
    2. Window
    3. Free-form
    4. Full-screen

8. How many editions of Windows 7 have been released by the Microsoft?

    1. Five
    2. Seven
    3. Six
    4. Four

9. Which of the following applications is no longer a part of Windows 7?

    1. Windows Media Player
    2. WinZip
    3. Windows Media Centre
    4. Windows Photo Gallery

10. An Aero theme is the combination of one or more desktop backgrounds, ..., ..., and screen savers.

    1. Window colour, Sounds
    2. Desktop icons, Gadgets
    3. Sounds, Gadgets
    4. Window colour, Gadgets

State True or False

1. You can make any open window active by clicking on its thumbnail.
2. Using the Aero Shake feature, dialog boxes cannot be minimized.
3. The taskbar in Windows 7 is smaller than it was in previous versions of Windows.
4. Aero Snap feature allows you to minimize all the open windows to let you view the desktop.
5. If you open multiple instances of the same program, Windows 7 creates a separate button for each instance on the taskbar.
6. The Math Input Panel basically has been designed to be used with a tablet pen on a tablet PC.
7. In Windows 7, the Show desktop button is placed at the left-most corner of the taskbar.
8. You can view the Jump list for a program on the taskbar by right-clicking on the program icon.
9. The taskbar button of the active program is always highlighted.
10. Windows 7 Home Basic edition includes features like touch-screen support, Aero effects, etc.

Descriptive Questions

1. Describe the various editions of Windows 7 released by Microsoft.
2. List the steps for creating a slide show of the pictures provided by Windows as a part of Aero themes.
3. Discuss some changes that you have noticed in the taskbar in Windows 7.
4. What is the purpose of the Aero Snap feature? How would you use it?
5. Mention the steps for using the Snipping Tool to capture a dialog box window.
6. Write short notes on:
    1. Gadgets
    2. Aero Shake
    3. Jump lists
    4. Thumbnails
7. What do you mean by Sticky Notes? Write down the steps for creating a Sticky Note.
8. What is Math Input Panel? How can we use it?

ANSWERS

Fill in the Blanks

1. Windows 7 Starter
2. Aero
3. Gadgets
4. Internet Explorer 8, Windows Explorer
5. Start menu
6. Notification area
7. Aero Peek
8. Start, Windows logo

9. Winword
10. Jump list

Multiple-choice Questions

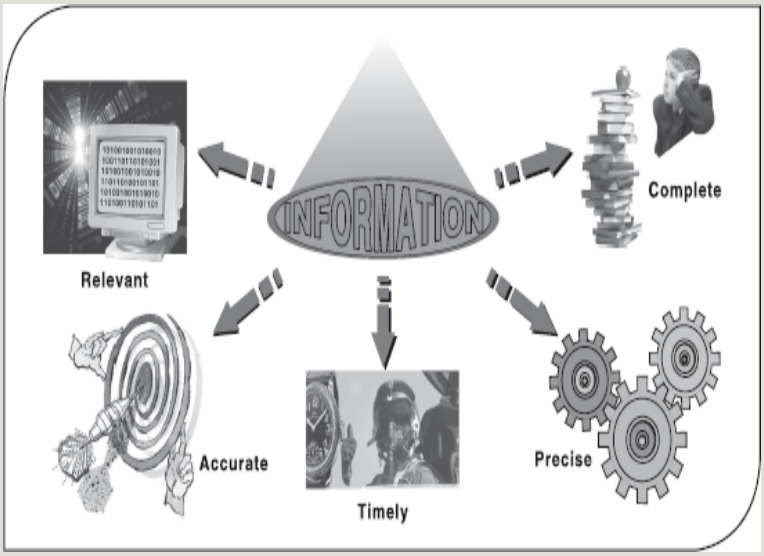
1. (a)
2. (c)
3. (a)
4. (b)
5. (a)
6. (d)
7. (b)
8. (c)
9. (d)
10. (a)

State True or False

1. True
2. True
3. False
4. False
5. False
6. True
7. False
8. True
9. True
10. False

Chapter 8

Information Technology Basics



This chapter introduces the reader to Information Technology (IT). Information technology is the combination of information and technology. IT comprises telecommunications and computers for obtaining, processing, storing and communicating information in various forms to different persons and places. This chapter begins with the discussion on information and technology and how these two unite to present information technology. The chapter describes the impact of IT in various areas such as business, education, defence, public sector, and media. The chapter concludes with a discussion on various job prospects in the IT sector, which is revolutionising the world.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The basics of Information Technology (IT)
- Information and its prerequisites and needs
- Technology and reasons for its popularity
- The meaning of IT and its components
- How IT fits into the present scenario
- The role of IT in various areas of our lives
- The relationship between IT and the Internet
- Careers available in the IT industry

8.1 INTRODUCTION

The world is changing at a very fast pace. The changes can be seen in every aspect of life, be it politics, society or economy. One of the most important factors of change is *technology*. The word "technology" originated from the Greek word *tecne* and *logia*, where "tecne" means skill and "logia" means study of science. Technology is the knowledge generated for the purpose of development of new systems to help in solving practical problems. Today, the impact of technology can be felt in all areas—be it home, office or market. Using technology, information can be easily gathered about various fields such as weather forecasting, space exploration and much more.

For the development of technology, information is essential. *Information* is the collection of facts gathered through various means of communication (for example, people, newspaper and television) and plays a vital role in drawing a conclusion. At the same time, technology

makes information gathering fast and easy. Information along with technology has created a new branch called *Information Technology* (*IT*). IT involves the processing of information by a computer. This is possible through the use of hardware, software, services and supporting infrastructure to manage and deliver information. IT has changed our daily lives radically over the recent years, for example, the use of mobile phones to make calls and send text messages, the use of websites to book flights and the use of ATM machines for banking, and so on.

8.2 INFORMATION

Today's information age began with the invention of the telegraph transmitter and receiver. It was the first instrument to transform information into electrical form and transmit it reliably over long distances. Information can be in any form, such as the answer to the question “why is the sky blue?”, the latest news updates, the top ten movies or the formula to solve a mathematical problem. Information can be considered as a critical tool for development. It is necessary to carry on research processes and enhance the lives of people in many ways.

8.2.1 Prerequisites of Information

Information is the processed data on which decisions are taken and subsequent actions are performed thereafter. For decisions to be meaningful and useful, the information must possess the following attributes (Figure 8.1).

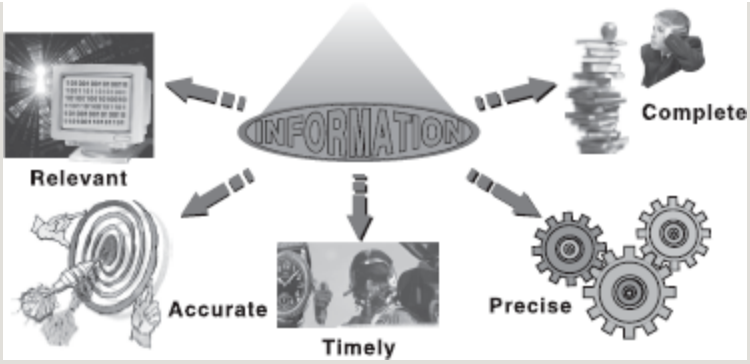


Figure 8.1 Prerequisites of Information

THINGS TO REMEMBER

Information

Information provides the power to find and evaluate problems and make decisions effectively and efficiently. Information can be identified regardless of the particulars of the physical realization. Information in one realization (such as ink on a sheet of paper) can be easily transferred to another form (such as spoken words). It also helps in determining how to use technology to one's advantage, learn the research process and enhance problem-solving skills. Information includes deterministic, probabilistic and quantum information.

- **Accurate:** To be useful, information must be accurate at all levels because all further developments are based on the available information. Accurate information provides a reliable and valid representation of raw facts. The cost of inaccurate or distorted information can be extremely high.
- **Timely:** Information is appreciated only if it is available on time. If the information is available ahead of time, its value may be diminished because the information might get archaic or the user may simply

forget it. Obviously, any availability after its due time simply has no significance.

- **Complete:** Complete information tends to be comprehensive in covering the issue or topic of interest. Without complete information, a decision maker may get a distorted view of reality. For example, incomplete market information can lead producers to introduce products and services not desired by the customers. As a result, the organization may incur huge losses.
- **Precise:** Apart from being complete, information must also be precise, containing all the essential elements of the relevant subject areas. In order to provide complete information, usually a large amount of data is gathered. Important information may be buried in the stacks of such data. One needs to work hard to get precise information out of it.
- **Relevant:** Information is relevant if it can be applied to a specific situation, problem or issue of interest. For example, operation managers need information on cost and productivity whereas marketing managers need information on sales projections and advertising rates. On the other hand, product inventory information is not relevant to a payroll manager.

8.2.2 Need for Information

Information is required for carrying out all the decision-making activities of an organization. The quantity of the information is important for the flow of activities in business. Some of the common needs for information are given below.

- **Information and Decision-Making:** Every job or task involves decision-making. Decision-making is the process of identifying, selecting and implementing the best possible alternative. The right information, in the right form, and at the right time, is essential to make correct decisions. For example, based on the information about customers, competitors and production capabilities, a manager may decide to inform top executives that a strategic decision needs to be made. The top executives would use this information to identify alternatives for consideration. Each alternative would then be evaluated based on feasibility, cost, time and other such criteria. Based on their assessment, the top executives would select the alternative that makes the best business sense and begin its implementation.
- **Information and Communication:** Information is vital for communication and a critical resource for performing work in organizations. Business managers spend most of their day in communicating with other managers, subordinates, customers, vendors and so on. Indeed, management in itself is an information process, which involves collecting, processing and distributing information. A manager must keep track of the information flow from the sources inside and outside the organization.
- **Information and Knowledge:** Information plays a vital role in the accumulation of knowledge. Within a corporation, the importance of efficient use of information (or knowledge), for normal functioning of

the corporation, has gained to the extent that it has become a major constituent for the growth and survival of a corporation in a competitive environment. The future is shaped by our actions today, and our actions today are based upon our knowledge. Therefore, for achieving higher levels of success, one must be well informed and should have clarity of information.

- **Information and Productivity:** Information helps in making sense of our environment, which assists in achieving the performance objectives. In fact, productivity is directly related to the availability, value and the application of information in the related context. Individuals and organizations need to ensure that their information-processing systems are properly integrated. They should also ensure that necessary information is being supplied to the right people at the right time in the organization. For example, it makes sense to share customer complaint data about a specific product with the members of the product development team responsible for redesign, thereby increasing efficiency and productivity.

8.3 TECHNOLOGY

Technology is the human innovation that involves the generation of knowledge and processes to develop systems capable of solving problems and extending human capabilities. Technology is concerned with the advancement of science, design, development, installation and implementation of information systems and applications. Technology may be referred to as a science used for practical purposes. It comprises specific information and knowledge that can be in physical form such as models, drawings, sketches, diagrams, blueprints, manuals, or in non-physical form such as training or technical services. It comprises mechanisms for distributing messages including postal systems, radio and television broadcasting companies, telephone, satellite, and computer networks.

Technology helps us to save time and energy by providing convenience at work. Technology is used in the production of goods and services that is considered useful in areas like computers, medicines and machines. Some examples of technology are the electronic media such as video, computers, compact disks, audiotape, cell phones and satellite equipment that are used as tools to create, learn, explain, document, analyse, communicate or present information. From household activities to

organizational activities, technology has been widely used as a part of our day-to-day lives. The reason that technology is becoming popular is due to the multi-faceted advantages that it provides. Some of the advantages of technology are discussed below.

- **Makes Work Easy:** Technology has provided many advanced appliances and electronic gadgets, which make work easier. For example, a washing machine is easier to use compared to manually washing clothes.
- **Fast:** Using technology, the results are obtained at a faster speed as compared to doing the same work manually. For example, washing machines take a few minutes to wash clothes; photographs are ready in just a minute, advanced technology in automobiles makes them more secure and faster, and cell phones help in making communication in seconds.
- **Helpful in Performing Creative Work:** Technology is useful in doing constructive work in various areas of life. For example, it is possible to edit an original picture for colour correction, text insertion or changing the background by using any graphic program. Similarly, technology is used to make a stereo picture based on scenes made within three-dimensional modelling programs like 3D MAX Studio.

THINGS TO REMEMBER

Computer Science

Some years back, a new science was born, which came to be known as “computer science”. Computer science is defined as a discipline that is concerned with methods and techniques relating to data processing performed by automatic means. It is important to understand that the development in IT is not restricted to computers only but also includes the fields like education, communication and science.

8.4 INFORMATION TECHNOLOGY

Information technology (IT) is a combination of telecommunications and computing to obtain, process,

store, transmit and output information in the form of voice, pictures or text. IT includes the following.

- Software applications and operating systems.
- Web-based information and applications such as distance learning.
- Telephones and means of telecommunications.
- Video equipment and multimedia products.
- World Wide Web.
- Electronic devices such as photocopiers.

IT exists from the days when telegraph was being used to move data from one place to another. With advancement in technology, the role of people responsible for handling data has also increased. A key to the technological advances in the IT sector is the interweaving of computation, communication and content supported by theories of information and computer science.

8.4.1 Components of Information Technology

The components of IT include hardware, software, data and people. These components are essential for the proper functioning of a system in IT that are discussed below and illustrated in Figure 8.2.

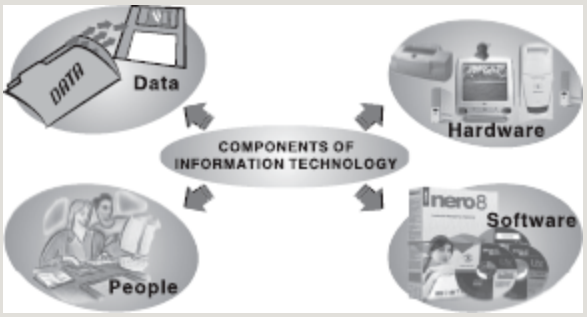


Figure 8.2 Components of IT

- **Hardware:** hardware consists of physical components that form a computer system or any other electronic gadget. To run a computer, both hardware and software are required. Hardware is the physical unit that stores and transmits information. Examples of hardware include monitor, keyboard, mouse, mobile phone and television, which are the popular inventions in the field of IT.
- **Software:** Software is a set of instructions in the form of programs, which control the sequence of operations (tasks). Software runs along with the hardware on the digital systems. For example, the software present in mobile phones allows users to perform various functions like sending messages, recording a person's voice and taking photographs.
- **Data:** Data are an unprocessed collection or representation of raw facts, concepts or instructions in a manner suitable for communication, interpretation and processing by humans or by automatic means. They include constants, variables, arrays and character strings. Data can be meaningful or meaningless. Data are gathered for the purpose of communication, interpretation or processing of tasks.
- **People:** People play an important role in IT. They perform various functions with the help of hardware and software to produce the desired output using IT.

8.5 PRESENT SCENARIO

Our society is being reshaped by rapid advances in IT, computers, telecommunications networks and other digital systems that have vastly increased our capacity to know, achieve and collaborate. These technologies allow people to transmit information quickly and widely, linking distant places and to create communities that just a decade ago were unimaginable.

It is difficult to appreciate how quickly IT is evolving. Five decades ago ENIAC, one of the earliest computers, stood 10 feet tall and stretched 80 feet wide; while today, one can buy a musical greeting card with a silicon chip that is 100 times faster than ENIAC. This extraordinary pace of IT evolution is bringing people and cultures together and creating new social dynamics in the process. It is leading to the formation of closely bonded, and widely dispersed communities of people united by their interest in doing business or in sharing experiences and intellectual pursuits. New forms of knowledge accumulation are developing, as are computer-based learning systems that open the way to innovative modes of instruction and learning.

The rapid evolution of digital technologies is creating not only new opportunities for the society but also challenges. Corporations and governments are reorganizing their work culture to enhance productivity, improve quality and control costs. Entire industries have been restructured to better align themselves with the realities of the digital age. It is no exaggeration to say that IT is fundamentally changing the relationship between people and knowledge.

8.6 ROLE OF INFORMATION TECHNOLOGY

IT plays a vital role in today's global economy. It has made a significant impact on research and development. Research is being carried out in networking, computing, data representation and many other areas for exploring the existing methodologies. Education is changing with the advent of new interactive online learning technologies and multimedia electronic libraries, which help in improving the sharing of knowledge and education practices. The major areas impacted by the advent of IT include business, media, education and so forth.

8.6.1 Information Technology in Business

IT plays an important role in business environment as every organization adopts it in some form or the other to perform tasks in an effective manner (see Figure 8.3). In the past few years, rapid development in IT, particularly in communications, electronic service networks and multimedia have opened up new opportunities for corporates. All these are contributing towards new and effective ways of processing business transactions, integrating business processes, transferring payments and delivering services electronically. *Electronic commerce* or *e-commerce* refers to the electronic means to conduct commerce between business communications and transactions over the Internet. It includes buying and selling over the Internet, electronic fund transfers,

smart cards, digital cash and all other ways of doing business over digital networks.

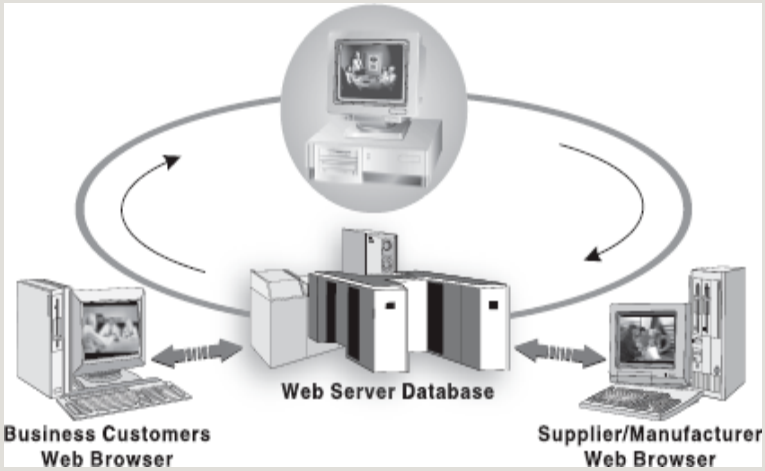


Figure 8.3 IT in Business

E-commerce websites are critical to industry for increasing e-commerce sales, improving productivity and lowering costs and user grievances. IT can provide the company a significant competitive advantage. The e-business approach consolidates a company's position by opening up new business opportunities and improving financial status. It has affected business in the following ways.

- **Competitiveness:** E-business offers a reliable and cost-effective means of doing business. Routine tasks can be automated. The customers can be provided support round the clock, that is, 24 hours a day. With advancement in the IT sector, corporates are spreading business around the world, thus increasing their presence and entering new markets.
- **Security:** With the growth of IT in business, there is more insecurity for data and important computer programs. Due to this factor, almost every organization has some security programs to avoid the illegal access of the company's information by unauthorized persons. The three fundamental attributes of a security program are confidentiality,

integrity and availability, which allow access to only authorized persons in an organization.

- **Cost Benefits:** The extensive availability of Internet-based information means that companies have a wider choice of suppliers, which leads to a more competitive pricing. Due to the presence of the Internet, the role of the middleman becomes less important as companies can sell their product or services directly to the customer.
- **Marketing:** Corporates engaged in e-business can take the help of their respective websites to create brand awareness of their products, thus creating new avenues of promotion of their products. In addition, companies' websites can also provide better services, such as after sales service to the customer.

8.6.2 Information Technology in Manufacturing

Corporates are developing a manufacturing society to provide manufacturers, distributors and researchers with structured methods and practices for implementing technologies in manufacturing environments (see Figure 8.4). Initial implementation of this collaborator-enhanced research in robotics are welding, which requires asynchronous and synchronous support and the use of diverse data formats (such as video, audio and image annotations).

Figure 8.4 IT in Manufacturing

Using IT, simulation systems have been developed, which enable manufacturers to virtually prepare prototype plant layouts, optimize raw material usage, and assess ergonomic factors prior to investing in plant redesigns or new factories. Corporates are investigating ways to augment commercial modelling and simulation software systems with programmable human modelling capabilities to harness manufacturing resources. Researchers are developing natural language interfaces that can help simulate human tasks in a manufacturing operations environment.

8.6.3 Information Technology in Mobile Computing

Mobile computing is a technology of wireless networking that uses common frequencies for transmission to permit networked devices move freely within the broadcast coverage area yet remain connected to the network. Mobile computing enables people to access network services anytime and anywhere and allows the transmission of data with the help of computers. This is done by integrating the cellular telephone technology with the portable computers having cabled network.

Mobile voice communication has been established throughout the world and has seen an increase in the number of subscribers to the various cellular networks. An extension of this technology is the ability to send and receive data across the cellular networks, which is the principle of mobile computing. Similarly, mobile data communication has become a rapid technology as it allows the users to transmit data from one remote location to another remote location. Thus, technology mobility has proved to be the solution to the biggest problem of business for the people on the move. The mobile data communications have a significant difference when compared with mobile voice

communications. In voice communication, mobile phones allow a user to move around and talk at the same time. One of the latest technologies used in mobile computing is *general packet radio service* (*GPRS*), which is a packet-based data transfer technology that provides increased capacity and higher data rates for access to the mobile Internet. It allows full mobility and wide area coverage as information is sent and received across a mobile network.

A cellular network consists of a number of mobile units that are linked together to switching equipment. The switching equipment interconnects the different parts of the network and allows access to the fixed *public switched telephone network* (*PSTN*). Every base station covers a given area. A number of adjacent cells grouped together form an area. Due to this reason, it is called *cellular communication*. The corresponding base stations communicate through a mobile switching centre (MSC), which is responsible for *routing* or *switching* calls from the originator to the destination (see Figure 8.5).

Figure 8.5 Mobile Computing

Mobile computing is used in many fields such as in emergency services, stock information control, credit card verification and e-mail. Mobile computing relieves the users from working at a desk and provides them access to information they need. In addition, it provides a better lifestyle through improved communication and personal data management.

8.6.4 Information Technology in Public Sector

In recognition of the serious staffing and equipment situation in the public sector, there is currently a great deal of activity in outsourcing of information services. In countries where public sector is still developing, and is under the control of the government, there is a great requirement to automate services such as booking of railway tickets and payment of electricity bills.

THINGS TO REMEMBER

Public Sector

As the public sector grows, one can expect to see a shift towards market level logistics systems, such as industry-wide distribution systems, some standardization of applications across industry, emphasis on global markets, outsourcing of more non-core processes such as finance and accounting, and home shopping. This can be possible if scarce resources are utilized in a proper and efficient manner.

The IT sector has become a boon for public sector companies by helping them increase their output and efficiency. In addition, IT has enabled the companies to have a greater quality control so that they are able to

meet the laid standards. This has been possible as IT has been able to do the following:

- Minimize duplication and sharing of scarce resources.
- Promote and ensure that IT solutions are cost-effective, efficient and business-driven.
- Promote transparency and accountability in public service by facilitating the availability, accessibility and sharing of information.

8.6.5 Information Technology in Defence Services

Today, military operators are restricted in their access to information. Most often, military operators use voice over radio or formatted text messages to communicate during crises. To help the military improve readiness and response, more flexible, timely and dependable access to information is needed (Figure 8.6). Key research areas in this field include dialogue management, context tracking, language generation, input language understanding, and hands-free and eyes-free interaction.

Figure 8.6 IT in Defence Services

New IT capability allows military personnel to literally converse with computers to create, access and manage information and solve problems. The presence of IT helps to gain relevant information in an organized and integrated manner that is readily usable by military personnel equipped with smart information devices. For example, pilotless warplanes have been developed to combat enemies at the time of war. This has been possible by gathering information and converting it into technology through means of research and development (R&D). Thus, the battlefield of the future will include weapons, unmanned combat vehicles and communication systems that can navigate, reconfigure and cooperate autonomously to accomplish time-critical military operations.

8.6.6 Information Technology in Media

With the advent of IT, two powerful forces have emerged to change the media. The first is the use of computers as a means of processing and analysing information. The second is the constantly accelerating capacity of technology to enhance communication in areas such as entertainment and social awakening. Some examples of IT used in media include audio-visual equipments like television, radio, VCR, video cameras and computers. Media can be broadly classified into two categories—*print media* and *electronic media*.

Print Media: Print media is the most widespread of all the media technologies. It remains the most widely adopted, portable, flexible and economical media. The most common example of print media is paper, such as newspaper and books (Figure 8.7). The reason for the popularity of the paper mode is that it can be read by anyone, anytime without any special knowledge or technology except the language in which it is printed. Due to this reason, the paper mode is still important. The presence of IT has greatly enhanced its (print media)

effectiveness making it an important source of information. For example, an event occurring in the remotest area of the world can be easily covered with the help of satellite links. The information provided by these links can then be converted into a readable form, enabling people around the world to know about the happenings.

Figure 8.7 Print Media

Electronic Media: Despite the growing popularity of the Internet, radio and television still represent the mode through which people around the world receive information about national and international events. In other words, when it comes to the delivery of information, electronic media such as television and radio are much ahead as compared to other types of media. This is mainly because of the speed with which information passes from one end to another. The improvement in technology and coming together of *information* and *technology* has greatly enhanced the prospects in this area. One can instantly view on

television any news that has occurred anywhere around the world (see Figure 8.8). For example, information on a natural calamity that has occurred thousands of miles away (for example, the Tsunami in December 2004) can be easily viewed.

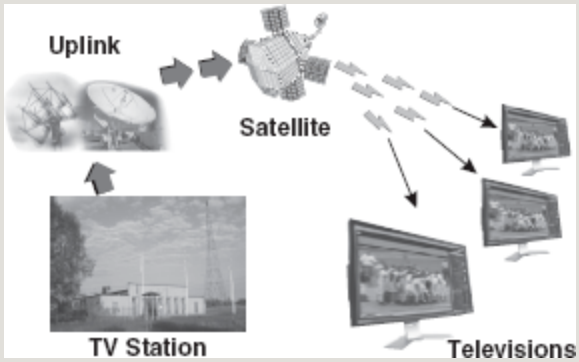


Figure 8.8 Information Broadcasting on Television

The process of receiving information and turning it into knowledge or entertainment is still at the development stage. However, multimedia applications (sounds, videos and images) are increasingly being used as part of the global revolution in electronic media. The Internet also forms a part of this media, which has been discussed in Section 8.7.

8.6.7 Information Technology in Education

The emerging technologies are constantly changing the world around us. These technologies have affected the way of learning and the way processes (like transactions) are carried out. The present education system is facing substantial pressure to prepare the students to learn, live and work in the digital age. Technology is becoming a powerful tool for communication, problem-solving, and

as a means of research for learning in the area of education. IT has helped learners to develop problem-solving, analytical and research skills.

Electronic Learning: *E-learning* or learning through computer technology is fast becoming an extremely versatile solution of providing learning with the tools necessary to address the needs of students in order to provide good and standardized education. Students can interact with the e-learning software, enjoy exciting colourful presentations, play games and answer quiz questions while learning a great deal through personal involvement.

In recent years, schools/universities and corporates have started to utilize the Web and local networks as a platform for distance e-learning. Tools like WebCT, Enspire, WebEx, HP Virtual Classroom or IBM Lotus Virtual Classroom, and many others allow students to understand, learn and interact effectively as the class experience is available online. Students can log in from home or a school computer, enjoy interactive presentations, watch a webcast (a video broadcast of the teacher over the network), or even participate in a full virtual classroom, with multiple students, live discussions, quizzes and voice/video interaction.

Students and teachers both gain considerable benefits by utilizing e-learning tools. E-learning is easily accessible round the clock, whether online or through special software. E-learning includes *computer-based training* (*CBT*) and *web-based training* (*WBT*).

- CBT is the training where a computer program provides motivation and feedback in place of a live instructor. It can be delivered through CD-ROM, LAN or the Internet.
- WBT is the training which delivers educational content through a Web browser over the public Internet or a private Intranet. Web-based training often provides links to other learning resources such as references, e-mail, bulletin boards and discussion groups. It may also

include a facilitator who can provide course guidelines, manage discussion boards, deliver lectures and so forth.

Students can work at their individual pace, as the option of recording and reviewing sessions exists with live, “virtual classroom” solutions.

Virtual Classrooms: This new technology helps captivate the student's interest. This concept interactivity involves a student with software that reacts and interacts, unlike a book. Thus, students can get immediate feedback on their actions, be it a quiz answer or a request for more information. With the help of virtual learning, students can do the following.

- Search databases for their views and of other students.
- Listen to pre-recorded audio comments to enhance knowledge.
- Converse about the subjects or related matters with the teacher and other students through e-mail and discussion groups.
- Use appropriate resources to seek, access and apply knowledge.

Through virtual classrooms, students can get online education from remote places (Figure 8.9). Virtual classrooms are beneficial to students who are unable to attend centralized classrooms due to some problems such as distant locations of institutes/universities.

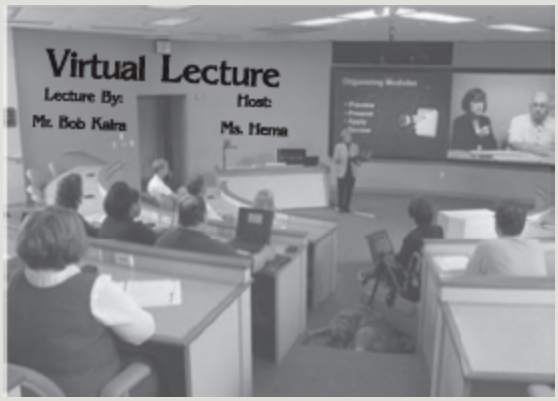


Figure 8.9 Virtual Classroom

8.6.8 Information Technology in Publication

In this ever-changing world, publishing houses, universities and e-book stores have started to use *information* written in the printed version of books on their computer screen through the use of *technology*. E-texts play an increasing role in education and research, as they offer unique benefits for both students and teachers. Websites like Project Gutenberg, NetLibrary and Bibliomania offer full texts of many classics. The files can be downloaded from these websites and can be read on the screen, printed and photocopied with minimal cost. In addition, portions of these can be copied and pasted into papers and used to prepare the examinations or home assignments for the students.

The digital revolution is gaining pace when it comes to books. Besides plain e-texts, a new format of digital books has evolved, usually referred to as *e-books*. These are special kinds of files, which can be viewed using specific software and offers more functionality for readers and researchers. Microsoft Reader, Adobe E-book Reader and MobiPocket (designed for PDAs) are some of the softwares available, which allow users to build their own libraries, read books comfortably on the screen and add side notes and bookmarks.

Many stores like Amazon, e-books and e-reader provide readers with affordable electronic books in one of the popular formats. They offer a wide selection of books to download for lesser price as compared to the printed editions. E-books can be installed on a library computer, and allow fast access and search to every student. For example, they can be installed on a teacher's laptop, and allow them to carry dozens of books everywhere including their own notes and bookmarks. Special e-book devices allow users to carry an entire library with the comfort and for the size of a single book.

The Internet's greatest strength is its ability to support simultaneous, interactive communications among different people. Unlike the telephone, which primarily supports one-to-one communication, or radio and television, where information flows in only one direction, from a single source to an audience who can only listen passively, the Internet allows information to flow back and forth among multiple users at the same time (Figure 8.10).

Figure 8.10 Internet

The Internet is the fastest growing information system for delivering a wide variety of things, such as products and services along with their links to other files over the Internet. Due to this, the Internet is also known as a *distributed system*. The information residing on different computers around the world can be easily linked together using hyperlinks, a method of organizing information on the system by linking one document to another. It provides a user access to a wide variety of services. These services include e-mail, newsgroup, file transfer, multimedia displays, real-time (live) broadcasting, shopping opportunities, access to the latest news and much more.

Electronic Mail: An important use of the Internet is sending and receiving electronic mail or e-mail. It allows users to send text messages across a network, which can either be an internal network (Intranet) or an external one like the Internet. The power of e-mail can be seen from the fact that traditional paper-based mailing list procedures are being replaced by mail-shots at targeted audiences, which can be transmitted instantly at little cost using electronic mail. Thus, information is sent across using this technology, which is considered fast and reliable (see "Internet Tools" Chapter 16).

Newsgroups: A *newsgroup* is a forum formed for the discussion on a particular subject in which the notes on a website are accessed and redistributed through a worldwide network of news discussion groups. They are organized into hierarchies of subjects, with the options in the newsgroup name indicating major subject categories and sub-categories. Some of the subject categories are news, society, science and computers. Users can create new newsgroups, post to existing newsgroups and respond to previous posts.

One of the Internet services is *electronic bulletin board* (*EBB*) that allows you to leave messages and access the required information through the software running on a computer. It uses a stand-alone computer connected to the Internet, which allows you to connect to the system, send and receive messages, files, and so on. An electronic bulletin board can contain directories of files and the users can read, download, upload and even modify the stored files depending upon their access privileges.

Webcast Services: The term Webcasting has been derived from the word "broadcasting". Webcasting refers to the use of the Web to deliver (or "broadcast") by streaming live or delayed versions of sound or video. Webcast has potential applications for academic courses,

presentations, using video on Web pages, on-demand programs, live events and training materials.

Formats of Real, QuickTime and Windows Media are commonly used to provide webcast services. Webcast services are being increasingly used by business houses to engage in teleconferencing with their branches or clients spread out in different areas, all over the world.

Intranet: The Intranet is a private computer network that is maintained by an organization for internal communication. It uses some of the technologies that the Internet uses, like the protocols, software and servers but cannot be viewed by an unauthorized person outside the organization (Figure 8.11). It uses different methods to ensure security over the network, such as access control list and so on. It is a cheap, fast and reliable networking system that links offices around the world, making it easy for the users working in a company to communicate with one another, and also to access information resources from the Internet.

Figure 8.11 Intranet

In an organization, printing and distribution of printed materials such as company letters, handbooks, manuals and training materials are expensive and time consuming. These traditional forms of printed material are replaced by the Intranet. Intranet can improve the quality of product, and provide a more efficient way to propagate information among employees at a minimal cost, time and effort.

Internet Chatting: *Chatting* refers to the live discussions conducted on the Internet, generally between two users with the help of computers, connected through a network, using the keyboard, webcam or microphone to communicate (Figure 8.12). Chatting is one of the most popular features of the Internet. This feature is offered by many online services or websites that allow

the users to chat by typing messages (text-based chat). These messages are displayed almost instantly on the screen of the recipient, as a scrolling dialog, who is using the chat room.

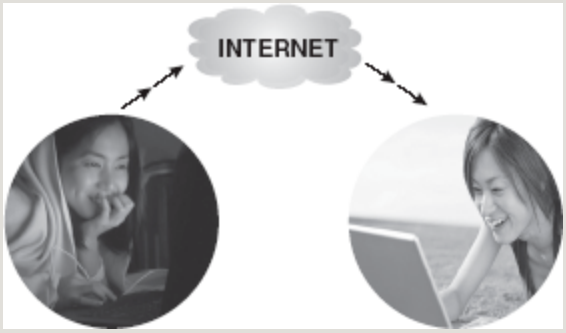


Figure 8.12 Internet Chatting

Internet chatting requires a *chat server* and a *client*. The chat server is usually provided by the Internet service provider (ISP) and the chat client is used by the persons chatting. An example of chat software is IRC that uses the Web browser (see "Internet Tools", Chapter 18).

8.8 CAREERS IN THE IT INDUSTRY

The fast spread of IT has generated a need for highly trained workers to design and develop new information systems and to integrate new technologies in them. It is creating the need for professional workers who can fulfil the high performance standards in the technology-based industries (Figure 8.13). IT professionals plan, design, develop, support and manage strategies and adopt them to fulfil basic objectives of the firm. The rapidly changing technology requires the employees to have more skills and education. The companies look out for IT professionals who not only have sound technical

knowledge of the subject but also have sound interpersonal and communication skills. Some of the career opportunities available in the IT industry are as follows.

Figure 8.13 Careers in IT

Programmer: A *programmer* is a person who writes the code for a computer program. Computer programs are the instructions given to the computer, telling the computer what to do, which information to identify, access and process. The programmers write programs, depending on the specifications given by the systems analysts and software engineers. After the completion of the design process, the programmer's job is to convert that design into a series of instructions that the computer can follow. These instructions are coded in a programming language such as C, C++ and Java, and then tested to ensure that the instructions are correct and produce the required result. Programmers perform functions such as system designing and its related services. They are involved in coding of a project and the successful implementation of the developed system.

There are two types of programmers—*application programmers* and *systems programmers*. The application programmers code programs for a specific task (application) in the organization. On the other hand, the systems programmers code programs for controlling and maintaining the computer system software.

Systems Analyst: A *systems analyst* plans, designs, develops and implements new systems or applies the existing system resources to perform additional operations. Most of the systems analysts work with specific type of systems like business, accounting, finance, science or engineering, and so on. Some systems analysts are also known as *system developers* or *system architects*.

A systems analyst determines the problem in the system by discussing it with the managers and users before starting an assignment. They study the information needs of the organization to determine the changes that are required, to deliver a solution to the problem of the users. To design a system, the systems analysts specify the inputs required by the system, plan the processing steps and finalize the output to meet the user's requirements. Next, the systems analyst determines the hardware and software needed to set up the system.

Systems analysts prepare the specifications, flowcharts and process diagrams for the computer programmers to follow, and then work with the programmers to debug and rectify any errors in the system. They also check the initial use of the system and organize tests to check the working of the system to ensure that it works as required, or planned. Systems analysts who perform more intricate testing of products are referred to as *software quality assurance analysts*. They not only perform tests but also diagnose the problems, suggest solutions and check whether the program requirements

have been met. The analysts who design, test and evaluate the Internet, Intranet, local area network (LAN), wide area network (WAN) and other data communication systems are called *network and data communication analysts*.

Network Administrator: A *network administrator* is a person who installs, configures and supports the local area network, wide area network, the Internet system or the segment of a network system in an organization. Every network in an organization requires at least one network administrator. Network administrators perform the functions such as installation, configuration, maintenance of computers and networking of equipment. They also check and ensure the availability of network to all the users. They supervise the support of other networks and plan, coordinate and implement network security measures.

Database Administrator: A *database administrator* (*DBA*) is an individual who is responsible for the storage of data, management of the database and implementation of the database management system. They work with database management systems software and determine ways to organize and store data. They identify user requirements, set up the computer databases, and test and organize the modifications to the systems. They also ensure the performance of the system in the organization, understand the platform on which the database is running and adds new users to the system. DBA is responsible for maintaining data integrity, back up systems and database security.

Web Designers: A *web designer* is an individual who provides an architectural structure for the effective communication of clients' content by logically organizing the information and making it accessible. With the rapid growth of the Internet, many jobs relating to design,

development and maintenance of the websites and servers have cropped up in the market. The web designer controls all aspects of design and programming of elements of text and visual design. They are responsible for maintaining all technical aspects of a website including its performance issues such as speed of access and approving the content of the website. Web designers create the artistic feel of the website, facilitate navigation and make content on the website accessible to the users.

Information Security Analyst: An *information security analyst* is an individual who performs highly complex analysis and technical tasks involving assignment and coordination of measures to provide information assurance, event detection and rapid response across various environments of the enterprise. Information security analysts design, implement and support the integration of information security solutions including security architectures, firewall administration, integrating security products, and developing and coordinating security implementation plans. They provide technical support on security projects that involve a wide range of issues including secure electronic data traffic, network security, platform and data security, and privacy. They review and contribute to the security activities of business applications and the development of project plans across all business units.

Business Process Outsourcing: *Business Process Outsourcing* (BPO) refers to the process of assigning some or all non-core processes to another party. The main motive of BPO is to allow the company to invest more time, money and human resources into core activities and building strategies, which helps in the company's growth. BPO jobs include various sectors like finance, accounts, customer support, IT services and medicines. Today, the global market is highly competitive and ever changing. A company must focus

on improving productivity and even cut down costs. Therefore, many tasks that use precious time, resources and energy, are being outsourced. BPOs, or the units to which work is being outsourced are flexible, quicker, cheaper and very efficient in performance.

LET US SUMMARIZE

1. *Information* is the collection of facts gathered through various means of communication (for example, people, newspaper and television) and plays a vital role in drawing a conclusion.
2. *Technology* is the knowledge generated for the purpose of development of new systems to help solve practical problems.
3. For information to be useful, it must be accurate, timely, complete, precise and relevant. Information is essential in our everyday lives for making decisions, communication, knowledge and productivity.
4. *Information technology* (IT) is a combination of telecommunications and computing to obtain, process, store, transmit and output information in the form of voice, pictures or text. The components of IT include hardware, software, data and people.
5. IT plays a vital role in today's global economy. The major areas impacted by the advent of IT include business, manufacturing, mobile computing, public sector, defence services, media, education and publication.
6. *Internet* is a massive network, which connects several computers globally forming a network in which any computer can communicate with the other.
7. *E-mail* allows users to send and receive messages anytime and anywhere in the world. It is popular as it is fast, flexible and reliable.
8. *Newsgroup* is a group formed for the discussion on a particular subject in which the notes on websites are accessed and redistributed through a worldwide network of news groups.
9. EBB is an Internet service that allows users to leave messages and access the required information through the software running on a computer.
10. *Webcasting* refers to the use of the Web to deliver (or "broadcast"), by streaming live or delayed, versions of sound or video.
11. *Intranet* is a private computer network maintained by an organization for internal communication.
12. *Chatting* refers to the live discussions conducted on the Internet, generally between two users with the help of computers, connected through a network, using the keyboard, webcam or microphone to communicate.
13. The fast spread of IT has generated versatile jobs such as programmer, systems analyst, network administrator, web designers, information security engineer and many more.

Fill in the Blanks

1. ..... is the collection of facts gathered through various means of communication.
2. ..... is the human innovation that involves the generation of knowledge and processes to develop systems capable of solving problems and extending human capabilities.
3. The components of IT include hardware, software, ... and .....
4. IT plays a vital role in areas such as ....., manufacturing, mobile computing, public sector, defence services, ..., ... and publication.
5. ..... refers to the use of the Web to deliver, by streaming live or delayed, versions of sound or video.
6. Through the Internet, information is sent in the form of files, audio, ... and .....
7. Media can be broadly classified into two categories—..... and .....
8. Mobile computing technology uses common ..... for transmission to permit networked devices move freely within the broadcast coverage area yet remain connected to the network.
9. E-learning includes ..... and .....
10. The IT industry has many career choices such as programmer, ..., network administrator, database administrator, ..., information security analyst and BPO jobs.

Multiple-choice Questions

1. Which of the following is an attribute of information?
    1. Timeliness
    2. Accuracy
    3. Relevancy
    4. All of these
2. Information is needed for .....
    1. Communication
    2. Decision-making
    3. Both (a) and (b)
    4. None of these
3. Which of the following tools is used to provide online learning experience to students?
    1. WebCT
    2. Enspire
    3. WebEx
    4. All of these
4. Virtual classroom is the online learning space meant for
    1. Listening to pre-recorded audio comments to enhance knowledge
    2. Learning from home using computers
    3. Providing teaching at home
    4. None of these
5. Which of the following option is included in E-learning?
    1. Computer-based training
    2. Web-based training

3. Both (a) and (b)
4. None of these

6. With the emerging role of IT, which of the following option has been affected?

    1. Cost Benefits
    2. Marketing
    3. Security
    4. All of these

7. Electronic Bulletin Board is an Internet service used for .....

    1. Accessing the Internet
    2. Leaving messages and accessing the required information through the software
    3. Collecting information using different software on the Internet
    4. Chatting on the Internet

8. One of the latest technologies used in mobile computing is

    1. PSTN
    2. Multimedia
    3. GPRS
    4. None of these

9. With the advancement of IT in business, it has affected

    1. The cost and competition
    2. Marketing and security
    3. Both (a) and (b)
    4. None of these

10. The forum formed for the discussion on a particular subject through the use of the Internet is known as .....

    1. News forum
    2. Newsgroup
    3. News guide
    4. News channel

State True or False

1. The conversion of printed books to e-books is an example of the role of IT.
2. If information is precise, then its time of delivery does not matter.
3. In this digital age, using the paper mode as print media is wastage of paper.
4. Intranet is a private computer network that is maintained by an organization for internal communication.
5. IT is concerned with the advancement of science and technology along with design, development and implementation of information system.
6. Electronic bulletin board allows users to leave and access the required information through the software running on the computer.
7. Computer-based training is delivered with the help of CD-ROM.
8. World Wide Web is not an example of print media.
9. Intranet based training is a kind of e-learning.
10. Due to advancement in IT, it has opened up many career options.

Descriptive Questions

1. What is information? What are its prerequisites?
2. Explain the need of information in our lives.
3. What is technology? Give some examples of areas where technology is being used.
4. Describe the essential components of IT.
5. Write short notes on
    1. IT in public sector
    2. Print media
    3. Electronic Learning
    4. Virtual classroom
    5. Intranet
    6. Internet chatting
6. Explain in your own words the advantages and impact of IT in our lives.

ANSWERS

Fill in the Blanks

1. Information
2. Technology
3. Data, People
4. Business, Media, Education
5. Webcasting
6. Video, Graphics
7. Print media, Electronic media
8. Frequencies
9. Computer-based training,
10. Systems analyst, Web designer

Multiple-choice Questions

1. (d)
2. (c)
3. (d)
4. (a)
5. (c)
6. (d)
7. (b)
8. (c)
9. (c)
10. (b)

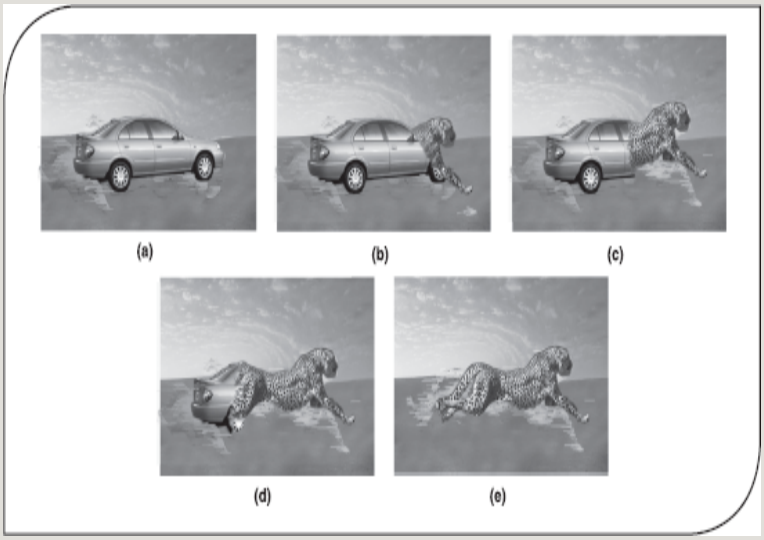
State True or False

1. True
2. False
3. False
4. True
5. True
6. True
7. True
8. True

9. False
10. True

Chapter 9

Multimedia Essentials



This chapter is written to give students a broad grounding on issues surrounding multimedia. Multimedia is the integration of multiple media such as graphics, text, sound, video, and animation elements, which together make an influence on the given information so that it can be presented in a pleasing and interactive manner. The chapter starts by discussing various components of multimedia—text, graphics, digital audio, video, and animation. The discussion is continued by providing an outlook on multimedia system and data compression. A detailed overview on multimedia applications and virtual reality completes the chapter.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The basics of multimedia and its influence on society
- The term “multimedia” and how it is used in our daily life
- An in-depth view on the building blocks of multimedia
- What a multimedia system is and what are its desirable features
- The various applications of multimedia in different walks of life
- Virtual reality (VR), its devices and characteristics

9.1 INTRODUCTION

The changing face of technology has taken humankind into an entirely new era, the “information age”, and the ability of state-of-the-art electronic tools to gather, refine and shift data have opened new vistas for data presentation. Presentation applications involving imaginative use of “multimedia” allow for new and more powerful and personalized modes of representing data. Multimedia is the use of a combination of text, graphic and audio, and aids in the presentation of information. Devices such as television, computer and video recorder are multimedia devices as they present information in combinations of text, graphics, animation, audio or video.

With the development of multimedia and Internet technology, people's outlook on information representation has changed completely. Growth in the Internet technology has emerged as a cost-effective means of communicating ideas across the world. Internet technology has helped multimedia in spreading its magic in almost every aspect of data communication. We feel and see multimedia in every aspect of our daily routine,

in cinema halls, discos, amusement parks and public kiosks. Nowadays, multimedia technology is extensively being used in schools, businesses, homes and public places to present information to people, employing both knowledge and entertainment together for interactive presentations. With the advancement in computer hardware and software, multimedia has the perfect launching pad to rule the field of information presentation and communication.

9.2 MULTIMEDIA: DEFINITION

The word *multimedia* is made up of two separate words, *multi* meaning many, and media meaning ways through which information may be transmitted. Therefore, *multimedia* can be described as an integration of multiple media elements together to influence the given information so that it can be presented in an attractive and interactive manner. In simple words, multimedia means being able to communicate in more than one way.

Traditionally, information was presented in a single medium only (unimedium presentation), be it newspaper, or radio, or television. With the development (and literally the most preferred means of information communication nowadays) of computers, presenting data has taken an entirely new meaning. Today, almost every computer can present data using a variety of mediums such as text, graphics, animation, audio and/or video. Hence, multimedia can be defined more appropriately as a technique that supports the physical and logical coexistence and interactive use of heterogeneous media classes, such as print, audio and video in specific application environments.

Newspapers were perhaps the first mass communication medium to employ multimedia— they used mostly text, graphics and images. Later radio was developed, which is now a major medium for audio broadcasting. Then the

television was invented that brought the video and changed the world of mass communication. After the invention of computers, multimedia is talked mainly in reference with computers. We can even say that multimedia is a computerized method of presenting information by combining audio and video components. For example, a multimedia database of musical instruments may allow a user not only to search and retrieve text about a particular instrument but also to see its pictures and play a piece of music. The advent of high-resolution screens and sound and video playback facilities for computers has resulted in the increased use of multimedia applications. A typical multimedia application is characterized by the following features.

- It possesses and makes use of different media
- It uses digital technology
- It demands interaction from the user
- It integrates sound, images, text and data into seamless applications.

9.3 BUILDING BLOCKS OF MULTIMEDIA

As discussed, multimedia is mainly concerned with the computer-controlled integration of text, graphics, drawings, still or moving images (video), animation and audio where every type of information can be represented, stored, transmitted and processed digitally (Figure 9.1). By definition, the media class print includes static media types such as text, graphics, facsimile and natural images; the media class audio includes dynamic media types such as natural and synthetic sound, speech and music; and finally, the media class video includes dynamic media types such as natural and synthetic image sequences.

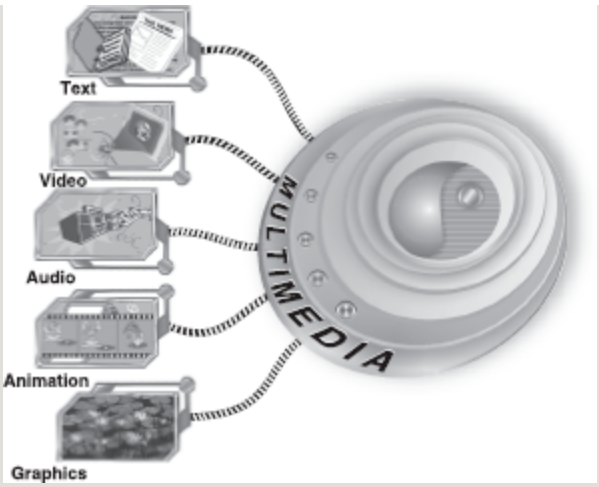


Figure 9.1 Building Blocks of Multimeida

In the case of computers, the most commonly used multimedia mediums are the following:

- **Text:** The on-screen display of words, that is, text, is the base of most applications. Different styles, fonts and colours can be used to emphasize specific points.
- **Graphics**: Seeing a picture of an object has more impact than merely reading about it. Graphics include conventional artwork, computer-generated artwork and photographs or captured video frames.
- **Audio**: This includes speech, audio effects, ambient sound (for example, the background sound of the sea) and music.
- **Video**: This includes displaying real-life events that are recorded using digital video recorder. For example, a movie.
- **Animation**: Animation is the timed sequence of a series of graphics or frames together to give the appearance of continuous movement.

9.3.1 Text

Humankind developed text and symbols so that they could communicate. Writing helped human to remember things for a much longer period than what is possible with the help of memory (brain). Soon the written messages became popular. Reading and writing skills are prevalent in most modern cultures. As has been the case throughout history, text still delivers information that is

concentrated and meaningful. Words and symbols in any form, spoken or written, are the most common means of communication.

The advancement in representing textual messages has completed a full circle. These days not only is the text capable of displaying information, but it has also gained interactivity; things like an electronic annotation, which assist in navigating through all the information presented on the Web by means of *hypertext*. A hypertext can be defined as an object in the text that allows the user to jump from one web page to another at a click of the mouse button (Figure 9.2). It acts like electronic footnotes and assists you in navigating through all the information.

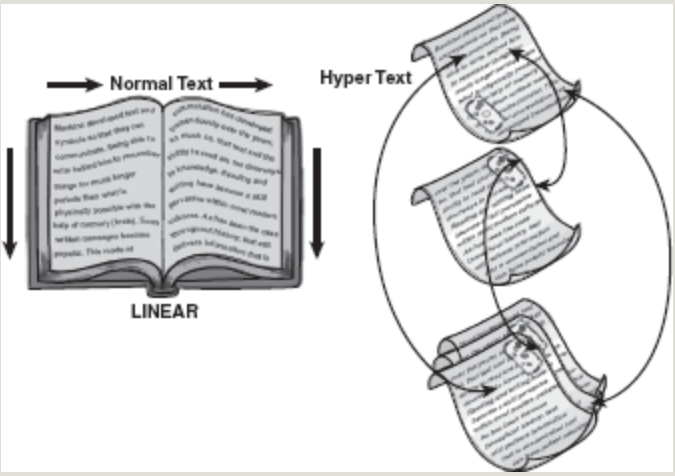


Figure 9.2 Normal Text and Hypertext

The amount of text to be included in an application, the type of font selection, the size, style and colour of the font make the application more meaningful. Text files are usually stored character by character. Files may

contain raw text or formatted text. For example, Hypertext Markup Language (HTML), a program language source (C, Pascal) or Rich Text Format (RTF)—a file format for text files that includes formatting like different font styles and typefaces and so on. The basic storage of text is one byte per character (text or format character). For other forms of data such as spreadsheet files, some may store format as text (with formatting), others may use binary encoding. Even though the storage requirement of this media is not high, when data is stored on disk, small files may take larger disk storage requirements due to block and sector sizes of disk partitions.

Text can be written using various text editing and processing tools. The most commonly used tool is a *Word Processing Software* such as Microsoft Word and Word Perfect where a text file can be developed and later imported into the multimedia authoring tools to integrate with other media elements (Figure 9.3). Typically, a word processor includes grammar and spelling checkers, template for letter writing, mail merging and many more features to enhance ordinary text into rich text.

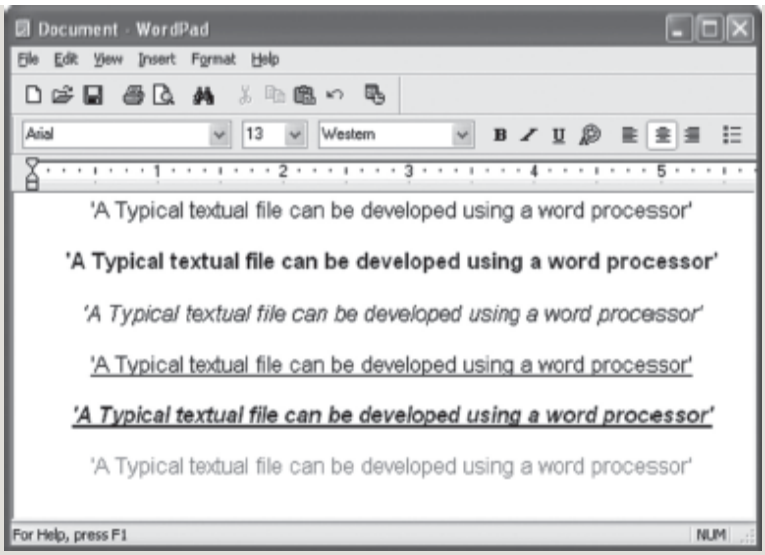


Figure 9.3 Word Processor

9.3.2 Graphics

Graphic is a generic term used to describe the pictorial representation of data. When we talk about graphics, we generally tend to refer those pictorial representations that are created by primitive objects, such as lines, polygons, circles, curves and arcs. For example, a picture of a cube can be drawn using lines and a pie chart can be drawn using lines and arcs. Images are still pictures that are represented as a bitmap (a grid of pixels [Figure 9.4]). Graphics are usually editable or revisable (unlike images). Generally, graphic input devices include keyboard, mouse, trackball or light pens. Graphic files usually store the primitive assembly and do not take up a very high overhead. Images may be generated by programs similar to graphics or animation programs. They can also be scanned using a digital scanner or from a digital camera. Images may be stored at one bit per pixel (black and white), eight bits per pixel (greyscale, colour map) or 24 bits per pixel (true colour). Typically, a 512 × 512 greyscale image takes up one-fourth of an MB, a 512 × 512 24-bit image takes three-fourth of an MB

with no compression. This overhead soon increases with image size, so compression is commonly applied.

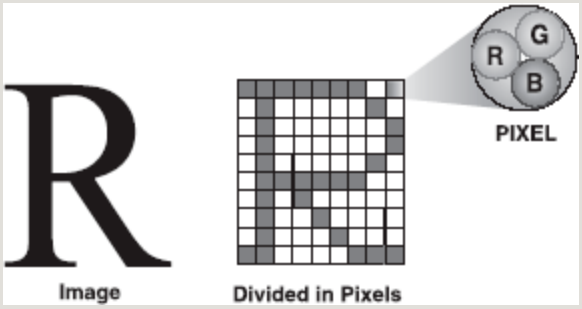


Figure 9.4 Composition of an Image

It is said that a picture is worth a thousand words; as a result, every multimedia presentation is predominantly based on graphics. By nature, human beings comprehend concepts more easily through graphics than textual information. Adding a graphic element into a text-based document can change the user's perception towards the document. Integration of the graphics element into the application makes it more elegant and understandable. Growing technology has produced a wealth of graphics and visual imagery for a multimedia project. For the development of a multimedia project, various images are captured from different sources like scanners, digital cameras and through the Internet. Once the images are transformed into a digitized form, they can be edited and modified with the help of image editing software like Adobe Photoshop, Corel Draw and Freehand (Figure 9.5).

Figure 9.5 Image Editing Software

In today's challenging world, representing thoughts effectively has become the basic need of the hour. No matter how good the multimedia presentation is, if the ideas of the presentation are not expressed in an effective manner then the presentation is a failure. One way of expressing imaginative ideas is by using graphics. Graphics can be presented in the following modes (Figure 9.6).

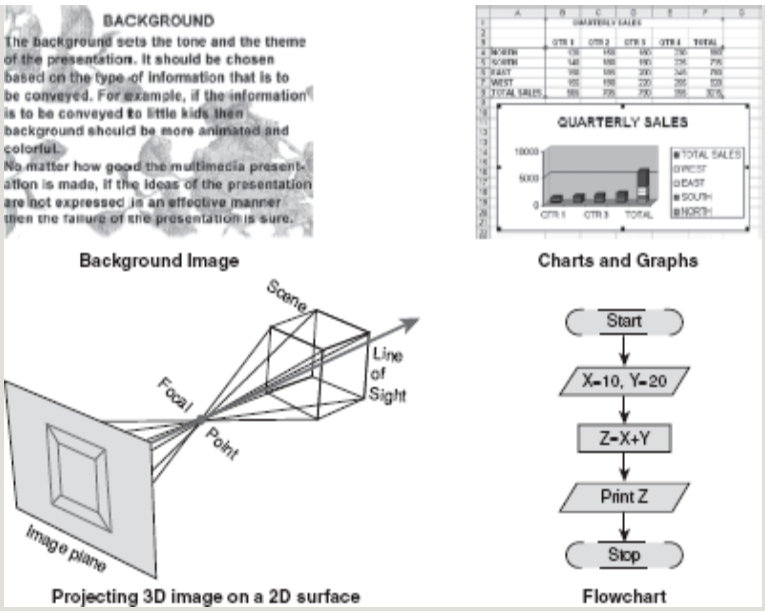


Figure 9.6 Types of Images

- **Background**: The background sets the tone and the theme of the presentation. It should be based on the type of information that is to be conveyed. For example, if the information is more text oriented, then simple background images should be used. In case the information is to be conveyed to little kids then the background should be more animated and colourful.
- **Two-dimensional (2D) Graphics**: Usually, when you look at a computer screen, you are looking at a two-dimensional image. The image spans from left to right and top to bottom but usually not in and out of the computer screen. A 2D image definition contains only information about how a scene appears from a particular point.
- **Three-dimensional (3D) Graphics:** The difference between 2D graphics and 3D graphics lies not in the way the images appear rather in the way they are defined. A 3D image contains information about the scene in three dimensions (x, y and z coordinates). However, this 3D data cannot be viewed directly. In order to view it, first the 3D data must be converted into a 2D image by the computer using a process known as *rendering*. The rendering process takes as input the 3D scene and the location and attributes of a camera and produces as output the 2D image that the camera sees from that particular point in space. The process is done by mathematically "projecting" a 3D scene onto a 2D surface. In photography, this is done by the camera lens, but in computer graphics, this is done by solving mathematical equations. In both cases, the result is a 2D image that has been generated from a 3D scene.
- **Charts and Graphs**: In business and training-related multimedia projects, charts and graphs are the most appropriate forms of graphics

used for presenting facts and figures related information (especially numerical).

- **Flowcharts**: The use of flowcharts expresses the systematic order of an event. They are developed mainly to break the complex problem-solving steps into smaller, less complicated steps so that developing logic becomes easier.

9.3.3 Audio

Audio means the sensation resulting from stimulation of the auditory mechanism by airwaves or vibrations. It is the vibration of air molecules in the atmosphere that can be sensed by the ear. It involves transmission of mechanical vibration by an elastic medium that is capable of being detected by the human ear. Today, a computer can play back most complex parts of a melody with great precision. Computers and digital technology have changed the face of the music industry. Today, digital music is used to produce special effects in televisions, films, in recording studios and in games. Music and sound effects not only add interactivity with the user but also play with the user's perception and emotions. For example, speech can be used to offer commentary or help without obscuring information on the screen. Audio can also be used to provide a sense of place or mood as done to perfection in the game. Mood-setting audio should employ quiet background sounds so that it does not compete with the main information for the user's attention.

Music is probably the most obvious form of sound. It can be used to set a mood, convey an idea and even tell a story. Music influences how we think and feel and how we perceive other forms of information. Speech is also a perfect way for teaching users the pronunciation of words as done by the speech learning software. Non-speech sound effects can be used in the user interface to inform users about background events. For example, the sound of the progress of a file download could be signalled by the sound of water pouring into a glass that

gradually fills up. These kinds of background sounds have to be non-intrusive. Sounds can be embedded into a web page and can be played. Common types of sound formats are MIDI and WAV. MIDI sound files are stored on a computer and contain musical notes, and sound information can be sent via a MIDI interface card to a musical instrument. MIDI files contain one or more MIDI streams with time information for each event. WAV sound files are the Microsoft standard file format for storing waveform audio data. These files are also known as *universal* files. A user can use these files through QuickTime and Windows Media Player. Once the audio files are recorded, they can be manipulated to produce the desired output using various sound-editing software like the Sound Forge (Figure 9.7).

Figure 9.7 Sound Editing Software

THINGS TO REMEMBER

Sound Editing Operations

The basic sound editing operations that a sound editing tool should possess are:

- **Trimming**: Removal of blank spaces from the front of the recording.
- **Splicing and Assembly**: Removal of extraneous noise that inevitably creeps into a recording.
- **Volume Adjustment:** To provide consistent volume level for different recordings in a single sound track.
- **Format Conversion**: To convert audio files into different formats.
- **Equalization**: To modify a recording frequency content to brighter sound.
- **Time Stretching**: To alter the length of sound file.
- **Reversing Sound**: To reverse all or small portions of a digital recording.

9.3.4 Video

The term video refers to the moving picture accompanied by sound such as a movie on television. It is a medium that is as direct as print and catches more attention. It adds richness to content and enhances the user's experience. It dramatizes and gives impact to the multimedia application. Integrating a video element in a multimedia application provides a great deal of information in a short time. If used in a planned and well-executed manner, it can make dramatic differences in the multimedia presentation. Digital video is useful in multimedia applications for showing real life objects, such as people talking or real life illustrations of concepts. Video places the highest performance demand

on the computer memory and on the bandwidth if placed on the Internet. Under these constraints, video has to serve as a supplement to text and images more often. Currently, video is good for the following:

- Promoting television shows, films or other non-computer media that are used as trailers in advertising.
- Showing things that move, such as a clip from a ballet and product demo of physical products.

Note: *Bandwidth is defined as the amount of data that can be transmitted from one computer to another in a given time frame.*

There are three main types of video files that are used: QuickTime, AVI, and MPEG (Figure 9.8). QuickTime and MPEG are the most commonly used file formats. AVI is a Windows-oriented video format that is not used because of problems like synchronizing both audio and video. Since QuickTime is readily available for Windows as well as Macintosh, the need for AVI is rapidly diminishing, at least on the Internet. MPEG's is the extremely high output quality. It delivers higher image resolution, picture quality, interlaced video format, multi-resolution scalability, and multi-channel audio features. It was developed as an international standard for use in CD-ROM, video games and other media that require high quality digital video. For the trade-off of using slightly larger files, one gets much higher-quality video, with up to 30 frames per second (the same as standard American television).

Figure 9.8 QuickTime Video Player

In terms of computers and Internet, the major problems with most videos are that their production quality is too low and they have a very large file size. To make video usable, video files need to be compressed in different file formats. One can even create one's own video files using video editing software such as Adobe Premier and Microsoft Moviemaker. Although it is time-consuming, the process of digitizing, editing and uploading video files is not an extremely complicated process. The only thing that should be of concern is the bandwidth that is being used. At 28.8 Kbps, a 1 MB file representing a few seconds of video will take about 10 minutes to download.

9.3.5 Animation

Animation is the process of sequencing still images in rapid succession to give the effect of live motion. The sequence of images is shown one after the other very quickly so that the human eye and brain blend the images to produce the effect of motion. It is a powerful

tool capable of communicating complex ideas. It is often much easier to show somebody how things work than to try to explain it. Animation is viable because of a natural phenomenon known as *persistence of vision*. An object seen by the human eye remains mapped on the eye's retina for a short time interval after viewing. This makes it possible for a series of images that are changed very slightly and very rapidly, one after the other, to seemingly merge into a visual illusion of movement. In other words, if we just slightly change the location or shape of an object rapidly enough, the eye will perceive the changes as motion (Figure 9.9). An animated sequence has to show a certain number of frames every second. For professional animation, one needs to have at least 30 frames per second.

Figure 9.9 Animation Effect

Animation can be useful because of the following:

- **Continuity in Transitions**: When something has two or more states, then changes between states will be much easier for users to understand if the transitions are animated instead of being instantaneous.
- Dimensionality in Transitions: Opposite animated transitions can be used to indicate movement back and forth along some navigational

dimension. For example, following a hypertext link to a footnote might be shown by a “down” animation.

- **Illustrating Change Over Time**: Since an animation is a time-varying display, it provides a one-to-one mapping to phenomena that changes over time. For example, population growth can be illustrated by showing a map with an animation of population changing over time.
- **Multiplexing the Display**: Animation can be used to show multiple information objects in the same space. For example, whenever the user moves the cursor over the various hyperlinks, an explanation for each is popped-up in the form of tool tip.
- **Visualizing 3D Structures**: Since the computer screen is two-dimensional, users can never get a full understanding of a 3D structure by a single illustration, no matter how well designed it is. Animation can be used to emphasize the three-dimensional nature of objects and makes it easier for users to visualize their spatial structure.
- **Attracting Attention**: Animation has the ability to control a user's visual awareness and this advantage can be moulded in the interface. If the goal is to draw the user's attention to a single element out of several or to alert the user to updated information then an animated headline will do the trick.

Animations can be done in two ways: *2D animation* and *3D animation*. In 2D animation, flat images are drawn one frame after another. In 3D animation, a mathematical model of a 3D object is created to portray its width and height as well as its depth. 2D animations are more popular; most cartoons that you watch on television are based on 2D animation technique. 3D animation is mostly used in computer-aided design (CAD). The two most commonly used animation effects are kinematics and morphing. *Kinematics* is the study of movement and motion of structures that have joints, such as a walking person. This is a tricky and carefully executed task, as one requires calculating position, rotation, velocity and acceleration of all the joints. *Morphing* involves transformation of one image into another (Figure 9.10). Morphing and other modelling applications are performed on still as well as in moving images.

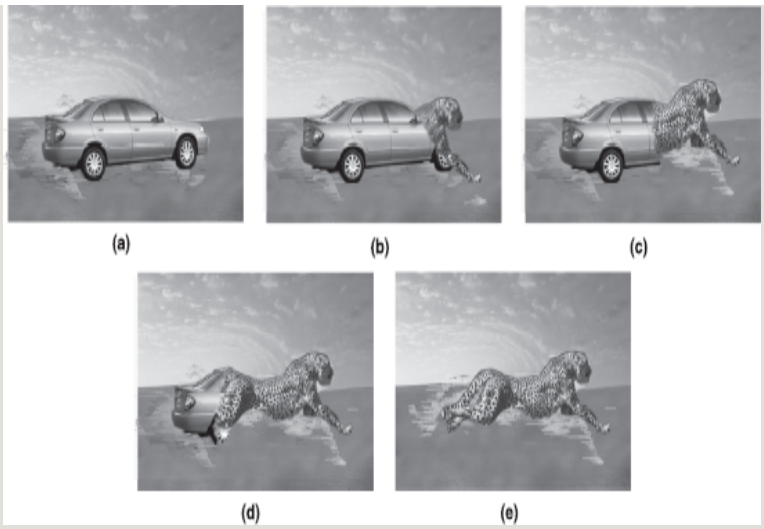


Figure 9.10 Morphing

9.4 MULTIMEDIA SYSTEM

A multimedia system is a system capable of processing multimedia data and applications. It is characterized by the processing, storage, generation, manipulation and rendition of multimedia information. Computer hardware and software play a vital role in supporting multimedia-based applications. These applications consume a large amount of computer resources like storage space and I/O devices. Permanent storage is required to keep digitized information for future retrieval and use. Many multimedia products require storage on a CD-ROM. Multimedia also requires a large amount of a computer's main memory and a large amount of bandwidth when distributed on a network. Optical disk emerged recently as the most cost-effective means for distribution of a multimedia project. A good quality multimedia project needs a high capacity of RAM that would enable smooth functioning of the project. It should support adequate storage of data that contains text, graphics, audio, video and animations. As vast amount of data is required for the processing of

multimedia elements, a faster processor speed is also essential along with the supporting RAM.

9.4.1 Desirable Features of Multimedia System

The minimum recommended hardware configuration needed by a multimedia personal computer to support a good multimedia application is as follows:

- **High Processing Power**: A multimedia system requires high processing power to deal with large data processing and real time delivery of media. A multimedia system should have a processor having a minimum of 500 MHz (or higher) of processing speed.
- **Storage and Memory**: Typically, all multimedia presentations are memory hungry, hence it requires large storage units (2–20 GB) and large memory (32–128 MB or more). For efficient management, a large cache memory, of Level 2 and 3 hierarchy, is also required.
- **Operating System**: Every multimedia system requires a special operating system that allows access to the file system and process data efficiently and quickly. It is also needed to support direct transfers to disk, real-time scheduling, fast interrupt processing, I/O streaming and so on.
- **Software Tools**: User friendly software tools are needed to handle media and design, develop applications, and to deliver the media. In addition, data representations/ file formats should be easy to handle and should also allow for compression/ decompression in real time.
- **Sound and Display Cards**: Since audio and video are basic components of any multimedia presentation, a good sound and colour display card (at least 8-bit) is necessary for any multimedia system.
- **Efficient Input/Output Devices**: Efficient input/output devices are needed to allow real-time recording as well as playback of data. Typically, a multimedia should be equipped with a high-resolution colour monitor, a CD-ROM drive (24× or higher), keyboard, three-button scroll mouse, a colour printer (300 dpi or more), scanner (300–500 dpi or more), video camera, audio and video recorders, a set of CD-quality speakers, and an audio microphone (see Figure 9.11).

Figure 9.11 Typical Components of Multimedia Computer

- **Network Support**: In a business conference, where a multimedia presentation is on, we may require associated peripherals connected to a main computer via communication devices (modems, hubs and so on) to carry out the supportive role in running an application.

9.4.2 Data Compression

In terms of computers, the only limitation of using a multimedia presentation is that it consumes a lot of storage space. In order to reduce the storage space, the multimedia components must be “compressed”. *Compressing* a file refers to the process of cutting down the size of the file by using special compression algorithms. There are two types of compression techniques: lossy and lossless. In the *lossy* compression technique, some data are deliberately discarded in order to achieve massive reductions in the size of the compressed file. In this compression format, we cannot recover all of its original data from the compressed version. JPEG image files and MPEG video files are

examples of lossy compressed files. On the other hand, in the *lossless* compression technique, the size of the file is reduced without permanently discarding any information of the original data. If an image that has undergone lossless compression is decompressed, the original data can be reconstructed exactly, bit-for-bit, that is, it will be identical to the digital image before compression. PNG image file formats use lossless compression. Similarly, text can be "losslessly" compressed using the ZIP tool.

Let us take some examples to illustrate both the compression techniques. Suppose you have two cups, one of 200 ml capacity (filled with water) and the other of 100 ml capacity (empty). Now pour water into the 100 ml cup and throw away the rest. This is lossy compression because you have kept only the necessary data and thrown away unimportant data. Now if you want to fill the 200 ml cup again with water from the 100 ml cup (recreating the original image), then you will find that the cup will not be filled because you have thrown away 100 ml of unwanted water (original image quality will not be retrieved). Lossless compression, on the other hand, does not use such technique. An analogy of this technique may be said to be the folding of paper. When the paper is folded, its size is reduced. However, you can get the original size of the paper by unfolding the paper, hence nothing is lost. Similarly, when an image is compressed losslessly, the quality of the image is retained even though the size is reduced.

Some Data Compression Techniques: There are many known lossy and lossless compression techniques. Some of them are as follows:

- **Predictive Encoding (Lossy):** In digital audio or video, successive samples or frames are usually similar to each other. Considering this fact in this technique, the initial frame and the difference values in the successive frames for all the frames are stored in the compressed form. As the size of the difference values between two frames is much smaller than the size of the frame itself, this encoding technique saves much

space. While decompressing, the previous frame and the difference value are used to reproduce the next frame.

- **Run-length Coding (Lossless):** In this technique, each repeated symbol in the given text is replaced with the symbol itself and the number of times it is repeated. For example, the text "ccccccbbbbuffffff" is compressed as "c7b4u1f6". This technique is effective only when symbols in the text are being repeated many times.
- **Variable-length Coding (Lossless):** Generally, the frequency of occurrence of characters varies in the text. For example, the letter e of the English alphabet is used much frequently than z in a regular text. Exploiting this fact, we may use variable number of bits per character than a fixed number of bits per character that most coding schemes (like ASCII, BCD, etc.) use. Fewer bits are used to represent the more frequently used character; the most frequently used character can be represented by one bit only. This helps in reducing the length of the compressed data.

To understand this technique, consider a language comprising three letters: *a, e* and *z* with the probability of occurrence of 0.42, 0.54 and 0.04, respectively. Since e has the highest of occurrence, we use one bit to represent it. The bits that can be used for *a, e* and *z* are 10, 0 and 110, respectively. It means each occurrence of a is replaced with 10, each occurrence of e with 0, and so on. For example, the text "aeezea" is compressed to 1000110010 requiring 10 bits only. Note that with a fixed number of bits per character, we require 2 bits for each of the three characters (as $3 < 2^2$). Hence, a total of 12 bits would be required to represent "aeezea". The variable-length coding technique is effective when the frequency of occurrence of characters is known and uneven.

9.5 MULTIMEDIA APPLICATIONS

The growing technology of computer's hardware and software, to hold several types of media, had made multimedia suitable for a wide range of applications and users. Here are some of the possible areas where the use of multimedia applications has replaced the conventional system with new technology.

9.5.1 Multimedia and Education

The goal of an instructor is to facilitate learning to help students gain a body of knowledge, acquire specific skills, and function successfully in society. One of the greatest challenges for instructors is the diversity of students, especially in the different ways they learn. Some students learn better through association, others by experimentation, some learn faster through visually oriented methods. Multimedia has the ability to accommodate different learning styles and can present material in a nonlinear manner. It is motivating, highly interactive and can provide feedback and evaluate skills. In the coming years, further advancement Figure 9.12 Using Multimedia in Education and acceptance of multimedia will change the face of teaching methodology. In some instances, teachers may find this as a new helping tool through which they can enhance their own teaching approach. They have found these tools appealing because of their ease of use and effective way of expressing information. Students prefer this new technology to the traditional text-based approach because of the multi-sensory experience they undergo, making them feel like having a game-like interface (Figure 9.12).

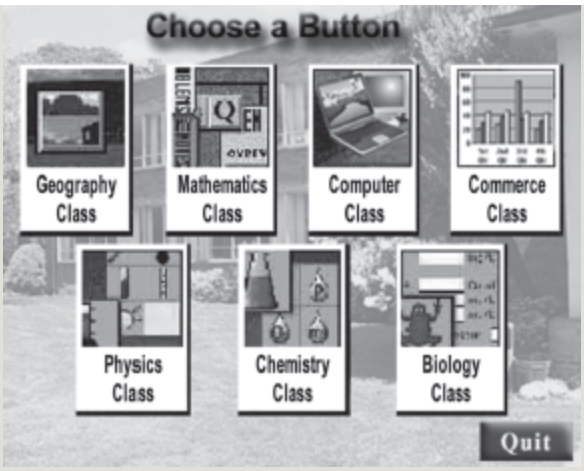


Figure 9.12 Using Multimedia in Education

Students are found to be more motivated when using multimedia. Even teachers have stated that multimedia enables students to learn in an efficient and effective manner. With multimedia coming into picture, the teacher has become a facilitator, problem solver and guide as opposed to playing a central role. Interactive multimedia is a powerful tool for education that can be used in many different ways, such as support to class presentations, tutorial for off-class consultancies or complementary activities, virtual laboratories, simulators, centre of activities to be used in a laboratory or at home, distance or local guided study, case studies and virtual conferences.

9.5.2 Multimedia and Entertainment

Drawing the line between education and entertainment in multimedia is almost impossible. Hence, a new term has been coined, “edutainment”, which means using multimedia to make learning entertaining. However, multimedia also has a purely entertainment side. Anything that is possible in sound and images is possible

on a multimedia CD. Multimedia means interaction, and to many interactive entertainment means games. In fact, game developers were the pioneers in the use of multimedia and they still provide the most innovative and interactive applications of multimedia. A game is one of those multimedia domains where all the elements of multimedia play an equally important role in making it realistic and interactive. In order to attract, engage, captivate and challenge the user, multimedia provides fast action, vivid colours, 3D animations and elaborate sound effects that are essential to entertainment (Figure 9.13). Today, many games have moved from the physical (hand/eye coordination) to the mental (solving the mystery or outwitting the opponent). On the other hand, hobbies and sports are examples of multimedia titles that provide the user with a vicarious experience such as being able to play the best golf courses in the world or simulate flying over 3D cityscapes.

Figure 9.13 Using Multimeida in Games

Nowadays, multimedia is extensively being used in the field of television broadcasting and movies. For example, while watching a live match on television, we are habitual of seeing replays, slow motion, charts analysis and other statistical information. These kinds of integrations on the sports make a novice understand and envisage the sports in a better way. In movies, you might have seen the “Titanic sinking” or “dinosaurs roaming around”. These are just some of the amazing effects that multimedia has offered.

9.5.3 Multimedia and Training

In a corporate world, many multinational organizations invest huge sums of money in providing training to their staff members and employees. Consider a case where a traditional method of training is being implemented. A reputed organization sends a group of trainers with necessary training tools and equipments to different centres across the world. This training team charges expenses related to travel, stay at the hotel, food, salaries and so on. Now consider the alternate case where a multimedia-based training is implemented. All the necessary data and teaching methodology are packed in a small storage media, that is, compact disk. This method of training can be easily provided in different parts of the world for a long time. Multimedia technology serves as a reference tool for the corporate executives. An employee can perform, self test and practice to improve his efficiency on the job.

Every company has a need to train its employees on a wide range of subjects from personnel policy to equipment maintenance. A number of companies are now training employees using multimedia-enhanced training materials. This can be done using off-the-shelf multimedia titles, but many companies are producing

their own in-house multimedia training products. For instance, the Boeing company has an entire training division dedicated to developing multimedia titles that instruct mechanics and pilots on new aircraft systems. Using multimedia tools, the trainee can perform a simulated job function in order to develop an advanced level without having touched the actual unit. The integration of audio and video permits this training technology to be a highly effective medium in areas such as flight and driving simulators (Figure 9.14). Similarly, NASA uses multimedia extensively for flight control training for astronauts.

Figure 9.14 Using Mulitimedia in Training

9.5.4 Multimedia and Business

Some of the biggest users of multimedia are business organizations. As businesses have to communicate with the outside world, multimedia offers a wide variety of options for business presentations, marketing and sales (Figure 9.15). Multimedia can be used at trade shows or to produce electronic catalogues. The marketing of new products can be greatly enhanced by using multimedia; these products can be marketed in a manner that

provides more detailed and stimulating information than print media. Auto manufacturers like Nissan provide interactive test drives of new vehicles on the Web or on CD. Macromedia now present their annual report, vision statement and product overview in print, on the Internet and in the CD-ROM form.

Figure 9.15 Using Multimedia in Business Presentation

Use of the multimedia application in the business world has created a boom among the business industry around the world. One can run his/her business in a much organized and professional manner. Using multimedia application as a tool for conducting presentations, training sessions, marketing, advertising and communications via videophone in running a business has made awesome changes in the style of carrying out business. As compared to a conventional presentation, a multimedia presentation can make the audience come alive with the presence of ambient sound and visual effects. Its widespread use in training programs has helped people in carrying out their tasks in a much better way. For example, an automobile mechanic can learn to repair the engine of a car; security personnel can learn to tackle a bank robbery through simulation. Keeping

records for the entire staff working in a big organization is a very cumbersome and challenging work. Integrating multimedia with the database management system, one can formulate the job in a much easier manner. Capturing images from the camera, storing voices along with the text data can make a secure database for the organization.

Multimedia technology marked its beginning with application in the advertising industry. Its power to boost communication at an affordable cost opened the way for the marketing and advertising personnel. Almost all advertisement and marketing agencies are well equipped with authoring tools and workstations. Presentations that incorporate flying banners, video transitions, animations and sound effects are some of the elements used in composing a multimedia-based advertisement to appeal to the consumer in a way never used before and promote the sale of the products. In addition, multimedia has incalculably increased the creativity of the marketing institutions. Today's computer has replaced high-priced sound and video editing facilities with reasonably priced equipment. The same computer is being used for a number of elements: sound, video, animation and so on.

9.5.5 Multimedia Interactive Kiosks

A *multimedia kiosk* is a kind of an automated information centre where one can get necessary information related to a particular context (Figure 9.16). Some of the common services provided by the kiosk are searching for necessary information, performing financial transactions, taking online shopping orders and many more. These kiosks are located near public places, such as hotels, railway stations, museums, historical monuments, bus stops and grocery stores, round the clock for the convenience of people, thereby eliminating the need for conventional-based information booths and

staff members. In such kiosks, printers are often attached so that a user can walk out with a printed copy of the information. For example, a kiosk at a bus stop can be useful in providing a list of all the bus numbers plying on the respective routes, the arrival and departure time of the buses, fare for the different destinations and so on.

Figure 9.16 Multimedia Kiosk

THINGS TO REMEMBER

Virtual Reality

Virtual reality, also referred to as *artificial reality,* or *synthetic reality,* is an interactive technology that totally controls sensory input and creates the convincing illusion that one is completely immersed

in a computer-generated world. To “enter” a virtual reality, a user wears special gloves, earphones and goggles, all of which receive their input from the computer system. In this way, at least three of the five senses are controlled by the computer. The goggles, for example, track how eyes move and respond accordingly by sending new video input.

Virtual reality can be described as an artificial environment that is created by using special computer hardware and software, which is presented to the user in such a manner that it appears and feels like a real three-dimensional world. The original term for virtual reality was “immersive virtual reality”, which means the user becomes fully immersed in an artificial, three-dimensional world that is completely generated by a computer. To enter in virtual reality, a user wears special gloves, earphones and head-mounted display device (HMD), all of which receive input from the computer system (Figure 9.17). In this way, at least three of the five human senses are controlled by the computer. In addition to feeding sensory input to the user, the devices also monitor the user's actions. The HMD, for example, tracks how the eyes move and respond accordingly by sending new video input. Virtual reality is used in health care and mass entertainment among other things. However, even today, virtual reality systems require very expensive hardware and software, and they are confined mainly to research laboratories. Virtual reality requires tremendous computing power to be lifelike.

Figure 9.17 Using HMD and Data Gloves

In virtual reality, cyber space is made up of thousands of geometric objects and points plotted in a three-dimensional space; the more the number of objects and points that describe the objects, the higher the resolution, therefore a more realistic view will be seen. Each motion or action requires the computer to perform recalculation on the position, angle and shape. Some of the devices that are used in virtual reality are as follows:

- **Head Mounted Display (HMD)**: The HMD was the first device providing its wearer with an immersive experience. A typical HMD has two miniature display screens and an optical system that channels the images from the screens to the eyes, thereby presenting a 3D view of the virtual world. A motion tracker continuously measures the position and orientation of the user's head and allows the image-generating computer to adjust the scene presentation to the current view. As a result, the viewer can look around and walk through the surrounding virtual environment.
- **Data Gloves**: The glove is equipped with sensors that sense the movements of the hand and interfaces those movements with a computer. Data gloves are commonly used in virtual reality environments where the user sees an image of the data glove and can manipulate the movements of the virtual environment using the glove.

- **Software**: Software gives the wearer of this special headgear and gloves the interactive sensory experience that feels like an alternative to real-world experiences.

9.6.1 Characteristics of Virtual Reality

Virtual reality presents many characteristics, some of which are as follows:

- Head-referenced viewing provides a natural interface for navigation in three-dimensional space and allows for look-around, walk-around and fly-through capabilities in virtual environments.
- Stereoscopic viewing enhances the perception of depth and the sense of space.
- The virtual world is present in a full scale and relates properly to the human size.
- Realistic interactions with virtual objects via data glove and similar devices allow for manipulation, operation and control of virtual worlds.
- The convincing illusion of being fully immersed in an artificial world can be enhanced by auditory and other non-visual technologies.
- Networked applications allow for shared virtual environments.

Note that a virtual environment can represent any three-dimensional world that is either real or abstract. This includes real systems like buildings, landscapes, underwater shipwrecks, spacecrafts, archaeological excavation sites, human anatomy, sculptures, crime scene reconstructions, solar systems and so on. Of special interest is the visual and sensual representation of abstract systems like magnetic fields, turbulent flow structures, molecular models, mathematical systems, auditorium acoustics, stock market behaviour, population densities, information flows, and any other conceivable system including artistic and creative work of abstract nature. These virtual worlds can be animated, interactive, shared and can expose behaviour and functionality.

LET US SUMMARIZE

1. *Multimedia* can be described as an integration of multiple media elements together to influence the given information so that it can be presented in an attractive and interactive manner.
2. Humankind developed text and symbols so that they could communicate. Writing helped humans to remember things for a much longer period than what is possible with the help of

memory (brain). Text can be written using various text-editing and processing tools such as word processing software.

3. *Graphics* is regarded as the key building block for multimedia projects. It can be implemented in a multimedia project in various forms like images, charts, flowcharts, 2D/3D graphics and so on.
4. *Audio* is the most sensuous element of multimedia. It is a physical phenomenon to stimulate our senses of hearing. It involves transmission of mechanical vibration by an elastic medium that is capable of being detected by the human ear.
5. *Video* refers to the moving picture accompanied by sound such as a movie on television. It is an excellent form of conveying a message to a large audience, adds richness to content and enhances the user experience. Integrating video elements in a multimedia application provides a great deal of information in a small duration of time. There are three main types of video file formats: QuickTime, AVI and MPEG.
6. *Animation* is the process of sequencing still images in rapid succession to give the effect of live motion. The sequence of images is shown one after the other very quickly so that the human eye and brain blend the images to produce the effect of motion. It is of two types, namely, 2D and 3D animation.
7. Kinematics and morphing are the two most commonly used animation effects. *Kinematics* is the study of movement and motion of structures that have joints, such as a walking man. It requires calculating position, rotation, velocity and acceleration of all the joints. *Morphing* involves transformation of one image into another.
8. A *multimedia system* is a system capable of processing multimedia data and applications. It is characterized by the processing, generation, manipulation and rendition of multimedia information.
9. There are two types of compression techniques: *lossy* and *lossless*. Lossy compression creates a smaller file by discarding some of the original information about the image whereas lossless compression reduces file size without permanently discarding information from the original file.
10. Multimedia is used in various fields like *education, entertainment, training, business* and *information kiosks*.
11. *Virtual reality* can be described as an artificial environment that is created by using special computer hardware and software, which is presented to the user in such a manner that it appears and feels like a real three-dimensional world. To enter in virtual reality, a user wears special gloves, earphones and HMD, all of which receive input from the computer system.

EXERCISES

Fill in the Blanks

1. ..... constitutes an integration of multiple media such as text, graphics, audio, video and animation elements.
2. .... is the object in the text that allows jumping from one web page to another page at the same site or to another site on the Web.
3. ..... means the removal of blank spaces from the front of the recording.

4. JPEG uses ..... compression technique whereas PNG uses ..... compression technique.
5. ..... files are the Microsoft standard file format for storing waveform audio data.
6. The process of removal of extraneous noise that inevitably creeps into a recording is called .....
7. ..... is extremely high quality file format that delivers higher image resolution, picture quality, interlaced video format, multi-resolution scalability, and multi-channel audio features.
8. ..... refers to the study of movement and motion of structures that have joints.
9. ..... can be described as an artificial environment created from the computer hardware and software, presented to the user in such a manner that it appears and feels like a real environment.
10. The amount of data that can be transmitted from one computer to another in a given time frame is called .....

Multiple-choice Questions

1. Text element of multimedia is predominant in .....
    1. Newspaper
    2. Movie
    3. Radio
    4. None of these
2. Text that allows skipping from one web page to another page at the same site or to another site somewhere else on the Web at a click of a button is called .....
    1. Normal text
    2. Animated text
    3. Hypertext
    4. None of these
3. Which of the following are the features of multimedia application?
    1. Use of digital technology
    2. Interaction from the user
    3. Integration of text, sound and video
    4. All of these
4. A flowchart is a kind of .....
    1. One-dimensional graphics
    2. Two-dimensional graphics
    3. Three-dimensional graphics
    4. None of these
5. The process of removal of blank spaces from the front of the recording is called.....
    1. Splicing
    2. Morphing
    3. Equalizing
    4. Trimming
6. A high-quality file format that delivers better picture quality for the motion picture is.....
    1. QuickTime
    2. AVI
    3. JPEG
    4. MPEG

7. Technique that involves transformation of one image into another is called .....
    1. Kinematics
    2. Compression
    3. Morphing
    4. Time stretching
8. The most appropriate form of graphics used for presenting facts and figure information is.....
    1. Flowchart
    2. Background
    3. Charts and graphs
    4. 2D graphics
9. A device having two miniature display screens and an optical system that channels the images from the screens into the user's eyes is called .....
    1. HMD
    2. BOOM
    3. CAVE
    4. None of these
10. Which of the following image file format uses lossless compression?
    1. MPEG
    2. JPEG
    3. PNG
    4. None of these

State True or False

1. Graphics is one of the building blocks of multimedia.
2. The text which allows users to jump from one web page to another at a click of a button is called rich text.
3. Adobe Photoshop is an image editing software.
4. Music can be used to set moods and convey ideas to the user.
5. Trimming operation in sound files refers to the modification of the recording frequency content into brighter sound.
6. JPEG file format is used for creating video.
7. Morphing and Kinematics are the two most important animation techniques.
8. Lossy compression gives better image quality than lossless compression.
9. Virtual reality is an artificial environment, created by special hardware and software.
10. HMD stands for Head Mounted Display.

Descriptive Questions

1. What is multimedia? State the role of elements that together constitute multimedia.
2. Differentiate between normal and rich text.
3. What are the characteristics of virtual reality?
4. Explain some of the common uses of multimedia in various fields.
5. What is a multimedia system? Explain all desirable features of a good multimedia system.
6. How does the virtual reality system work? What components are used in the development of a virtual reality environment?

7. Explain the importance and role of animation in making a multimedia application.

ANSWERS

Fill in the Blanks

1. Multimedia
2. Hypertext
3. Trimming
4. Lossy, Lossless
5. WAV
6. Splicing
7. MPEG
8. Kinematics
9. Virtual reality
10. Bandwidth

Multiple-choice Questions

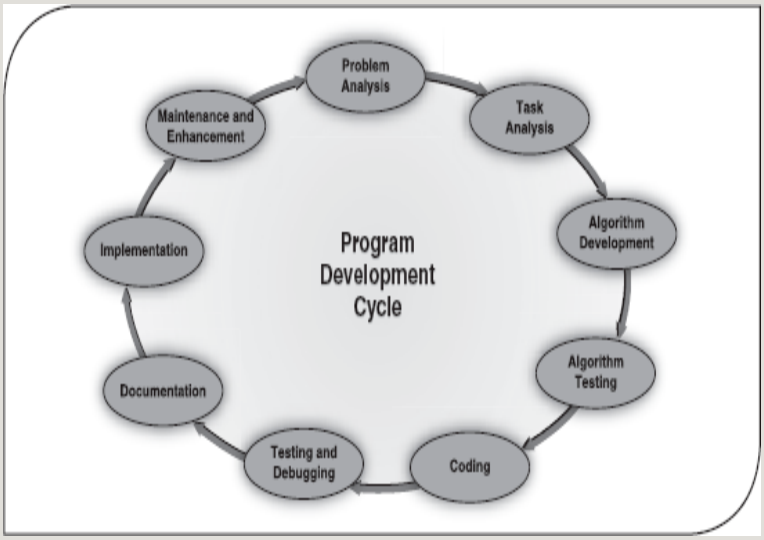
1. (a)
2. (c)
3. (d)
4. (b)
5. (d)
6. (d)
7. (c)
8. (c)
9. (a)
10. (c)

State True or False

1. True
2. False
3. True
4. True
5. False
6. False
7. True
8. False
9. True
10. True

Chapter 10

Computer Programming and Languages



This chapter introduces the reader with the fundamentals of computer programming and languages. A computer program refers to the set of instructions needed for performing a particular task. Computer or programming language refers to the vocabulary and set of grammatical rules that are required to instruct a computer to perform a specific task. The chapter begins with the life cycle of program development and various programming tools which are essential for making a computer program. The discussion continues with various programming control structures and programming paradigms. Next, the chapter explains the evolution of various programming languages and generations of programming languages. The reader is also familiarised with the working of

languge translators, namely, assembler, compiler, and interpreter.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The prominent concepts related to computer programming and languages
- The algorithm that is a basic tool used for developing logic of a given problem
- The flowchart—a pictorial representation of an algorithm
- The decision table—a way of representing an algorithm in a tabular form
- The pseudocode—a generic way of describing an algorithm without using any specific programming language
- How control structures specify the work to be done that affects the flow of the simulation code
- Programming paradigm, which refers to the way a program is written in order to solve a problem
- Various programming languages used to communicate with computers
- Different generations of programming languages
- Features of a good programming language

10.1 INTRODUCTION

The computer has emerged as the most useful machine in recent times. It can perform a variety of tasks like receiving data, processing it and producing useful results. However, being a machine, the computer cannot perform on its own. A computer needs to be instructed to perform even a simple task like adding two numbers. Computers work on a set of instructions called *computer program*, which clearly specify the ways to carry out a

task. An analogy of this may be thought of as the instructions given by the manager or team leader to its team. The team members follow those instructions and accordingly perform their duties. Similarly, a computer also takes instructions in the form of computer programs to carry out the requested task.

Now the question that arises is how human beings instruct computers. We, as human beings, use natural languages such as English, Spanish or French to communicate. Similarly, a user communicates with the computer in a language understood by it. Note that human beings cannot interact directly with the computer using natural languages because thus far we have not developed such computers that can comprehend natural languages. Rather, the instructions, provided in the form of computer programs, are developed using *computer or programming languages*. This chapter provides some of the most prominent concepts related to computer programming and languages.

10.1.1 Developing a Program

As discussed earlier, a program consists of a series of instructions that a computer processes to perform the required operation. In addition, it also includes some fixed data required to perform the instructions and the process of defining those instructions and data. Thus, in order to design a program, a programmer must determine three basic rudiments:

- The instructions to be performed.
- The order in which those instructions are to be performed.
- The data required to perform those instructions.

To perform a task using a program, a programmer has to consider various inputs of the program along with the process, which is required to convert the input into desired output. Suppose we want to calculate the sum of two numbers, A and B, and store the sum in C; here A

and B are the inputs, addition is the process, and C is the output of the program (see Figure 10.1).

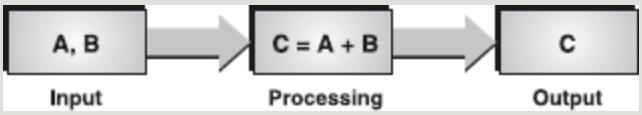


Figure 10.1 Program Performing a Task

FACT FILE

Development of a Program

A program is a set of instructions that instructs a computer on how to perform a specific task. It is written in a programming language that can be implemented on a number of different processors. A collection of programs can be compared to a recipe book, where each recipe can be assumed as a program. Every recipe has a list of ingredients (fixed data) and a list of instructions detailing exactly what to do with those ingredients. When you follow a recipe, you are actually executing a program.

10.1.2 Program Development Cycle

Before starting the process of writing a program (coding), the programmer has to determine the problem that needs to be solved. There are different approaches to problem solving. One common approach is to use the *program development cycle*, with the number of steps that may vary according to the person who has

formalized the development. Often the process runs in a loop, for example, as the current process is completed, new demands appear and the development process commences again. Illustrated in Figure 10.2 is the development cycle of a program, which includes the following phases.

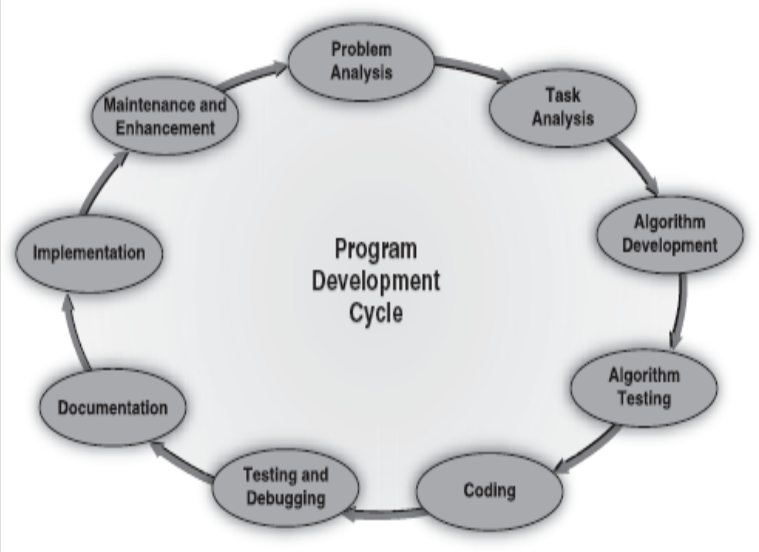


Figure 10.2 Program Development Cycle

- **Problem Analysis:** The problem is analysed precisely and completely. Based on understanding, the developer knows about the scope within which the problem needs to be developed.
- **Task Analysis:** After analysing the problem, the developer needs to develop various solutions to solve the given problem. From these solutions, the optimum solution is chosen, which can solve the problem comfortably and economically.
- **Algorithm Development:** After selecting the appropriate solution, an algorithm is developed to depict the basic logic of the selected solution. An algorithm depicts the solution in logical steps (sequence of instructions). Further, an algorithm is represented by flowcharts, decision tables and pseudocodes. These tools make the program logic clear and they eventually help in coding.
- **Algorithm Testing:** Before converting the algorithms into actual code, it should be checked for accuracy. The main purpose of checking the algorithm is to identify major logical errors at an early stage because logical errors are often difficult to detect and correct at later stages. The testing also ensures that the algorithm is a “true” one and it should work for both normal as well as unusual data.

- **Coding:** After meeting all the design considerations, the actual coding of the program takes place in the chosen programming language. Depending upon the application domain and available resources, a program can be written by using computer languages of different levels such as machine, assembly or high-level languages (HLL).
- **Testing and Debugging:** It is common for the initial program code to contain errors. A program compiler and programmer-designed test data machine tests the code for syntax errors. The results obtained are compared with results calculated manually from these test data. Depending upon the complexity of the program, several rounds of testing may be required.
- **Documentation:** Once the program is free from all the errors, it is the duty of the program developers to ensure that the program is supported by suitable documentation. These documents should be supplied to the program users. Documenting a program enables the user to operate the program correctly. It also enables other persons to understand the program clearly so that it may, if necessary, be modified, or corrected by someone other than the original programmer.
- **Implementation:** After documentation, the program is installed on the end user's machine and the user is also provided with all the essential documents in order to understand how the program works. The implementation can be viewed as the final testing because only after using the program the user can point out the drawbacks (if any) and report them to the developers. Based on the feedback from users, the programmers can modify or enhance the program.
- **Maintenance and Enhancement:** After the program is implemented, it should be properly maintained by taking care of the changing requirements of its users and the system. The program should be regularly enhanced by adding additional capabilities. This phase is also concerned with detecting and fixing the errors, which were missed in the testing phase. Since this step generates user feedback, the programming cycle continues as the program is modified or reconstructed to meet the changing needs.

10.2 ALGORITHM

Algorithms are one of the most basic tools that are used to develop the problem-solving logic. An *algorithm* is defined as a finite sequence of explicit instructions that when provided with a set of input values produces an output and then terminates. To be an algorithm, the steps must be unambiguous and after a finite number of steps, the solution of the problem should be achieved. However, algorithms can have steps that repeat (iterate) or require decisions (logic and comparison) until the task is completed.

THINGS TO REMEMBER

Algorithm Properties

Algorithms are not computer programs, as they cannot be executed by a computer. Some properties of algorithm are as follows.

- There must be no ambiguity in any instruction.
- There should not be any uncertainty about which instruction is to be executed next.
- The description of the algorithm must be finite. An algorithm cannot be open ended.
- The algorithm should terminate after a finite number of steps.
- The algorithm must be general enough to deal with any contingency.

Different algorithms may accomplish the same task, with a different set of instructions, in more or less the same time, space and efforts. For example, two different recipes for preparing tea: one “add the sugar” while “boiling the water” and the other “after boiling the water” produce the same result. However, performing an algorithm correctly does not guarantee a solution, if the algorithm is flawed or not appropriate to the context. For example, preparing the tea algorithm will fail if there are no tea leaves present; even if all the motions of preparing the tea are performed as if the tea leaves were there. We use algorithms in our daily life. For example, to determine the largest number out of three numbers A, B and C, the following algorithm may be used.

Algorithm: To Determine Largest of Three Numbers

1. `Start`

2. Read three numbers A, B, C
3. Find the larger number between A and B and store it in MAX_AB
4. Find the larger number between MAX_AB and C and store it in MAX
5. Display MAX
6. Stop

The above-mentioned algorithm terminates after six steps. This explains the feature of finiteness. Every action of the algorithm is precisely defined, hence there is no scope for ambiguity.

Once an algorithm has been designed, it can be represented as a flowchart, pseudocode or decision table, which are further expressed in programming language to develop computer programs.

10.3 FLOWCHART

A *flowchart* is a pictorial representation of an algorithm in which the steps are drawn in the form of different shapes of boxes and the logical flow is indicated by interconnecting arrows. The boxes represent operations and the arrows represent the sequence in which the operations are implemented. The primary purpose of the flowchart is to help the programmer in understanding the logic of the program. Therefore, it is always not necessary to include all the required steps in detail. Flowcharts outline the general procedure. Since they provide an alternative, visual way of representing the information flow in a program, program developers often find them very valuable.

10.3.1 Flowchart Symbols

Flowcharts can be compared with the blueprint of a building. Just as an architect draws a blueprint before

starting the construction of a building, a programmer draws a flowchart before writing a computer program. As in the case of the drawing of a blueprint, the flowchart is drawn according to defined rules and using standard flowchart symbols prescribed by American National Standard Institute (ANSI). Some standard symbols that are frequently required for flowcharts are shown in Table 10.1.

Table 10.1 Flowchart Symbols

Symbol	Symbol Name	Description
	Flow Lines	Flow lines are used to connect symbols. These lines indicate the sequence of steps and the direction of flow of control.
	Terminal	This symbol is used to represent the beginning (start), the termination (end) or halt (pause) in the program logic.
	Input/Output	This symbol represents information entering or leaving the system such as customer order (input) and servicing (output).
	Processing	This symbol is used for representing arithmetic and data movement instructions. It can represent a single step (“add two cups of flour”), or an entire sub-process (“make bread”) within a larger process.
	Decision	This symbol denotes a decision (or branch) to be made. The program should continue along one of the two routes (IF/ELSE). This symbol has one entry and two exit paths. The path chosen depends on whether the answer to a question is “yes” or “no”.
	Connector	This symbol is used to join different flow lines.
	Off-page Connector	This symbol is used to indicate that the flowchart continues on the next page.
	Document	This symbol is used to represent a paper document produced during the flowchart process.
	Annotation	This symbol is used to provide additional information about another flowchart symbol. The content may be in the form of descriptive comments, remarks or explanatory notes.
	Manual Input	This symbol represents input to be given by a developer/programmer.
	Manual Operation	This symbol shows that the process has to be done by a developer/programmer.
	Online Storage	This symbol represents the online data storage such as hard disks, magnetic drums, or other storage devices.

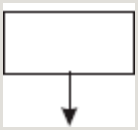
Symbol	Symbol Name	Description
	Offline Storage	This symbol represents the offline data storage such as sales on OCR and data on punched cards.
	Communication Link	This symbol is used to represent data received or to be transmitted from an external system.
	Magnetic Disk	This symbol is used to represent read from or write to a magnetic disk.

Note: *Some sample flowchart's have been represented in Section 10.6.*

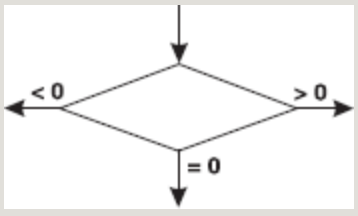
10.3.2 Guidelines for Preparing Flowcharts

The following guidelines should be used for creating a flowchart:

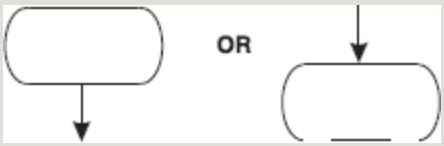
- The flowchart should be clear, neat, and easy to follow.
- The flowchart must have a logical start and finish.
- In drawing a proper flowchart, all necessary requirements should be listed in logical order.
- Only one flow line should come out from a process symbol.



- Only one flow line should enter a decision symbol. However, two or three flow lines (one for each possible answer) may leave the decision symbol.



- Only one flow line is used with a terminal symbol.



- Write briefly within standard symbols. If necessary, use the annotation symbol to describe data or process more clearly.
- In case of complex flowcharts, connector symbols are used to reduce the number of flow lines.
- Intersection of flow lines should be avoided to make it a more effective and better way of representing communication.
- It is useful to test the validity of the flowchart by passing through it with normal/ unusual test data.

10.3.3 Benefits of Flowcharts

A flowchart helps to clarify how things are currently working and how they could be improved. It also assists in finding the key elements of a process by drawing clear lines between the end of one process and the start of next one. Developing a flowchart establishes a common understanding among the participants about the process. Flowcharts help in revealing redundant or misplaced steps. They also help in establishing important areas for monitoring or data collection and to identify areas for improvement or increase in efficiency. The reasons for using flowcharts as a problem-solving tool are given below.

- **Makes Logic Clear:** The main advantage of using a flowchart to plan a task is that it provides a pictorial representation of the task, which makes the logic easier to follow. The symbols are connected in such a way that they show the movement (flow) of information through the system visibly. The steps and how each step is connected to the next can be clearly seen. Even less experienced personnel can trace the actions represented by a flowchart, that is, flowcharts are ideal for visualizing fundamental control structures employed in computer programming.
- **Communication:** Being a graphical representation of a problem-solving logic, flowcharts are a better way of communicating the logic of a system to all concerned. The diagrammatical representation of logic is easier to communicate to all the interested parties as compared to actual program code as the users may not be aware of all the programming techniques and jargons.

- **Effective Analysis:** With the help of a flowchart, the problem can be analysed in an effective way. This is because the analysing duties of the programmers can be delegated to other persons, who may or may not know the programming techniques, as they have a broad idea about the logic. Being outsiders, they often tend to test and analyse the logic in an unbiased manner.
- **Useful in Coding:** The flowcharts act as a guide or blueprint during the analysis and program development phase. Once the flowcharts are ready, the programmers can plan the coding process effectively as they know where to begin and where to end, making sure that no steps are omitted. As a result, error-free programs are developed in HLL and that too at a faster rate.
- **Proper Testing and Debugging:** By nature, a flowchart helps in detecting the errors in a program, as the developers know exactly what the logic should do. Developers can test various data for a process so that the program can handle every contingency.
- **Appropriate Documentation:** Flowcharts serve as a good program documentation tool. Since normally programs are developed for novice users, they can take the help of the program documentation to know what the program actually does and how to use the program.

10.3.4 Limitations of Flowcharts

A flowchart can be used for designing the basic concept of the program in pictorial form but cannot be used for programming purposes. Some of the limitations of the flowchart are given below.

- **Complex:** The major disadvantage in using flowcharts is that when a program is very large, the flowcharts may continue for many pages, making them hard to follow. Flowcharts tend to get large very quickly and it is difficult to follow the represented process. It is also very laborious to draw a flowchart for a large program. You can very well imagine the nightmare when a flowchart is to be developed for a program, consisting of thousands of statements.
- **Costly:** Drawing flowcharts are viable only if the problem-solving logic is straightforward and not very lengthy. However, if flowcharts are to be drawn for a huge application, the time and cost factor of program development may get out of proportion, making it a costly affair.
- **Difficult to Modify:** Due to its symbolic nature, any changes or modification to a flowchart usually requires redrawing the entire logic again, and redrawing a complex flowchart is not a simple task. It is not easy to draw thousands of flow lines and symbols along with proper spacing, especially for a large complex program.
- **No Update:** Usually programs are updated regularly. However, the corresponding update of flowcharts may not take place, especially in the case of large programs. As a result, the logic used in the flowchart may not match with the actual program's logic. This inconsistency in flowchart update defeats the main purpose of the flowcharts, that is, to give the users the basic idea about the program's logic.

Decision tables are generally useful in situations where a number of decisions or actions have to be taken depending on the input provided. A *decision table* is a way of representing an algorithm in a tabular form. It contains all the possible conditions for a specific problem and the corresponding results using condition rules that connect conditions with results. It consists of rows and columns, divided into four separate quadrants as shown in Figure 10.3.

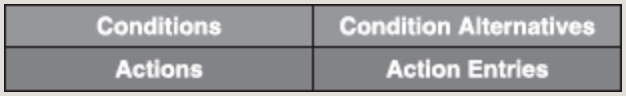


Figure 10.3 Four Quadrants of a Decision Table

- **Conditions:** Contains a list of all the possible conditions pertaining to a problem.
- **Condition Alternatives:** Contains the condition rules (set of possible values for each condition) of alternatives.
- **Actions:** Contains actions, which can be a procedure or operation to be performed.
- **Action Entries:** Contains the action rules, which specify the actions to be performed on the basis of the set of condition alternatives corresponding to that action entry.

Note that each column in the quadrant can be interpreted as a *processing rule*. The basic four-quadrant structure is common for all the decision tables; however, they may differ in the way of representation of the condition alternatives and action entries. For example, in some decision tables, condition alternatives are represented using true/false values while in other tables, Yes/No or numbers (0 or 1) may be used. Similarly, in decision tables, action entries are represented by placing check marks, cross marks or numbering the actions to be performed.

10.4.1 Steps to Develop a Decision Table

In order to build decision tables, first the maximum size of the table is determined. Any impossible situations, inconsistencies or redundancies are then eliminated from the table. To develop a decision table, follow the steps given below.

1. Determine all actions that can be associated with a specific module or procedure.
2. Determine all the conditions affecting the decision.
3. Associate the specified conditions with specific actions. In the simplest case, each condition in the decision table will be associated with two alternatives (Yes or No).
4. Define rules by indicating actions that occur for the specified conditions.

To understand the concept of a decision table, let us consider an example to find the largest of the three numbers A, B and C. The decision table for this example is shown in Figure 10.4.

Conditions	Rules		
	1	2	3
A > B	Y	N	—
A > C	Y	—	N
B > C	—	Y	N
Actions			
A is Largest	X		
B is Largest		X	
C is Largest			X

Figure 10.4 Decision Table for Finding the Largest of Three Numbers

10.4.2 Benefits of Decision Tables

A decision table is generally preferred over the flowchart because of the following benefits.

- It is easier to draw and Modify a decision table as compared to a flowchart.
- For the algorithms having complex problem-solving logic, a flowchart may take a number of pages; however using decision tables, such

alogorithms can be represented in a compact form.

- Each column in a decision table represents a unique combination of conditions and their corresponding actions. This helps to better understand the decision logic of the program.

10.4.3 Limitations of Decision Tables

- Most programmers and beginners are more familiar to flowcharts rather than decision tables, and thus prefer to use flowcharts.
- Greater the size of a decision table, more difficult will be to Modify it.

10.5 PSEUDOCODE

Pseudocode is made up of two words: *pseudo* and *code*. Pseudo means imitation and code refers to instructions, written in a programming language. As the name suggests, pseudocode is not a real programming code, but it models and may even look like programming code. It is a generic way of describing an algorithm without using any specific programming language-related notations. Simply put, pseudocode is an outline of a program, written in a form that can be easily converted into actual programming statements. Pseudocode uses plain English statements rather than symbols to represent the processes in a computer program. It is also known as *PDL* (*program design language*), as it emphasizes more on the design aspect of a computer program or structured English, because usually pseudocode instructions are written in normal English, but in a structured way.

Pseudocode strikes a fine balance between the understandability and informality of a natural language like English and the precision of a computer program code. It is somewhat halfway between English and a programming language. If an algorithm is written in English, the description may be at such a high level that it may prove difficult to analyse the algorithm and then to transform it into actual code. If instead, the algorithm is written in code, the programmer has to invest a lot of time in determining the details of an algorithm, which he may choose not to implement (since, typically,

algorithms are analysed before deciding which one to implement). Therefore, the goal of writing pseudocode is to provide a high-level description of an algorithm, which facilitates analysis and eventual coding, but at the same time suppresses many of the details that are insignificant.

Pseudocode uses some keywords to denote programming processes. Some of them are as follows.

- **Input:** READ, OBTAIN, GET and PROMPT
- **Output:** PRINT, DISPLAY and SHOW
- **Compute:** COMPUTE, CALCULATE and DETERMINE
- **Initialize:** SET and INITIALIZE
- **Add One:** INCREMENT

The pseudocode given below calculates the area of a rectangle.

Pseudocode: To Calculate the Area of A Rectangle

```
PROMPT the user to enter the height of
the rectangle

PROMPT theuser to enter the width of
the rectangle

COMPUTE the area by multiplying the
height with width

DISPLAY the area

STOP
```

Since pseudocode is detailed yet readable, it can be inspected by the team of designers and programmers as a way to ensure that actual programming is likely to match the design specifications. It is better to catch

errors at the pseudocode stage rather than correcting them in later stages as it would prove expensive. Once the pseudocode is accepted, it is transformed into actual program code using the vocabulary and syntax of the chosen programming language. The benefit of pseudocode is that it enables the programmer to concentrate on the algorithms without worrying about all the syntactic details of a particular programming language. In fact, you can write pseudocode without even knowing what programming language you will use for the final implementation. Often computer textbooks use pseudocode in their examples so that all programmers can understand them, even if they do not know the same programming languages.

10.5.1 Pseudocode Guidelines

Writing pseudocode is not a difficult task. Even if you do not know anything about computers or computer languages, you can still develop effective and efficient pseudocodes, if you are writing in an organized manner. Although there are no established standards for pseudocode construction, a few general guidelines for developing pseudocodes are given as follows.

- Statements should be written in simple English (or any preferable natural language) and should be programming language independent. Remember that pseudocodes only describe the logic plan to develop a program, it is not programming.
- Steps must be understandable, and when the steps (instructions) are followed, they must produce a solution to the specified problem. If the pseudocode is difficult for a person to read or translate into code, then something is wrong with the level of detail you have chosen to use.
- Pseudocodes should be concise.
- Each instruction should be written in a separate line and each statement in pseudocode should express just one action for the computer. If the task list is properly drawn, then in most cases each task will correspond to one line of pseudocode.
- Capitalize keywords such as READ, PRINT and so on.
- Each set of instructions is written from top to bottom, with only one entry and one exit.
- It should allow for easy transition from design to coding in the programming language.

10.5.2 Benefits of Pseudocode

Programming can be a complicated process when the program requirements are complex in nature. Pseudocode provides a simple method of developing the program logic as it uses everyday language to prepare a brief set of instructions in the order in which they appear in the completed program. It allows the programmer to focus on the steps required to solve a program rather than on how to use the computer language. Some of the most significant benefits of pseudocode are as follows.

- Since it is language independent, it can be used by most programmers. It allows the developer to express the problem logic in plain natural language.
- It is easier to develop a program from a pseudocode rather than from a flowchart or decision table. Programmers do not have to think about syntaxes; they simply have to concentrate on the underlying logic. The focus is on the steps to solve a problem rather than on how to use the computer language.
- The words and phrases used in pseudocode are in line with basic computer operations. This simplifies the translation from the pseudocode to a specific programming language.
- Unlike flowcharts, pseudocode is compact and does not tend to run over many pages. Its simple structure and readability makes it easier to Modify.

10.5.3 Limitations of Pseudocode

Although pseudocode is a very simple mechanism to simplify problem-solving logic, it has its limitations. Some of the most notable limitations are as follows.

- It does not provide a visual representation of the program's logic.
- There are no accepted standards for writing pseudocodes. Programmers use their own style of writing pseudocode.
- It is quite difficult for the beginners to write pseudocode as compared to drawing a flowchart.

10.6 PROGRAM CONTROL STRUCTURES

A program is usually not limited to a linear sequence of instructions. During its process, it may diverge, repeat code or take decisions. For these purposes, control structures specify what has to be done to perform the program. Program statements that affect the order in

which statements are executed, or that affect whether statements are executed, are called *control structures*. They affect the flow of simulation code since a control structure evaluates statements and then executes code according to the result.

Control structures are used to express algorithms, flowcharts, decision tables and pseudocodes as actual computer programs. Essentially, there are three control structures.

- **Sequence:** where information flows in a straight line.
- **Selection:** (decision or branched), where the decisions are made according to some predefined condition.
- **Repetition:** (looping), where the logic (sequence of steps) is repeated in a loop until the desired output is obtained.

10.6.1 Sequence Control Structure

As the name implies, in a sequence structure, the instructions to be computed simply follow one another in a logical progression. This structure is denoted by writing one action after another, each action on a line by itself, and all actions aligned with the logical indent. The actions are performed in the same sequence (top to bottom) in which they are written (see Figure 10.5).

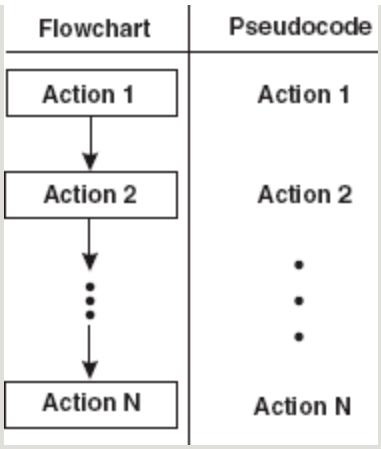
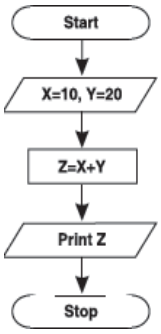


Figure 10.5 Flowchart and Pseudocode for Sequence Construct

Typical sequence operations consist of process and I/O steps. The following example illustrates a simple sequence control structure that adds two numbers. Note that the actions are flowing in a logical manner from top to bottom, as shown in Table 10.2. Not a single process branches out and no process is repeated. Each process is contributing something to the next process.

Table 10.2 Sequence Control Structure

Example Flowchart	Step-by-Step Instructions
	Begin the flowchart.
	Initialize value for memory variables X and Y.
	Add X and Y and store the result in variable Z.
	Print the value of variable Z as output.
	End the flowchart.

10.6.2 Selection Control Structure

A selection structure allows the program to make a choice between two alternate paths, whether it is *true* or *false*. The first statement of a selection structure is a conditional statement. Once the sequence of steps in the selected path has been carried out, the paths are rejoined and then the next instruction is carried out. Thus, the selection structure has only a single entry and a single exit. From Figure 10.6 it is clear that if the condition is true, Action 1 will be performed, otherwise Action 2 will be performed. After performing the list of actions, the control of the program moves on to the other actions in the process flow.

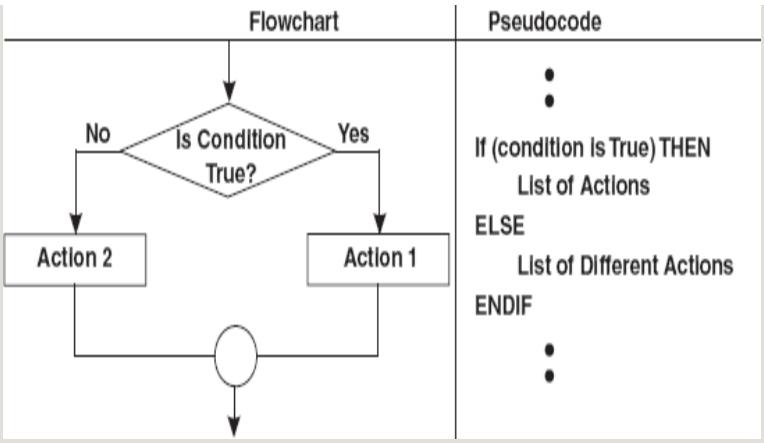
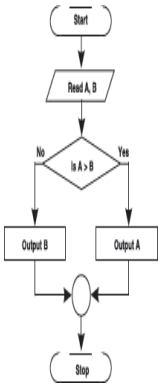


Figure 10.6 Flowchart and Psedocode for Selection Construct

The example given in Table 10.3 finds the larger of the two numbers.

Table 10.3 Selection Control Structure

Example Flowchart	Step-by-Step Instructions
	Begin the flowchart.
	Read two numbers and store the value in memory variables A and B, respectively.
	Check whether the value of A is greater than the value of B.
	If A is greater than B then display A, otherwise display B.
	Connector symbol is used to join the flow lines.
	End the flowchart.

10.6.3 Repetition Control Structure

Repetition or loop pattern causes an interruption in the normal sequence of processing. It directs the system to move back to a previous statement in the program, repeating the same sequence over and again, usually with new data. When a sequence of statements is repeated against a condition, it is said to be in a *loop*. Using looping, the programmer avoids writing the same set of instructions again (see Figure 10.7). The looping process can either be one time or multiple times until the desired output is obtained within a single program.

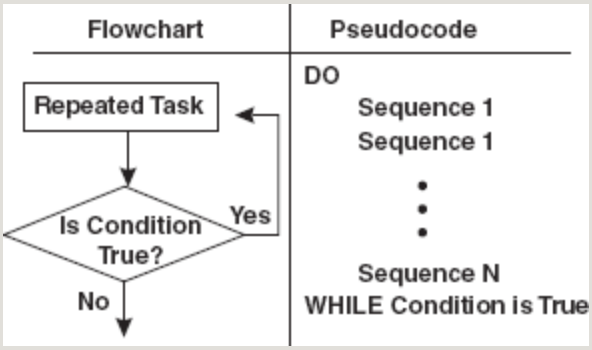
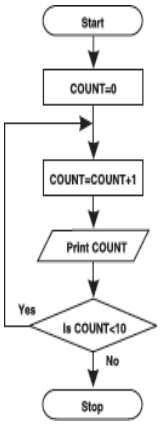


Figure 10.7 Flowchart and Pseudocode for Repetition Construct

The example shown in Table 10.4 prints the first 10 natural numbers. Notice that at the beginning, the value of COUNT variable is initialized to zero. After that, the value of COUNT is increased by one and is printed. If the value of COUNT is less than 10 (which represents the count of numbers to be displayed) then the same process is repeated. At the end of each iteration, if the condition evaluates to true, the loop repeats; otherwise the loop terminates.

Table 10.4 Repetition Control Structure

Example Flowchart	Step-by-Step Instructions
	Begin the flowchart.
	Initialize the memory variable COUNT to zero.
	Increment the value of COUNT by 1.
	Print the value of COUNT.
	Check whether the value of COUNT is less than 10. If yes, then move back to the third step otherwise go to the next step.
	End the flowchart.

In all the above examples, we have used a term known as memory variable. *Memory variable* is a programming terminology used to represent the location in the memory where the values are stored.

10.7 PROGRAMMING PARADIGMS

Previously, we have discussed that programming involves many hardships in terms of labour, time and money. Usually a programmer is busy in maintaining an existing application rather than creating a new one. Hence, different types of programming paradigms have been developed in order to minimize the programming efforts. *Programming paradigm* refers to the approach used to develop the program for solving a problem. It provides (and determines) the view that the programmer has of the execution of the program.

Broadly, programming can be classified in the following three categories: *unstructured programming, structured*

programming and *object-oriented programming*.

10.7.1 Unstructured Programming

Unstructured programming refers to writing small and simple programs consisting of only one main program. All the actions, such as providing input, processing and displaying output, are done within one program only. This style of programming is generally restricted for developing a small application, but if the application becomes large then it poses real difficulties in terms of clarity of the code, modifiability and ease of use. Although this type of programming style is not recommended, still most programmers start to learn programming using this technique.

10.7.2 Structured Programming

The programs generated using unstructured approach are meant for simple and small problems. If the problem gets lengthy, this approach becomes too complex and obscure. After some time, even the programmers themselves may find it difficult to understand their own program. Hence, programming should be performed using a “structured” (organized) approach. Using structured programming, a program is broken down into small independent tasks that are small enough to be understood easily, without having to understand the whole program at once. Each task has its own functionality and performs a specific part of the actual processing. These tasks are developed independently, and each task can carry out a specified function on its own. When these tasks are completed, they are combined together to solve the problem. In this way, designers map out the large-scale structure of a program in terms of smaller operations, implement and test the smaller operations, and then tie them together into a whole program. Structured programming can be performed in two ways as follows.

- **Procedural Programming:** This programming has a single program that is divided into small segments called *procedures* (also known as *functions, routines* or *subroutines*). From the main or controlling procedure, a procedure call is used to invoke the required procedure. After the sequence is processed, the flow of control continues from where the call was made. The main program coordinates calls to procedures and hands over appropriate data as parameters. The data is processed by the procedures and once the program has finished, the resultant data is displayed.
- **Modular Programming:** The programs coded with procedural paradigms usually fits into a single code file and it is meant for relatively smaller programs. However, if the program gets large, then modular way of programming is recommended. In case of modular programming, large programs are broken down into a number of smaller program units known as *modules* (see Figure 10.8). Each module is designed to perform a specific function. Note that each module in itself is a small program which can be compiled and tested by an individual program and all modules are combined to build the whole program.

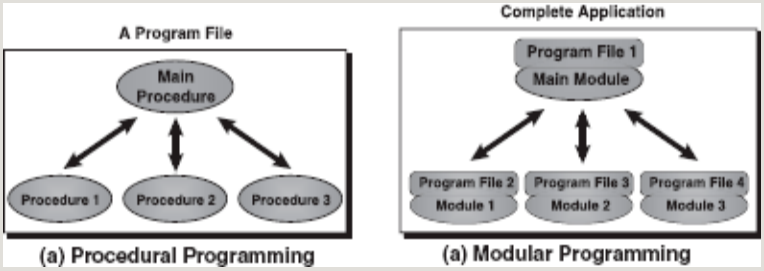


Figure 10.8 Structure Programming

10.7.3 Object-oriented Programming

Object-oriented programming (OOP) is a style of computer programming which promotes building of independent pieces of code that interact with each other. It is an evolutionary form of modular programming, with more formal rules. It allows pieces of programming code to be reused and interchanged between programs. In OOP, programs are organized as cooperative collections of *objects*, each of which represents an instance of some *class*, and whose classes are all members of a hierarchy of classes united by way of *inheritance* relationships. In such programs, classes are generally viewed as static,

whereas objects typically have a much more dynamic nature, which is encouraged by the existence of *polymorphism*. OOP paradigm emphasizes on the following aspects:

- Objects
- Abstraction
- Inheritance
- Classes
- Encapsulation
- Polymorphism

Object: *Objects* are key to understanding the object-oriented paradigm. An object is like a “black box” which contains *code* (sequence of computer instructions) and *data* (information which the instructions operates on). Traditionally, code and data have been kept apart. For example, in a procedural language like C, units of code are called *procedures*, while units of data are called *structures*. Procedures and structures are not formally connected in C. A procedure can operate on more than one type of structure, and more than one procedure can operate on the same structure. As a procedural program grows in size, the network of interaction between procedures and data becomes increasingly complex and hard to manage. However, in OOP, code and data are merged into a single item—an object. Objects can be combined into structured networks to form a complete program, similar to how the pieces in a puzzle fit together to create a picture. The data in an object specifies its attributes, while the code specifies the actions. For example, a car (object) has attributes like four wheels and five gears, and actions like accelerating and changing gears.

Class: The concept of class is fundamental to the notion of OOP. A *class* is an abstraction that captures the common structure and common behaviour of a set of objects. It defines an object or group of objects and how

the objects should behave. In other words, the class is a template or blueprint that defines the characteristics of an object and describes how the object should look and behave. All objects in a given class are identical in form and behaviour but contain different data in their variables. It means that a class does not represent an object; it represents all the information a typical object should have as well as all the methods it should have.

In the real world, there are many objects of the same kind. For example, the object *car* falls under the class *vehicle*. It has some attributes (*gears, wheels*) and actions (*change gears, apply brakes*) that are common with other vehicle objects like *bus* or *truck*. However, each vehicle's attribute is independent and can be different from that of the other one. For example, a *car* has four wheels while a *bus* can have six wheels (see Figure 10.9).

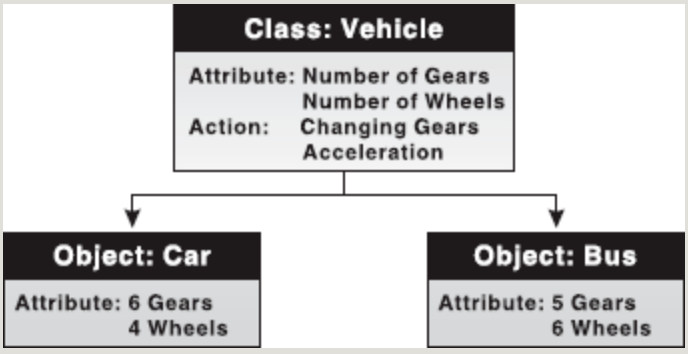


Figure 10.9 Class and Objects

Abstraction: Abstraction is the ability of a program to ignore some aspects of the information it is manipulating and focus only on the essential elements. According to this principle, a general description of an entire class of

objects is defined instead of defining each object individually. For example, we do not think of a car as a set of hundreds of individual parts. Rather, we view a car as an object with its own unique behaviour. This abstraction allows us to drive a car without being bothered by the complexity of the parts involved in making of the car. Hence, we can ignore the details of how the engine and braking system works and focus only on using the object, in this case, driving.

Encapsulation: Previously, we used the “black box” metaphor for explaining the concept of object. A primary rule of OOP is that as the user of an object, one should never need to peek inside the box. The main reason behind it is that objects communicate with each other using *messages*. Messages define the *interface* to the object, that is, the only thing that an object knows about another object is that it exists. Each object's attribute and method is “encapsulated” from other objects. This allows the developer to separate an object's implementation from its behaviour. This separation creates a “black box” effect where the user is isolated from implementation of changes. As long as the interface remains the same, any change to the internal implementation is transparent. For example, if the wheel message is sent to the car object, it does not matter to the user how the developer implemented the code to handle this message. All the sending object needs is the correct protocol for interacting with the car object. The developer can change the implementation at any time, but the same message would still work because the interface is the same (Figure 10.10).

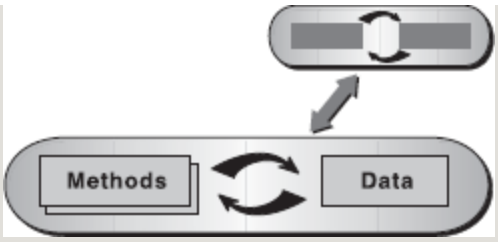


Figure 10.10 Encapsulation and Objects

Polymorphism: Polymorphism is the idea of allowing the same code to be used with different types, resulting in more general and abstract implementations. It allows two or more objects to respond to the same message. An analogy of polymorphism in daily life is how students respond to a school bell. When the bell (message) rings, however, it has its own meaning to different students (objects). Some students go home, some go to the library, and some go to other classes. Every student responds to the bell, but their response is different. The sending object does not have to know how the receiving object implements the message. Only the receiving objects worries about that.

Inheritance: Inheritance is the way to extend or change an already existing parent class (also known as *super class*), and make a *child* class (also known as *subclass* or *derived* class) that stays linked with its parent class. It allows a class to have the same behaviour as another class and extend or tailor that behaviour to provide special action for specific needs. The subclass inherits all the existing messages, and therefore, all the behaviour of the original class. This is the key for code-reusability. Programmers do not have to start from scratch when writing a new program. They can simply reuse an existing repertoire of classes that have behaviours similar to what they need in the new program. For example, after creating the class shapes, one might make a

subclass called triangle, which defines some triangle-specific messages like hypotenuse (Figure 10.11).

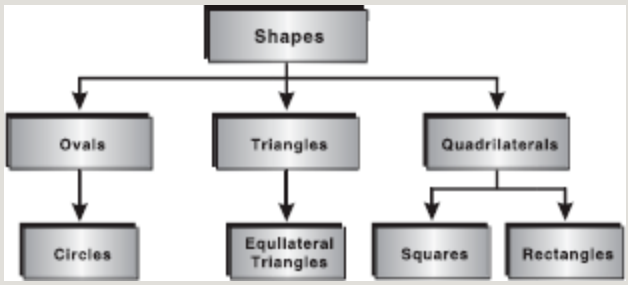


Figure 10.11 Inheritance

10.8 CHARACTERISTICS OF A GOOD PROGRAM

Every computer requires appropriate instruction set (programs) to perform the required task. The quality of the processing depends upon the given instructions. If the instructions are improper or incorrect then it is obvious that the result will also be superfluous. Therefore, proper and correct instructions should be provided to the computer so that it can provide the desired output. Hence, a program should be developed in such a way that it ensures proper functionality of the computer. In addition, a program should be written in such a manner that it is easier to understand the underlying logic. A few important characteristics that a computer program should possess are as follows.

- **Portability:** Portability refers to the ability of an application to run on different platforms (operating systems) with or without minimal changes. Due to rapid development in hardware and software, nowadays platform change is a common phenomenon. Hence, if a program is developed for a particular platform then the life span of the program is severely affected.
- **Readability:** The program should be written in such a way that it makes other programmers or users to follow the logic of the program without much effort. If a program is written structurally, it helps the programmers to understand their own program in a better way. Even if some computational efficiency needs to be sacrificed for better

readability, it is advisable to use a more user-friendly approach, unless the application's processing is of utmost importance.

- **Efficiency:** Every program requires certain processing time and memory to process the instructions and data. As you must have realized, processing power and memory are the most precious resources of a computer; a program should be laid out in such a manner that it utilizes the least amount of memory and processing time.
- **Structural:** To develop a program, the task must be broken down into a number of subtasks. These subtasks are developed independently, and each subtask is able to perform the assigned job without the help of any other subtask. If a program is developed structurally, the program not only becomes more readable, but the testing and documentation processes also get easier.
- **Flexibility:** A program should be flexible enough to handle most of the changes without having to rewrite the entire program. Most of the programs are developed for a certain period and they require modifications from time to time. For example, in case of payroll management, as the time progresses, some employees may leave the company while some others may join. Hence, the payroll application should be flexible enough to incorporate all the changes without having to reconstruct the entire application.
- **Generality:** Apart from flexibility, the program should also be general. By generality, we mean that if a program is developed for a particular task then it should also be used for all similar tasks of the same domain. For example, if a program is developed for a particular organization then it should suit all the other similar organizations.
- **Documentation:** Documentation is one of the most important components of an application development. Even if a program is developed following the best programming practices, it will be rendered useless if the end user is not able to fully utilize the functionality of the application. A well-documented application is also useful for other programmers because even in the absence of the author, they can understand it.

10.9 PROGRAMMING LANGUAGES

Previously, we discussed that a computer needs to be instructed using computer programs to perform all its tasks. To write computer programs, a computer or programming language is required, as computers do not understand natural languages (like English). A *programming language* consists of a set of characters, symbols and usage rules that allow the user to communicate with computers just as natural languages are used for communication among human beings.

You might be wondering why natural languages are not used to instruct the computers. There are mainly two

reasons behind it. Firstly, the natural languages (English, Spanish) are ambiguous, vaguely structured, and have very large (and ever changing) vocabularies. In contrast, computer languages have relatively few, exactly defined rules for composition of programs, and strictly controlled vocabularies in which unknown words must be defined before they can be used. Secondly, in case of natural languages, we can understand even while using poor grammar and vocabulary. However, in the case of programming languages, the rules are very rigid and the programmer has to follow all the specified rules to create an accurate program and to obtain the desired results.

10.9.1 Types of Programming Languages

Computers understand only one language and that is binary language or the language of 0s and 1s. In the initial years of computer programming, all the instructions were given in binary form only. Although these programs were easily understood by the computer, it proved too difficult for a normal human being to remember all the instructions in the form of 0s and 1s. Therefore, the computer remained a mystery to a common person until other languages such as assembly and HLL were developed which were easier to learn and understand. These languages use commands that have some degree of similarity with English (such as “if else”, “exit”). Programming languages can be divided into three major categories:

- **Machine Language:** It is the native language of computers. It uses only 0s and 1s to represent data and instructions.
- **Assembly Language:** It correspondences symbolic instructions and executable machine codes and was created to use letters instead of 0s and 1s.
- **High-level Language:** These languages are written using a set of words and symbols following some rules similar to a natural language such as English. The programs written in HLL are known as *source programs* and these programs are converted into machine-readable form by using compilers or interpreters.

Note: *Together, machine and assembly language are also known as low-level languages.*

Since early 1950s, programming languages have evolved tremendously. This evolution has resulted in the development of hundreds of different languages. With each passing year, the languages become user-friendly and more powerful. We can illustrate the development of all the languages in five generations.

First generation languages are machine languages, in which instructions are particular sequences of 0s and 1s that digital computers can understand. Second generation languages, assembly languages, allow programmers to use meaningful abbreviations for machine-specific instructions in place of the incomprehensible 0s and 1s form. Each instruction in an assembly language translates directly to a machine language instruction using a tool called an *assembler*. Programming became somewhat easier, but many users still wanted floating-point numbers and array indexing. Since these capabilities were not supported in hardware, HLL had to be developed to support them. The next few sections discuss the five generations of languages and how they revolutionized the computer industry.

10.10.1 First Generation: Machine Language

The first language was binary, also known as *machine language*, which was used in the earliest computers and machines. We know that computers are digital devices, which have only two states, ON and OFF (1 and 0). Therefore, every instruction and data should be written using 0s and 1s. Machine language is also known as the computer's “native” language as this system of codes is directly understood by the computer.

THINGS TO REMEMBER

First Programmers

Lady Lovelace Ada Augusta (officially the first programmer) suggested binary numbers for computer storage instead of decimals. A British mathematician, Alan Mathison Turing, was the first person to recognize that programming in machine language is less time consuming. In 1952, John von Neuman proposed to have new programs loaded from a magnetic tape to read. With differences in magnetic polarities, it can mean either ON or OFF states.

Instruction in machine language consists of two parts (see Figure 10.12). The first part is an operation which tells the computer what functions are to be performed. The second part of the instruction is the operand, which tells the computer where to find or store the data on which the desired operation is to be performed.

Figure 10.12 Machine Language Instruction Format

A binary program is a long list of instructions that are executed by the CPU. Normally, instructions are executed one after the other, but program flow may be influenced by special jump instructions that transfer

execution to an instruction other than the following one. Each computer has its own set of instructions based on its architecture. Hence, machine language may differ from computer to computer.

Advantages of Machine Language: Even though machine language is not a human-friendly language, it offers the following advantages.

- **Translation Free:** Machine language is the only language that computers can directly execute without the need for conversion. Even an application using HLL has to be converted into machine-readable form so that the computer can understand the instructions.
- **High Speed:** Since no conversion is needed, the applications developed using machine language are extremely fast. It is usually used for complex applications such as space control system, nuclear reactors and chemical processing.

Disadvantages of Machine Language: There are many disadvantages in using machine language to develop programs. Some of these are as follows.

- **Machine Dependent:** Every computer type differs from the other, based on its architecture. Hence, an application developed for a particular type of computer may not run on another type of computer. This may prove costly as well as difficult for the organizations.
- **Complex Language:** Machine language is very difficult to read and write. Since all the data and instructions must be converted to binary code, it is almost impossible to remember the instructions. A programmer must specify each operation, and the specific location for each piece of data and instruction to be stored. It means that a programmer practically needs to be a hardware expert to have proper control over the machine language.
- **Error Prone:** Since the programmer has to remember all the opcodes and memory locations, machine language is bound to be error prone. It requires a superhuman effort to keep track of the logic of the problem, and therefore results in frequent programming errors.
- **Tedious:** Machine language poses real problems while Modifying and correcting a program. Sometimes the programming becomes too complex to Modify and the programmer has to rewrite the entire logic again. Therefore, it is very tedious and time consuming, and since time is a precious commodity, programming using machine language tends to be costly.

Due to its overwhelming limitations, machine language is rarely used nowadays.

10.10.2 Second Generation: Assembly Language

The complexities of machine language led to the search of another language: the assembly language, developed in the early 1950s and its main developer was IBM. However, Jack Powell, Bob Nevelen, Clement and Michael Bradly also helped in the development of the assembly language. It was a stepping-stone for all subsequent language development. Assembly language allows the programmer to interact directly with the hardware. This language assigns a mnemonic code to each machine language instruction to make it easier to remember or write. It allows better human-readable method of writing programs as compared to writing in binary bit patterns. However, unlike other programming languages, assembly language is not a single language, but a group of languages. Each processor family (and sometimes individual processors within a processor family) has its own assembly language.

An assembly language provides a mnemonic instruction, usually three letters long, corresponding to each machine instruction. The letters are usually abbreviated indicating what the instruction does. For example, ADD is used to perform an addition operation, SUB for subtraction, and so on. Assembly languages make it easier for humans to remember how to write instructions to the computer, but an assembly language is still a representation of the computer's native instruction set. Since each type of computer uses a different native instruction set, assembly languages cannot be standardized from one machine to another, and instructions from one computer cannot be expected to work on another.

The basic unit of an assembly language program is a line of code. It allows the use of symbols and set of rules that can be used and combined to form a line of code. Each line of an assembly language program consists of four

columns called *fields*. The general format of an assembly instruction is as follows.

```
[Label] <Opcode> <Operands> [; Comment]
```

[...] brackets indicate that enclosed specification may or may not appear in a statement. If a label is specified, it is associated as a symbolic name with the machine words generated for the assembly statement. If multiple operands are used, each of them is separated by a comma. The text after semicolon (;) is just comments. Comments are not a part of actual program, but are used just for reference purposes, that is, to specify what actually the statement will do. Although comments are optional, they are included to facilitate proper documentation. For example,

Label	Opcode	Operands	Comments
BEGIN	ADD	A,B	;ADD B to A

Note: *The first character of an assembly language variable should be an alphabet. The rest of the characters may be alphabets or digits. However, the total number of characters should not exceed 8.*

Assembler: This language is nothing more than a symbolic representation of machine code, which allows symbolic designation of memory locations. However, no matter how close assembly language is to machine code, the computer still cannot understand it. The assembly language program must be translated into machine code by a separate program called an *assembler*. The assembler program recognizes the character strings that make up the symbolic names of the various machine operations, and substitutes the required machine code for each instruction. At the same time, it also calculates

the required address in memory for each symbolic name of a memory location, and substitutes those addresses for the names resulting in a machine language program that can run on its own at any time. In short, an assembler converts the assembly codes into binary codes and then it assembles the machine understandable code into the main memory of the computer, making it ready for execution (see Figure 10.13).

Figure 10.13 Working of an Assembler

The original assembly language program is also known as the source code, while the final machine language program is designated the *object code*. If an assembly language program needs to be changed or corrected, it is necessary to make the changes to the source code and then re-assemble it to create a new object program. The functions of an assembler are given below.

- It translates mnemonic operation codes to machine code and corresponding register addresses to system addresses.
- It checks the syntax of the assembly program and generates diagnostic messages on syntax errors.
- It assembles all the instructions in the main memory for execution.
- In case of large assembly programs, it also provides linking facility among the subroutines.
- It facilitates the generation of output on the required output medium.

Advantages of Assembly Language: The advantages of using assembly language to develop a

program are as follows.

- **Easy to Understand and Use:** Assembly language uses mnemonics instead of using numerical opcodes and memory locations used in machine language. Hence, the programs written in assembly language are much more easier to understand and use as compared to its machine language counterpart. Being a more user-friendly language as compared to machine language, assembly programs are easier to Modify.
- **Less Error Prone:** Since mnemonic codes and symbolic addresses are used, the programmer does not have to keep track of the storage locations of the information and instructions. Hence, there are fewer errors while writing an assembly language program. Even in case of errors, assembly programs provide better facility to locate and correct them as compared to machine language programs. Moreover, assemblers also provide various mechanisms to locate the errors. For example, in case of adding two variables such as ADD A, B if the variables (A and B) are not defined in the program, the assembler will give an error indicating the same so that the programmer can easily correct the mistake.
- **Efficiency:** Assembly programs can run much faster and use less memory and other resources than a similar program written in a high-level language. Speed increment of 2–20 times faster is common, and occasionally, an increase of hundreds of times faster is also possible.
- **More Control on Hardware:** Assembly language also gives direct access to key machine features essential for implementing certain kinds of low-level routines such as an operating system kernel or micro-kernel, device drivers and machine control.

Disadvantages of Assembly Language: The disadvantages in using assembly language to develop a program are as follows.

- **Machine Dependent:** Different computer architectures have their own machine and assembly languages, which means that programs written in these languages are not portable to other, incompatible systems. This makes it a low-level language. If an assembly program is to be shifted to a different type of computer, it has to be modified to suit the new environment.
- **Harder to Learn:** The source code for an assembly language is cryptic and in a very low machine-specific form. Being a machine-dependent language, every type of computer architecture requires a different assembly language, making it hard for a programmer to remember and understand every dialect of assembly. More skilled and highly trained programmers, who know all about the logical structure of the computer, can only create applications using assembly language.
- **Slow Development Time:** Even with highly skilled programmers, assembly generated applications are slower to develop as compared to high-level-language-based applications. In the case of assembly language, the development time can be 10–100 times as compared to a high-level-language-generated application.
- **Less Efficient:** A program written in assembly language is less efficient than machine language because every assembly instruction

has to be converted into machine language. Therefore, the execution of assembly language program takes more time than a machine language program. Moreover, before executing an assembly program, the assembler has to be loaded in the computer's memory for translation and occupies a sizeable memory.

- **No Standardization:** Assembly languages cannot be standardized because each type of computer has a different instruction set, and therefore, a different assembly language.
- **No Support for Modern Software Engineering Technology:** Assembly languages provide no inherent support for software engineering technology. They work with just machine-level specifics, not with abstractions. Assembly language does not provide inherent support for safety-critical systems. It provides very little opportunity for reuse and there is no OOP support. There is also no specific support for distributed systems. The tools available for working with assembly languages are typically very low-level tools.

10.10.3 Third Generation: High-level Language

During 1960s, computers started to gain popularity and it became necessary to develop languages that were more like natural languages, such as English, so that a common user could use the computer efficiently. Since assembly language required deep knowledge of computer architecture, it demanded programming as well as hardware skills to use computers. Due to the computer's widespread usage, the early 1960s saw the emergence of the third generation programming languages (3GL). Languages such as COBOL, FORTRAN, BASIC and C are examples of 3GLs and are considered HLL.

HLL are similar to the English language. Programs written using these languages can be machine independent. A single high-level statement can substitute several instructions in machine or assembly language. Unlike assembly and machine programs, high-level programs may be used with different types of computers with little or no modification, thus reducing the re-programming time.

In a high-level language, programs are written in a sequence of statements to solve a problem. For example, the following BASIC code snippet will calculate the sum of two numbers.

```
LET X = 10
LET Y = 20
LET SUM = X + Y
PRINT SUM
```

The first two statements store 10 in variable X (memory location name) and 20 in variable Y, respectively. The third statement again creates a variable named SUM, which will store the summation of X and Y value. Finally, the output is printed, that is, the value stored in SUM is printed on the screen.

THINGS TO REMEMBER

Compilers

Compiler is a program which translates source code written in a particular programming language into computer-readable machine code that can be directly loaded and executed. For each high-level language, a separate compiler is required. For example, a compiler for C language cannot translate a program written in FORTRAN. Hence, to execute both language programs, the host computer must have the compilers of both languages.

Translating High-level Language into Machine Language: Since computers understand only machine language, it is necessary to convert the high-level language programs into machine language codes. This is achieved by using language translators or language processors, generally known as compilers, interpreters or other routines that accept statements in one language and produces equivalent statements in another language (see Figure 10.14).

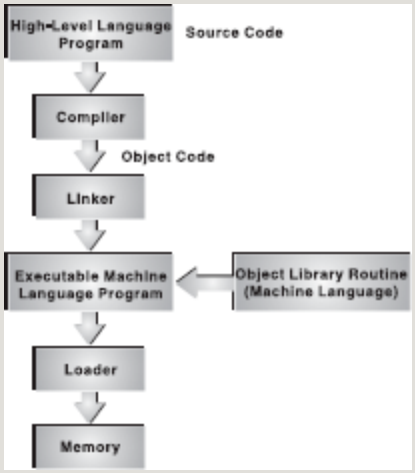


Figure 10.14 Program Translation Hierarchy

Compiler A compiler is a kind of translator that translates a program into another program, known as *target language*. Usually, the term compiler is used for language translator of high-level language into machine language. The compiler replaces a single high-level statement with a series of machine language instructions. A compiler usually resides on a disk or other storage media. When a program is to be compiled, its compiler is loaded into main memory. The compiler stores the entire high-level program, scans it and translates the whole program into an equivalent machine language program. During the translation process, the compiler reads the source program and checks the syntax (grammatical) errors. If there is any error, the compiler generates an error message, which is usually displayed on the screen. In case of errors, the compiler will not create the object code until all the errors are rectified.

Once the program has been compiled, the resulting machine code is saved in an executable file, which can be run on its own at any time. To be precise, once the executable is generated, there is no need for the actual

source code file. Anyway, it is worthwhile to keep the source file(s) because if the source code is modified, it is necessary to recompile the program again to regenerate the executable file containing amendments.

Interpreter Unlike compilers, an interpreter translates a statement in a program and executes the statement immediately, before translating the next source language statement (see Figure 10.15). When an error is encountered in the program, the execution of the program is halted and an error message is displayed. Similar to compilers, every interpreted language such as BASIC and LISP has its own interpreters.

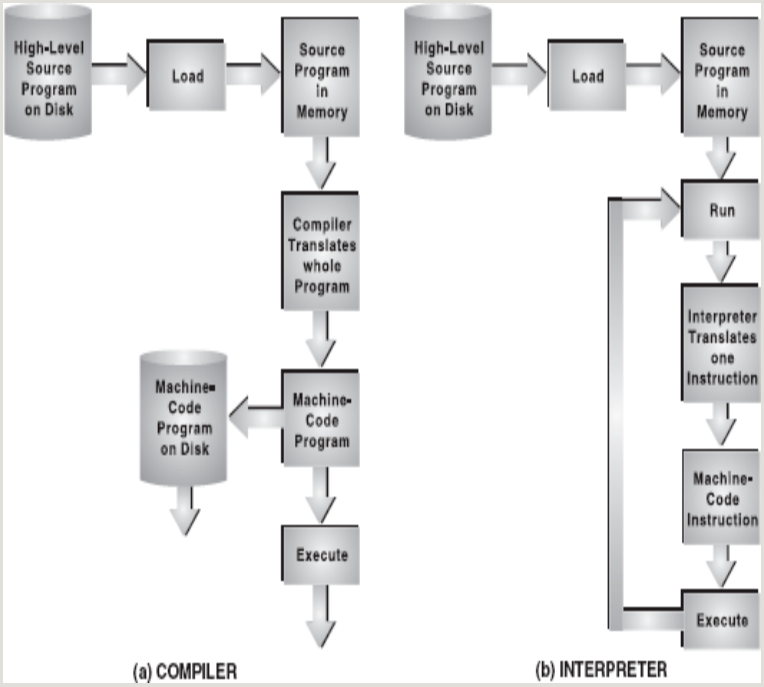


Figure 10.15 (a) Working of a Compiler (b) Working of an Interpreter

Linker: An application usually consists of hundreds or thousands of lines of codes. The codes are divided into logical groups and stored in different modules so that the debugging and maintenance of the codes becomes easier. Hence, for an application, it is always advisable to adhere to structural (modular) programming practices. When a program is broken into several modules, each module can be modified and compiled independently. In such a case, these modules have to be linked together to create a complete application. This job is done by a tool known as *linker*. A linker is a program that links several object modules and libraries to form a single, coherent program (executable). Object modules are the machine code output from an assembler or compiler and contain executable machine code and data, together with information that allows the linker to combine the modules together to form a program.

Generally, all HLL use some in-built functions like calculating square roots, finding logarithm values and so on. These functions are usually provided by the language itself, the programmer does not need to code them separately. During the program execution process, when a program invokes any in-built function, the linker transfers the control to that program where the function is defined, by making the addresses of these functions known to the calling program.

Loader: *Loaders* are a part of the operating system that brings an executable file residing on disk into memory and starts its execution. It is responsible for loading, linking and relocation. In computing, a loader is a program that performs the functions of a linker and then immediately schedules the executable code for execution, without necessarily creating an executable file as an output. A loader performs four basic tasks as follows.

1. **Allocation:** It allocates memory space for the programs.

2. **Linking:** It combines two or more separate object programs and supplies the information needed to allow references between them.
3. **Relocation:** It prepares a program to execute properly from its storage area.
4. **Loading:** It places data and machine instructions into the memory.

There are two types of loaders.

- **Absolute Loader:** It loads the file into memory at the location specified by the beginning portion (header) of the file and then passes control to the program. If the memory space specified by the header is currently in use, execution cannot proceed, and the user must wait until the requested memory becomes free. Moreover, this type of loader performs only the loading function. It does not perform linking and program relocation.
- **Relocating Loader:** This loader loads the program in memory, altering the various addresses as required to ensure correct referencing. The decision as to where in memory the program is placed, is made by the operating system, not the file's header. It is a more efficient loader, but there is a slight overhead in terms of a small delay whilst all the relative offsets are calculated. The relocating loader can only relocate code that has been produced by a linker capable of producing relative code.

Advantages of High-level Languages: HLL are useful in developing complex software, as they support complex data structures. It increases the programmer's productivity (the number of lines of code generated per hour). Unlike assembly language, the programmer does not need to learn the instruction set of each computer being worked with. The various advantages of using HLL are discussed below.

- **Readability:** Since HLL are closer to natural languages, they are easier to learn and understand. In addition, a programmer does not need to be aware of computer architecture; even a lay person can use it without much difficulty. This is the main reason of HLL's popularity.
- **Machine Independent:** HLL are machine independent in the sense that a program created using HLL can be used on different platforms with very little or no change at all.
- **Easy Debugging:** HLL include the support for ideas of abstraction so that programmers can concentrate on finding the solution to the problem rapidly, rather than on low-level details of data representation, which results in fewer errors. Moreover, the compilers and interpreters are designed in such a way that they detect and point out the errors instantaneously.
- **Easier to Maintain:** As compared to low-level languages, the programs written in HLL are easy to Modify and maintain because HLL programs are easier to understand.

- **Low Development Cost:** HLL permit faster development of programs. Although a high-level program may not be as efficient as an equivalent low-level program, but the savings in programmer's time generally outweighs the ineffciencies of the application. This is because the cost of writing a program is nearly constant for each line of code, regardless of the language. Thus, an HLL, where each line of code translates to 10 machine instructions, costs only a fraction as compared to the program developed in a low-level language.
- **Easy Documentation:** Since the statements written in HLL are similar to natural languages, they are easier to understand as compared to low-level languages.

Disadvantages of High-level Languages: The main disadvantages of this language are as follows.

- **Poor Control on Hardware:** HLL are developed to ease the pressure on programmers so that they do not have to know the intricacies of hardware. As a result, sometimes the applications written in HLL cannot completely harness the total power available at the hardware level.
- **Less Efficient:** The HLL applications are less efficient as far as computation time is concerned. This is because, unlike low-level languages, HLL must be created and sent through another processing program known as a compiler. This process of translation increases the execution time of an application. Programs written in HLL take more time to execute, and require more memory space. Hence, critical applications are generally written in low-level languages.

10.10.4 Fourth Generation

Fourth generation languages (4GLs) have simple, English-like syntax rules, commonly used to access databases. The 3GLs are considered as procedural languages because the programmer must list each step and must use logical control structures to indicate the order in which instructions are to be executed. 4GLs, on the other hand, are non-procedural languages. The non-procedural method is simply to state the needed output instead of specifying each step one after another to perform a task. In other words, the computer is instructed *what* it must do rather than *how* a computer must perform a task.

The non-procedural method is easier to write, but has less control over how each task is actually performed. When using non-procedural languages, the methods used and the order in which each task is carried out is

left to the language itself; the user does not have any control over it. In addition, 4GLs sacrifice computer efficiency in order to make programs easier to write. Hence, they require more computer power and processing time. However, with the increase in power and speed of hardware and with diminishing costs, the uses of 4GLs have spread.

4GLs have a minimum number of syntax rules. Hence, common people can also use such languages to write application programs. This saves time and allows professional programmers for more complex tasks. The 4GLs are divided into three categories:

- **Query Languages:** They allow the user to retrieve information from databases by following simple syntax rules. For example, the database may be requested to locate details of all employees drawing a salary of more than $10,000. Structured Query Language (SQL) and IBM's Query-By-Example (QBE) are examples of query languages.
- **Report Generators:** They produce customized reports using data stored in a database. The user specifies the data to be in the report, the report's format, and whether any subtotals and totals are needed. Often report specifications are selected from pull-down menus, making report generators very easy to use. Examples of report generators are Easytrieve Plus by Pansophic and R&R Relational Report Writer by Concentric Data Systems.
- **Application Generators:** With application generators, the user writes programs to allow data to be entered into the database. The program prompts the user to enter the needed data. It also checks the data for validity. Cincom System's MANTIS and ADS by Cullinet are examples of application generators.

Advantages of 4GLs: The main advantage of 4GLs is that a user can create an application in a much shorter time for development and debugging than with other programming languages. The programmer is only interested in what has to be done and that too at a very high level. Being non-procedural in nature, it does not require the programmers to provide the logic to perform a task. As a result, lot of programming effort is saved. Use of procedural templates and data dictionaries allow automatic type checking (for the programmer and for user input) and this results in fewer errors. Using application generators, the routine tasks are automated.

Disadvantages of 4GLs: Since programs written in a 4GL are quite lengthy, they need more disk space and a large memory capacity as compared to 3GLs. These languages are inflexible also because the programmers control over language and resources is limited as compared to other languages. These languages cannot directly utilize the computer power available at the hardware level as compared to other levels of languages.

10.10.5 Fifth Generation: Very High-level Languages

Fifth generation languages are just the conceptual view of what might be the future of programming languages. These languages will be able to process natural languages. The computers would be able to accept, interpret and execute instructions in the native or natural language of the end users. The users will be free from learning any programming language to communicate with the computers. The programmers may simply type the instruction or simply tell the computer by way of microphones what it needs to do. Since these languages are still in their infancy, only a few are currently commercially available. They are closely linked to artificial intelligence and expert systems.

10.11 FEATURES OF A GOOD PROGRAMMING LANGUAGE

The features of one programming language may differ from the other. One can be easy and simple while others can be difficult and complex. We judge the success and strength of a programming language with respect to standard features. To begin the language selection process, it is important to establish some criteria for what makes a language good.

Ease of Use: The language should be easy in writing codes for the programs and executing them. The ease and clarity of a language depends upon its syntax. It should be capable enough to provide clear, simple and unified set of concepts. The vocabulary of the language

should resemble English (or some other natural language). Symbols, abbreviations and jargon should be avoided unless they are already known to most people. Any concept that cannot easily be explained to amateurs should not be included in the language.

Simplicity helps in the readability of the language. A simple language is easier to grasp and code. Developing and implementing a compiler or interpreter is also easier for simple languages as compared to complex ones.

Portability: The language should support the construction of code in a way that it could be distributed across multiple platforms (operating systems). Computer languages should be independent of any particular hardware or operating system, that is, programs written on one platform should be able to be tested or transferred to any other computer or platform and there it should perform accurately.

Naturalness for the Application: The language should have a syntax which allows the program structure to show the underlying logical structure of algorithm. A programming language should provide a conceptual framework for thinking through algorithms and means of expressing those algorithms through flowcharts. The advantage of displaying a program in algorithms and flowcharts is that even a novice can understand them easily without learning any programming language.

Reliability: The language should support construction of components that can be expected to perform their intended functions in a satisfactory manner throughout its lifetime. Reliability is concerned with making a system failure free, and thus is concerned with all possible errors. The language should provide support for preventing errors. For example, some errors can be prevented by strict syntax checking. Apart from prevention, the language should also be able to detect

and report errors in the program. For example, errors such as arithmetic overflow and assertions should be detected properly and reported to the programmers immediately so that the error can be rectified. The language should provide reliability by supporting explicit mechanisms for dealing with problems that are detected when the system is in operation (exception handling).

Safety: It is concerned with the extent to which the language supports the construction of safety-critical systems, yielding systems that are fault-tolerant, fail-safe, or robust in the face of systemic failures. The system must always do what is expected and be able to recover from any situation that might lead to a mishap or actual system hazard. Thus, safety tries to ensure that any failure that results in minor consequences, and even potentially dangerous failures are handled in a fail-safe fashion. Language can facilitate this through such features as built-in consistency checking and exceptional handling.

Performance: By performance, we mean that the language should not only be capable of interacting with the end users, but also with the hardware. The language should also support software engineering mechanism, discouraging or prohibiting poor practices and supporting maintenance activities. Nowadays, the hardware has become very sophisticated and quiet fast. Hence, the application developed using a good language should tap the maximum resources of the available hardware power in terms of speed and memory efficiency.

Cost: Cost component is a primary concern before deploying a language at a commercial level. It includes several costs such as the following.

- Program execution and translation cost.
- Program creation, testing and usage cost.

- Program maintenance cost.

Promote Structured Programming: We know that structured programming is the best way to create an application. Hence, a good language should be capable of supporting structured programming. Since, by nature, structured programming facilitates ease in understanding the code, a program is easier to create, debug and maintain. A structured program also helps programmers to visualize the problem in a logical way, thereby reducing the probability of errors in the code.

Compact Code: A language should promote compact coding, that is, the intended operations should be coded in a minimum number of lines. Even if the language is powerful, and is not able to perform the tasks in a small amount of code, then it is of no use. The main reason behind it is that a large code requires more testing and developing time, thereby increasing the cost of developing an application.

Maintainability: Creating an application is not the end of the system development. It should be maintained regularly so that it can be easily modified to satisfy new requirements or to correct deficiencies. Maintainability is actually facilitated by most of the languages, which makes it easier to understand and then change the software. Maintainability is closely linked with the structure of the code. If the original code is written in an organized way (structural programming) then it would be easy to Modify or add new changes.

Reusability: The language should facilitate the adaptation of code for use in other applications. Code is reusable when it is independent of other codes. It is very common, for example, to reuse common data structures such as stacks, queues and trees. When these have been defined with common operations on the structures, these abstract data types are easy to reuse.

Provides Interface to Other Language: From the perspective of the language, interface to other language refers to the extent to which the selected language supports the interfacing feature to other languages. This type of support can have a significant impact on the reliability of the data, which is exchanged between two applications, developed with different languages. In case of data exchange between units of different languages, without specific language support, no checking may be done on the data or even on their existence. Hence, the potential for unreliability becomes high.

Concurrency Support: Concurrency support refers to the extent to which inherent language supports the construction of code with multiple threads of control (also known as *parallel processing*). For some applications, multiple threads of control are very useful or even necessary. This is particularly true for real-time systems and those running on architecture with multiple processors. Although, concurrency is rarely supported by a language directly, language support can make concurrent processing more straightforward and understandable, and it can also provide the programmer with more control over its implementation.

Standardization: Standardization means the extent to which the language definition has been formally standardized (by recognized bodies such as ANSI and ISO) and the extent to which it can be reasonably expected that this standard will be followed in a language translator. Non-standard languages may soon become obsolete, rendering it inferior and difficult to use, producing inferior code, which is difficult to maintain. Lack of appropriate support compromises developer productivity and system quality. It also compromises the ease of development, as well as performance and reliability of the code. If the reliability of the code is compromised, not only will the system perform below

expectations, but it will also become much more costly during its lifetime.

LET US SUMMARIZE

1. A computer needs to be instructed to perform all its tasks. These instructions are provided in the form of a *computer program*. Programming involves many stages such as problem definition, task analysis, developing algorithm, testing algorithm, coding, program testing, documentation, implementation, and maintenance and enhancement.
2. An *algorithm* is defined as a finite sequence of explicit instructions that when provided with a set of input values produces an output and then terminates. Algorithms can be represented as flowchart, pseudocode or decision table.
3. A *flowchart* is a pictorial representation of an algorithm in which the steps are drawn in the form of different shapes of boxes and the logical flow is indicated by interconnecting arrows. The boxes represent operations and the arrows represent the sequence in which the operations are implemented.
4. A *decision table* is a way of representing the problem statement in a tabular form. It contains all the possible conditions for a specific problem and the corresponding results using condition rules that connect conditions with results.
5. *Pseudocode* is a generic way of describing an algorithm without the use of any specific programming language-related notations. It is an outline of a program, written in a form, which can easily be converted into real programming statements. Pseudocode uses plain English statements rather than symbols to represent the processes in a computer program.
6. Program statements that affect the order in which statements are executed, or that affect whether statements are executed, are called *control structures*. They affect the flow of simulation code since a control structure evaluates statements and then executes code according to the result.
7. There are three control structures, which are used in the case of designing a solution for a problem. These are *sequential*, *selection* and *repetition*. A *sequence* structure is a linear progression where one task is performed sequentially after another. A *selection* structure allows the program to make a choice between two alternate paths depending upon whether a conditional statement is true or false. A *repetition* or *loop* logic causes an interruption in the normal sequence of processing and directs the system to move back to a previous statement in the program, repeating the same sequence again.
8. *Programming paradigm* refers to how a program is written in order to solve a problem. It provides (and determines) the view that the programmer has of the execution of the program. Broadly, programming can be classified in the following three categories: *unstructured programming, structured programming* and *object-oriented programming*.
9. *Unstructured programming* refers to writing small and simple programs consisting of only one main program. All the actions such as providing input, processing and displaying output are done within one program only.

10. Using *structured programming*, a program is broken down into small independent tasks, each having a single point of entry and exit. Each task has its own functionality and performs a specific part of the actual processing.
11. In *object-oriented programming*, the programs are organized as cooperative collections of objects, each of which represents an instance of some *class*, and whose classes are members of a hierarchy of classes united by way of an inheritance relationship.
12. *Programming language* is a language that a user employs to interact with the computer. It can be divided into three major categories: *machine, assembly* and *high-level language*.
13. First generation language is the *machine language* in which instructions are in the form of 0s and 1s. Second generation language is the *assembly language* in which mnemonic code is assigned to each machine language instruction to make it easier to remember and write. 3GLs are machine independent, use HLL to write programs, and thus need language translators to translate the high-level code into machine code. 4GLs, also called as *non-procedural languages*, use minimum syntax rules and are categorized into *query languages, report generators* and *application generators*. Fifth generation language is in the development stage. The computer will be able to accept, interpret and execute the instruction in the natural language of the user.
14. *Linker* is a program that links together several object modules and libraries to form a single, coherent executable program. *Loader* is a program that brings an executable file residing on a disk into a memory. It is also responsible for performing basic tasks such as linking, loading and relocation.
15. A good programming language should be easier to use and be portable. It should incur the minimum cost and provide concurrent support.

EXERCISES

Fill in the Blanks

1. The sequence of instructions that instruct the computer to carry out a specific task is called.....
2. An algorithm can be represented as.....or.....
3. Broadly, programming can be classified into.....,.....and.....approaches.
4. A program that translates the assembly language code into machine language code is called.....
5. The part of operating system that brings an executable file residing on disk into memory and start running is.....
6. Compiler and interpreter are also called.....
7. The two-part forms of machine language instructions are.....and.....
8. The two categories of loader are.....and.....
9. .....language is very close to the hardware.
10. COBOL, FORTRAN and C are examples of.....

Multiple-choice Questions

1. The semantic and syntax errors in the program are checked in.....
    1. Coding Phase
    2. Testing Phase
    3. Implementation Phase
    4. Analysis Phase
2. Choose the odd one out.....
    1. Abstraction
    2. Encapsulation
    3. Inheritance
    4. Linking
3. Flow lines in a flowchart are used to connect.....
    1. Two Terminals
    2. Two Connectors
    3. Two Input/Output
    4. All of these
4. Programming tools helps in.....
    1. Producing readable code
    2. Making the program logic easier
    3. Reducing the size of programs
    4. None of these
5. A high-level language is a programming language that is.....
    1. Easier to use
    2. Considered a first generation language
    3. Considered a third generation language
    4. Both (a) and (c)
6. What consists of instructions written as a series of 1s and 0s?
    1. Query Language
    2. Assembly Language
    3. Machine Language
    4. None of these
7. Which of the following was a disadvantage of the first generation language?
    1. Machine dependency
    2. Less efficient
    3. Need for more disk space
    4. All of these
8. Good programming language has the following features:.....
    1. Portable, Safe
    2. Reliable, Good Performance
    3. High Cost
    4. Only (a) and (b)
9. Query languages comes under.....
    1. Third Generation
    2. Fourth Generation
    3. Fifth Generation
    4. None of these
10. Which of the following program control structure allows the program to make a choice between two alternate paths depending upon the condition?
    1. Sequence

2. Loop
3. Selection
4. None of these

State True or False

1. The graphical representation of an algorithm is called flowchart.
2. The generic way of describing an algorithm without the use of programming language is known as program.
3. The keyword, DETERMINE, used in a pseudocode is an input operation.
4. The terminal symbol in a flowchart represents start of program logic.
5. Programming is the art and science of creating computer programs.
6. A program that links several object modules and libraries to form a single, coherent program is called linker.
7. The language translator that translates the programs written in high-level language into its corresponding machine code is called assembler.
8. Machine language is also called low-level language.
9. Pseudocode can be compiled by a program compiler.
10. The diamond-shape symbol in the flowchart signifies processing.

Descriptive Questions

1. Explain the program development cycle with the help of a block diagram.
2. Define an algorithm. List the characteristics of a good algorithm.
3. Define a flowchart. List some important reasons for using flowcharts.
4. What benefits do decision tables offer over flowcharts?
5. What are the advantages and disadvantages of using a pseudocode?
6. Discuss the three basic program control structures with suitable examples.
7. Explain the features of a good programming language.
8. Describe the classification of programming languages.
9. Write short notes on the following.

    1. Compiler
    2. Interpreter
    3. Loader
    4. Linker

ANSWERS

Fill in the Blanks

1. Computer program
2. Flowchart, Decision table, Pseudocode
3. Unstructured, Structured, Object-oriented
4. Assembler
5. Loader
6. Language translators
7. Opcode, Operand

8. Absolute loader, Relocating loader
9. Machine language
10. Third generation languages

Multiple-choice Questions

1. (b)
2. (d)
3. (d)
4. (b)
5. (d)
6. (c)
7. (a)
8. (d)
9. (b)
10. (c)

State True or False

1. True
2. False
3. False
4. True
5. True
6. True
7. False
8. True
9. False
10. False

Chapter 11

Computer Sofware

This chapter introduces the reader to the essentials of computer software. Computer software refers to computer program(s)that causes a computer to carry out particular operations. The chapter starts with the definition of software, followed by its relationship with the hardware. The discussion continues with the types of software, namely, system software and application software, their roles along with the hardware in building up a fully functional computer system. In the later part of the chapter, piracy and techniques of its prevention are explained. The chapter concludes by providing an overview of software terminology.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The computer system, which consists of hardware and software
- Software, which is defined as collection of data and instructions in a computer
- The categories of software, namely system software and application software
- The process of installing and uninstalling software
- Software piracy—an unauthorized way of copying an internally developed software

11.1 INTRODUCTION

A computer system consists of *hardware*, the electronic devices that are capable of computing and manipulating information, and *software* (set of instructions) that carries out predefined tasks to complete a given job. As we know, a computer cannot think or perform on its own. It performs operations like addition, subtraction, multiplication, and division only when the user instructs it to do so. The user issues instructions and the CPU acts in accordance with the instructions. The sets of instructions, which control the sequence of operations, are known as *programs*, and collectively programs are called *software*. It is an intangible commodity, that is, the part of a computer system that users cannot touch.

We can equate hardware and software with human body and human intelligence, respectively. All human physical actions such as walking and eating are based on the thoughts and feelings, which is raised by the brain. If the brain does not raise thoughts and feelings, we do not perform any physical activity. Similarly, the actions and functioning of every hardware equipment are driven by

software. The combination of physical equipment (hardware) and logical instructions (software) gives modern computing systems their power and versatility.

11.2 SOFTWARE: DEFINITION

Software is a generic term for organized collection of computer data and instructions. It is responsible for controlling, integrating, and managing the hardware components of a computer and to accomplish specific tasks. In other words, a software tells the computer what to do and how to do it. For example, software instructs the hardware what to display on the user's screen, what kinds of input to take from the user, and what kinds of output to generate. Thus, software communicates with the hardware by organizing the control sequences, and the hardware carries out the instructions defined by the software.

As discussed earlier, a computer needs to be instructed to perform any task. These instructions are given in the form of computer programs, which are written in computer programming languages. A program controls the activity of the processor. The moment the hardware (processor, memory, etc.) acts as per the instructions of a program, the program is said to be in *running or executing* state.

A set of programs, which are specifically written to provide the user a precise functionality like solving a specific problem, is termed as a *software package*. For example, word-processing software package provides functionality to the computer so that it can be used to create text documents like letters and mailing lists. Similarly, an image-processing software package assists a user in drawing and manipulating graphics.

11.2.1 Relationship between Software and Hardware

Software refers to the computer programs that are loaded into a computer system, and hardware refers to all the visible devices, which are assembled together to build a computer system (see Figure 11.1). Both software and hardware go hand in hand; you cannot have one without the other. Even though hardware is the physical part of a computer, it is nothing unless it has software to control it. In a way, hardware is like a car without a driver; one needs both to get the work done. Software is a set of instructions that tells the hardware *what* to do and *how* to perform the requested actions. Thus, hardware and software share a special relationship. If hardware is the "heart" of a computer system, software is its "soul." Both are complementary to each other.

Figure 11.1 Hardware and Software

An analogy can be taken of a video game system, which comprises a console, games cassettes, joystick, and display screen as the hardware. The games in the cassettes can be considered as the software. To play a particular game, the cassette of that game has to be

loaded on the console and then the game can be played. Similarly, to get a particular job done by a computer, the relevant software is loaded in the storage device, which makes a computer perform the desired functions. Therefore, it is evident that the software is vital. Another inference from this analogy is that different software can be used on the same hardware to perform different jobs, just as different games can be played on the same console by using different cassettes.

11.3 CATEGORIES OF SOFTWARE

Software can be categorized as *system software* and *application software* (see Figure 11.2). System software is a generic term for referring to any computer program whose purpose is to help the user to run the computer system, whereas application software employs the capabilities of a computer directly to a task that the user wishes to perform. As an analogy, we can equate an electric bulb to an application and the electric power generation plant to a system. The power plant merely generates electricity. It is not really of any use until harnessed to an application like the electric bulb, which performs a service that the user desires.

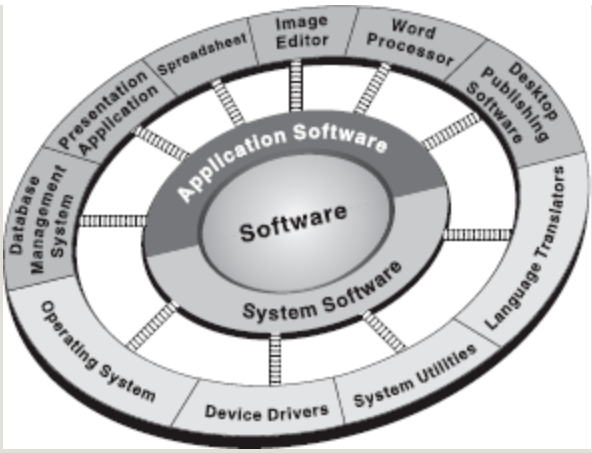


Figure 11.2 Software Categories

11.3.1 System Software

System software consists of several programs, which are directly responsible for controlling, integrating, and managing the individual hardware components of a computer system. You must have noticed that a new computer system is always accompanied with some software stored in either a CD or a DVD, which is supplied by the manufacturer. This software manages and supports the computer system and its information processing activities.

System software is more transparent and less noticed by the users as it usually interacts with the hardware or the applications. This software provides a programming environment in which programmers can create applications to accommodate their needs. This environment provides functions that are not available at the hardware level and performs the tasks related to the execution of an application program. Hence, system software acts as an interface between the hardware of the computer and the software applications (see Figure 11.3).

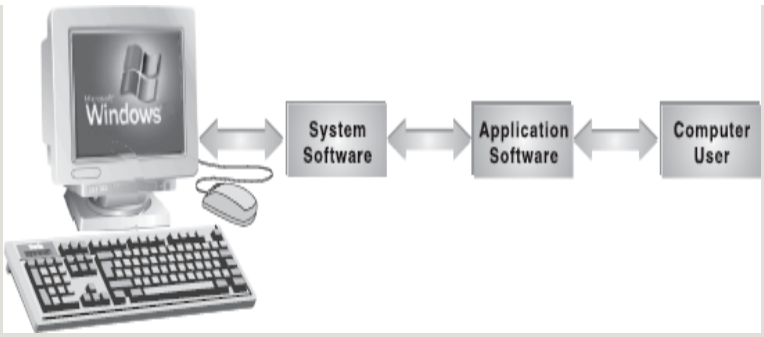


Figure 11.3 System Software, Application Software, and Hardware

In simple terms, system software makes the computer functional. It provides basic functionality like file management, visual display, keyboard input, etc., and are used by application software to accomplish these functions. Some examples of system software are *operating systems, device drivers, language translators*, and *system utilities*.

Operating System: It is the first layer of software loaded into the computer memory when it starts up. As the first software layer, all other software that gets loaded after it depends on it for various common core services. These common core services include disk access, memory management, task scheduling, and user interfacing. In addition, the operating system ensures that different programs executing at the same time do not interfere with each other. It provides a software platform on top of which other programs can run. In simple words, the operating system organizes and controls the hardware. Examples of operating systems are Windows XP, UNIX, and Linux. The basic functions of an operating system are (see Figure 11.4 also):

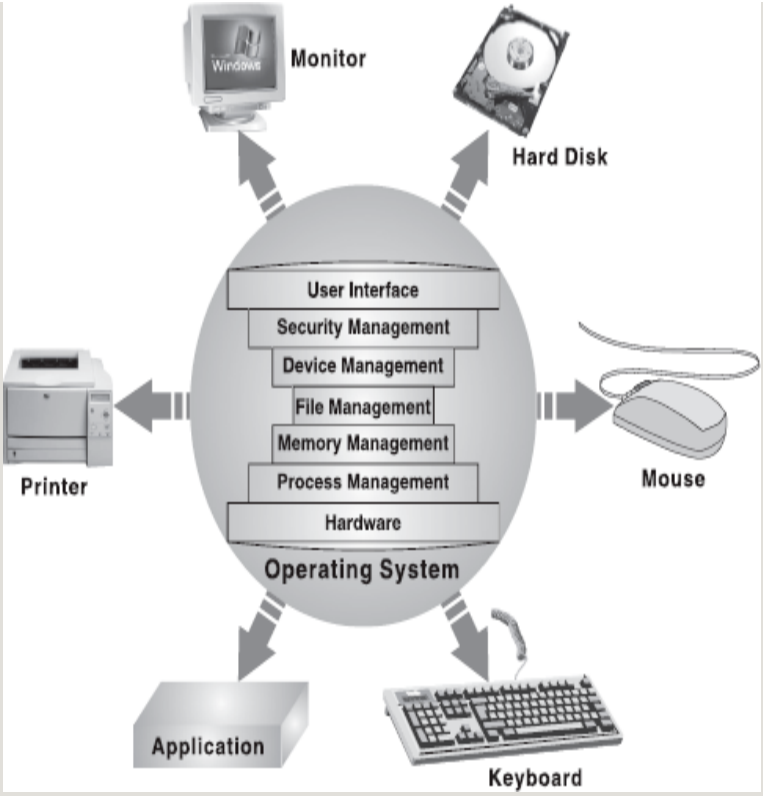


Figure 11.4 Operating System

- **Process Management:** It handles the creation, deletion, suspension, resumption, and synchronization of processes.
- **Memory Management:** It handles allocation and de-allocation of memory space as required by various programs.
- **File Management:** It is responsible for the creation and deletion of files and directories. It also organizes, stores, retrieves, names, and protects all the files.
- **Device Management:** It manages all the devices of the computer system such as printers and modems. If any device fails, it detects the device failure and notifies the same to the user.
- **Security Management:** It protects system resources and information against destruction and unauthorized use.
- **User Interface:** It provides the interface between the user and the hardware.

Device Drivers: Device drivers are system programs, which are responsible for proper functioning of devices. Every device, whether it is a printer, monitor, mouse, or

keyboard, has a driver program associated with it for its proper functioning. Whenever a new device is added to the computer system, a new device driver must be installed before the device is used. In modern operating systems, most hardware drivers, such as the keyboard driver, come with the operating system. A driver acts like a translator between the device and program (typically, an operating system) that uses the device. For example, when a user prints a document, the processor issues a set of generic commands to the printer driver, and the driver translates those commands into specialized instructions that the printer understands. Note that each device has its own set of specialized commands that only its driver understands. A device driver is not an independent program; it assists and is assisted by the operating system for the proper functioning of the device (see Figure 11.5).

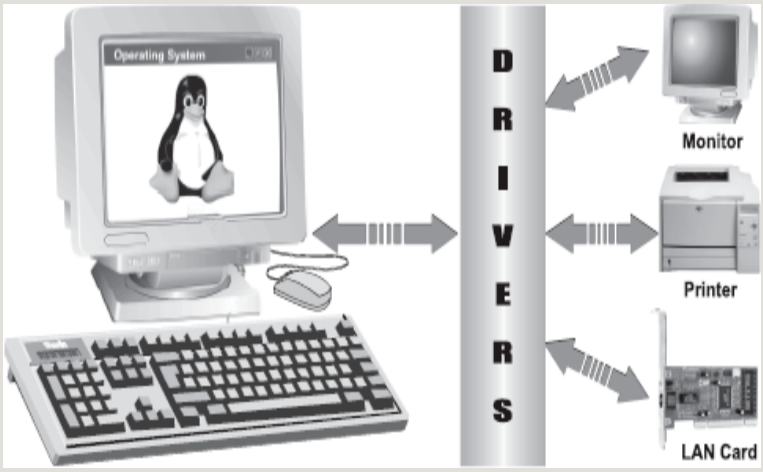


Figure 11.5 Device Drivers

Language Translators: Computers only understand *machine language* that is a language consisting of 0's and 1's. However, it is difficult for a programmer to understand the machine language. To ease the burden of

a programmer, special programming languages called *high-level programming languages* were developed that resemble natural languages like English. Since computer understands only machine language, a tool is required which could translate a program written in a high-level programming language to machine language. Hence with every programming language, a language translator is also developed, which accepts the programs written in the programming language and transforms them into a form suitable for execution (see Figure 11.6). In other words, language translators help in converting programming languages into machine language. To be precise, they convert programming statements into the 0's and 1's that the computer is able to process. Language translators are divided into three major categories: *compiler, interpreter,* and *assembler*, which have been discussed in Chapter 10, "Computer Programming and Languages."

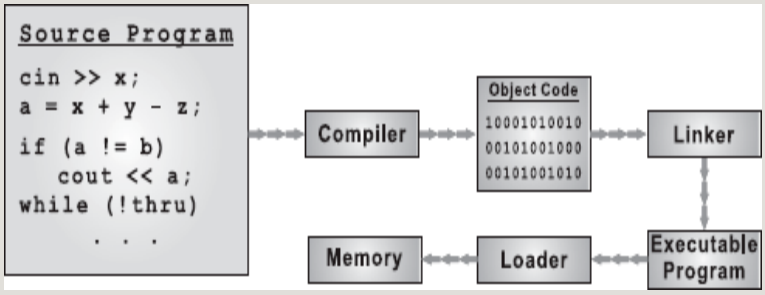


Figure 11.6 Translating Source Code into Executable Code

Typical software generally comprises millions of lines of programming statements or code. The code is divided into logical groups and stored in different independent modules so that debugging and maintenance of the code becomes easier. Before execution, different object codes

resulting from the independent modules have to be linked together to create an executable program. A *linker* is a system program that links together several objects, modules, and libraries to form a single, coherent program (executable program). The part of the operating system that brings an executable file residing on disk into memory and executes it is known as *loader*.

System Utility: System utility programs perform day-to-day tasks related to the maintenance of the computer system. They are used to support, enhance, and secure existing programs and data in the computer system. They are generally small programs, having specific tasks to perform. Some utility programs are usually provided along with the operating system, some are free while some need to be purchased from the third-party commercial vendors. Most common functions of system utilities include:

- **File Management:** These utilities make it easier to manage data files. Many programs are written to help users to find the files, create and organize directories, copy, move, and remove files. For example, the Windows Explorer in Microsoft Windows operating system does all the said activities in a user-friendly interface.
- **Backup:** Sometimes, it may happen that some data files are corrupted or accidentally deleted. In such a case, data backups become very useful. A backup system utility is essential for those organizations that want to keep their data intact.
- **Data Recovery:** It is the process of retrieving deleted or inaccessible data from failed electronic storage media such as computer hard disk drives, removable media, optical devices, and tape cartridges. The data might become inaccessible due to a software problem, computer virus, mechanical or electrical malfunction, or a deliberate human act. Using data recovery tools, an experienced technician can successfully recover 80 to even 100 per cent of lost data.
- **Virus Protection:** Anti-virus programs are essential system utilities for a computer system functioning in a network. They provide security to the system from viruses that can damage the computer system. Viruses are small programs written with malicious intent, which copy themselves to hard disk from the Internet or other infected systems. Viruses keep on spreading to other computers through the network or exchange of infected storage devices such as floppies and CDs. To protect system from viruses, one needs anti-virus software. Once installed on the system, anti-virus software scans the hard disk for any kind of virus and, if found, removes them. In addition, it continuously monitors the computer and prevents viruses from getting inside.

Examples of some anti-virus programs are Norton anti-virus and McAfee anti-virus.

- **Disk Management:** Disk management programs include various software like *disk defragmenter, data compressor,* and *disk formatting tools*. De-fragmentation implies putting fragments of files in a sequential order onto the disk, which reduces the time to access the file. It recognizes the data stored on the disk so that it can be arranged more effectively. Data compression programs squeeze out the slack space generated by the formatting schemes. Formatting tools format the hard drive in tracks and sectors for orderly storing of data in the disk.
- **Firewalls:** Firewalls form a barrier between networked computers within an organization and those outside the organization. They are commonly used to protect information such as e-mail and data files within a physical building or an organization. Essentially, a firewall is designed to protect a computer from unauthorized access, especially via network.
- **Disk Cleanup:** To keep a computer running smoothly, its regular maintenance is vital. Therefore, one should use the disk cleanup utility, which easily determines which files on the hard drive are no longer needed, and delete those files. In addition to freeing up potentially significant hard disk space; using disk cleanup on a regular basis can significantly improve the system performance.

11.3.2 Application Software

The most often seen software by a user is the application software. It is used to accomplish specific tasks rather than just managing a computer system. For a user, the computer system has no specific use without the application software. Application software may consist of a single program, such as Microsoft's Notepad (for writing and editing simple text), or a collection of programs, which work together to accomplish a task such as database management software. Application software may also include a larger collection of related but independent programs and packages (a software suite), which have a common user interface or shared data format, such as Microsoft Office suite.

Application software is dependent on system software. A system software (like operating system) acts as an interface between the user and the computer hardware, while application software performs specific tasks. Applications are softwares that perform tasks for the user besides helping the computer operate, which is the

task of system software. Application software is controlled by system software, which manages hardware devices and performs background tasks for them. The distinction between the two is important. Without system software, the computer will not run, and without application software, the computer, no matter how powerful, will not be helpful in meeting user requirements. Think of it this way—applications apply the computer's thinking power to business tasks such as tracking the general ledger or billing your customers. Figure 11.7 illustrates that application software executes on the system software, which lies on the hardware layer.

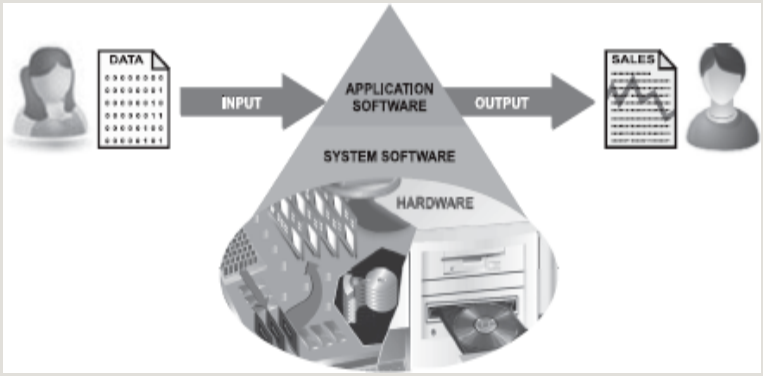


Figure 11.7 Relationship between Application and System Software

Application software ranges from games, calculators, and word processors (document creating programs), to programs that paint images on screen (image editors). Applications represent real world tasks. They can be easily divided by looking at exactly what function they serve. Some of the most commonly used application software are discussed below.

Word Processors: A word processor is a software used to compose, format, edit, and print electronic documents (see Figure 11.8). Word processing is one of the earliest applications for Office productivity and the personal computer. It involves not only typing, but also checking the spelling and grammar of the text and arranging it correctly on the page. A variety of different typefaces are available for a variety of effects. It is possible to include pictures, graphs, charts, and many other things within the text of the document. It also allows for changes in margins, fonts, and colour. Nowadays, virtually all personal computers are equipped with a word-processing program, which has functions for writing letters, reports or other documents, and printing. Examples of some well-known word processors are Microsoft Word and WordPerfect.

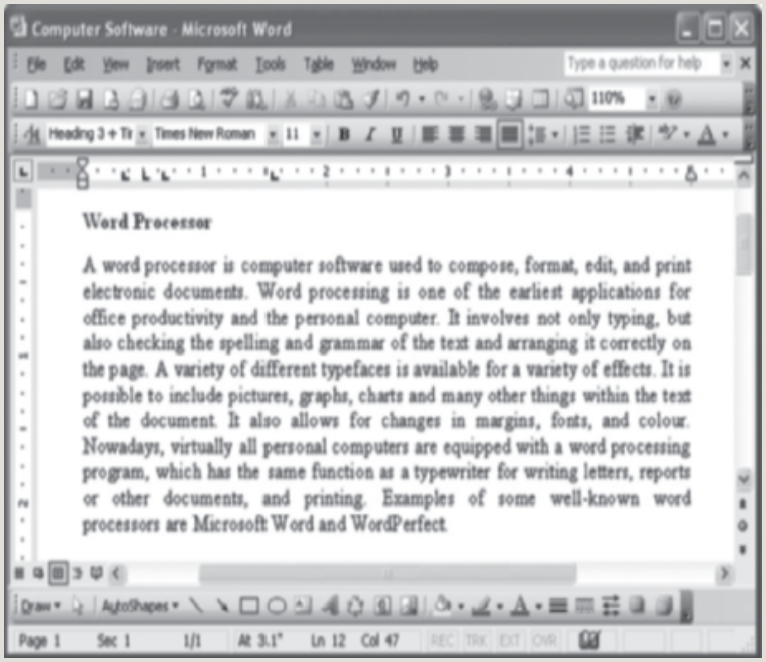
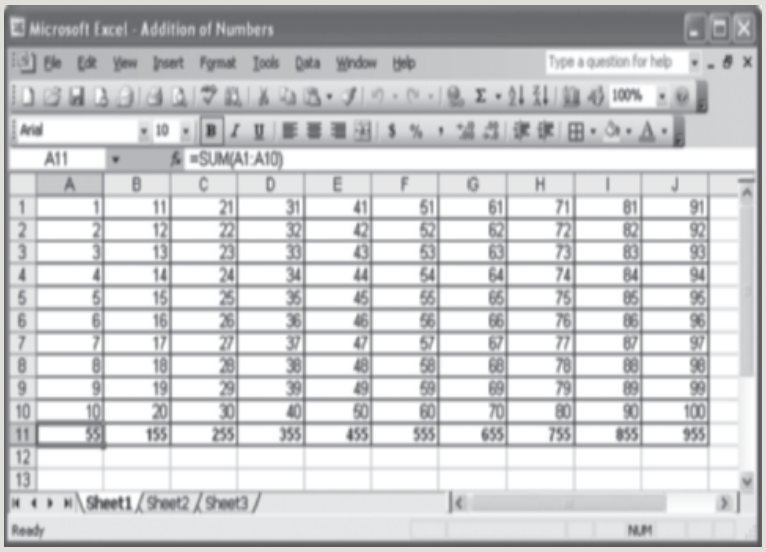


Figure 11.8 Word Processor

Spreadsheets: One of the first commercial uses of computers was in processing payroll and other financial records, so the programs were designed to generate reports in the standard "spreadsheet" format used by bookkeepers and accountants. A spreadsheet application is a rectangular grid, which allows text, numbers, and complex functions to be entered into a matrix of thousands of individual cells. The spreadsheet provides sheets containing *cells*, each of which may contain text and/or numbers. Cells may also contain equations that calculate the results from the data placed in other cells or a series of cells. A simple example might be a column of numbers totalled in a single cell containing an equation relating to that column (see Figure 11.9).



	A	B	C	D	E	F	G	H	I	J
1	1	11	21	31	41	51	61	71	81	91
2	2	12	22	32	42	52	62	72	82	92
3	3	13	23	33	43	53	63	73	83	93
4	4	14	24	34	44	54	64	74	84	94
5	5	15	25	35	45	55	65	75	85	95
6	6	16	26	36	46	56	66	76	86	96
7	7	17	27	37	47	57	67	77	87	97
8	8	18	28	38	48	58	68	78	88	98
9	9	19	29	39	49	59	69	79	89	99
10	10	20	30	40	50	60	70	80	90	100
11	55	155	255	355	455	555	655	755	855	955
12										
13										

Figure 11.9 Spreadsheet Application

Spreadsheet software is used predominantly for accounting and other numerically based tasks, because financial and mathematical data and analysis are often

managed in a tabular style, with columns of financial figures being manipulated and then influencing other data computations. Microsoft Excel and Lotus 1-2-3 are examples of spreadsheet applications.

Image Editors: Image editor programs are designed specifically for capturing, creating, editing, and manipulating images. These graphic programs provide a variety of special features for creating and altering images. In addition to offering a host of filters and image transformation algorithms, some image editors also enable the user to create and superimpose layers. Most graphic programs have the ability to import and export one or more graphic file formats. These computer programs enable the user to adjust an image to improve its appearance. With image editing software, one can darken or lighten an image, rotate it, adjust its contrast, crop out extraneous detail, and much more. Examples of these programs are Adobe Photoshop, Adobe Illustrator, and CorelDraw (see Figure 11.10).

Figure 11.10 Image Editing Software

Database Management Systems: Database management software is a collection of computer programs that allow storage, modification, and extraction of information from a database (which is a collection of interrelated records) in an efficient manner. It supports the structuring of the database in a standard format and provides tools for data input, verification, storage, retrieval, query, and manipulation. When such software is used, information systems can be changed much more easily as the organization's information requirements change. New categories of data can be added to the database without disrupting the existing system. It also controls the security and integrity of the database from unauthorized access. FoxPro, MS-Access, and Oracle are database management systems (see Figure 11.11).

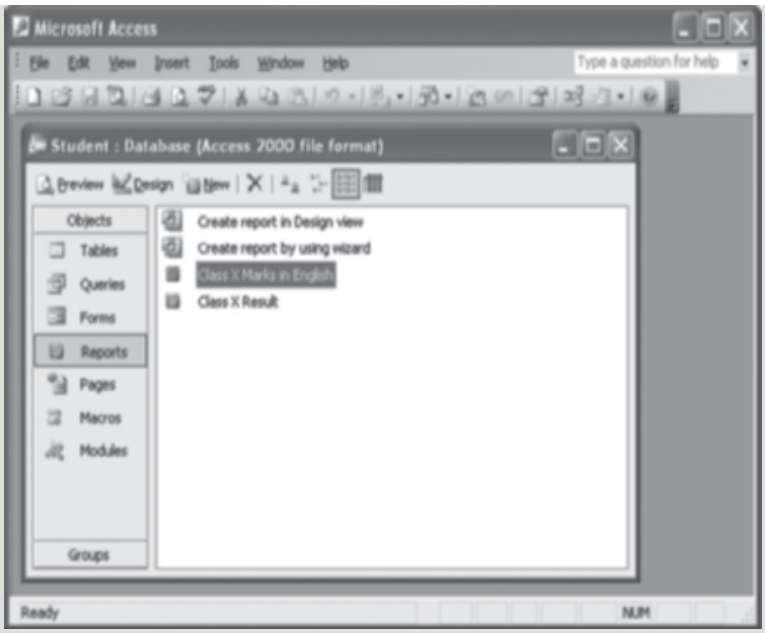


Figure 11.11 DBMS Application

Presentation Applications: A presentation is a means of expressing ideas in the form of visual images, audio, and animation. To prepare presentations, presentation software is used. Presentation software allows the users to create presentation by producing slides or handouts for presentation of projects. Essentially, such computer programs allow users to create a variety of visually appealing electronic slides for presentations. Microsoft PowerPoint is one of the most famous presentation applications (see Figure 11.12).

Figure 11.12 File System

Desktop Publishing Software: The term desktop publishing is usually used to describe the creation of printed documents using a desktop computer. It is a technique of using a personal computer to design images and pages, assemble type and graphics, and then using a laser printer or image setter to output the assembled pages onto a paper, film, or printing plate. Desktop publishing software is used for creating magazines, books, newsletters, and so on. Such software assist in creating sophisticated documents including complicated page designs, detailed illustrations, and camera-ready typefaces. Quark Express and Adobe PageMaker are desktop publishing software (see Figure 11.13).

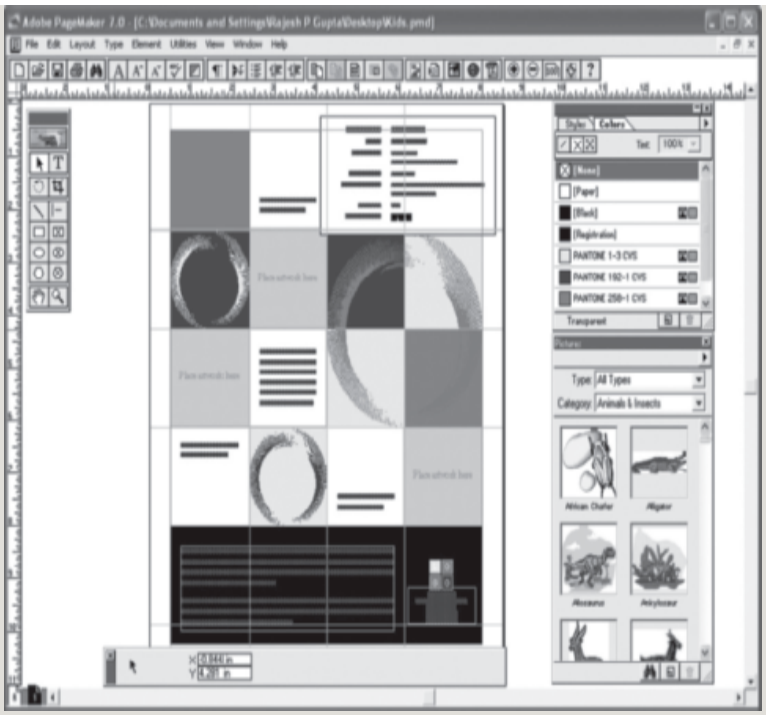


Figure 11.13 Desktop Publishing Software

11.4 INSTALLING AND UNINSTALLING SOFTWARE

Once software (system or application) is carefully selected and purchased, it should be installed on the computer to start working with it. During the installation-process, files are copied to their appropriate folders and icons are added to the system menus and/or desktop. Software has recently become much more user friendly in this area. Most softwares handle the installation chores without a lot of input from the user.

THINGS TO REMEMBER

Zip Files

The zip files format was originally created by Phil Katz in January 1989. A zip file is a collection of files compressed into a single unit (file) to take up less

space and reduce download times. Before using a zip file, the user needs to unzip (uncompress) it. To zip or unzip a file, utility such as WINZIP, PKZIP, and NETZIP is used. In some cases, a self-extracting (exe) zip file can be created that does not require any utility to open and uncompress the file.

The process of copying software from CD-ROM (or floppy disks) to the hard disk and configuring it to work with the computer hardware is known as *software installation* (see Figure 11.14). Software vendors provide special programs, known as installer, to carry out the installation process. An *installer* decompresses the required files within it and places them in proper folders. It is the process of arranging a program in a computer system such that the program works as desired by the user. The first step in installation is to locate the actual installation file. Installation files usually have the extension .EXE or .ZIP. If the installation is of .EXE extension (usually *setup.exe* or *install.exe*), simply double-click it and follow the instructions that appear on the screen. On the other hand, if the file has a .ZIP extension, the user needs to unzip the compressed files and then install the software.

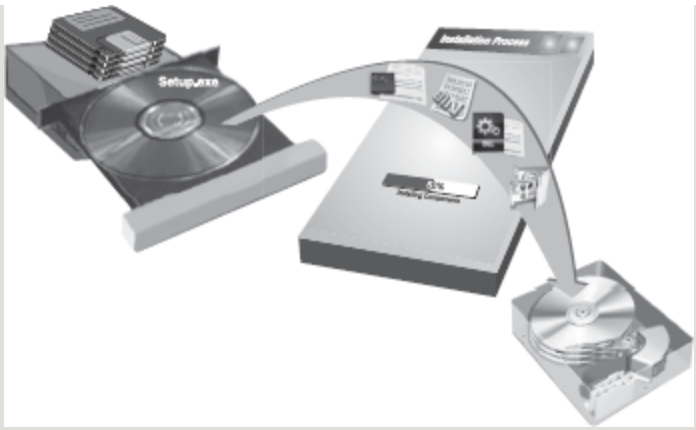


Figure 11.14 Software Installation Process

Usually software needs timely maintenance and upgrades to cater to the growing needs of the users and hardware changes. Sometimes the older version of the software (or parts of it) is replaced by their newer versions. This process of adopting a newer version of existing software with an enhanced feature is known as *software upgrade*. Generally, a new release (or version) of software contains new functionalities and upgrades are numbered. For example, various upgrades of Microsoft Internet Explorer had the versions IE4.0, IE5.0, IE5.5, and IE6.0.

Apart from software upgrade, the existing software also undergoes the process of update. *Software update* is a new release (or version) of software that is generally understood to be an error-correction release and does not contain new functionality. A software update usually adds relatively minor new features to a product or corrects errors (bugs)found since the program has released. Another important term related to software upgrade and update is software patch. A software patch is a collection of one or more files that correct flaws in the performance, reliability, or security of a specific

software product. Each patch remedies a specific set of reported errors in the version of the product to which the patch applies. Every patch is identified by a unique number, such as “patch567.” The number is used in manufacturing to identify the errors that the patch corrects (see Figure 11.15).

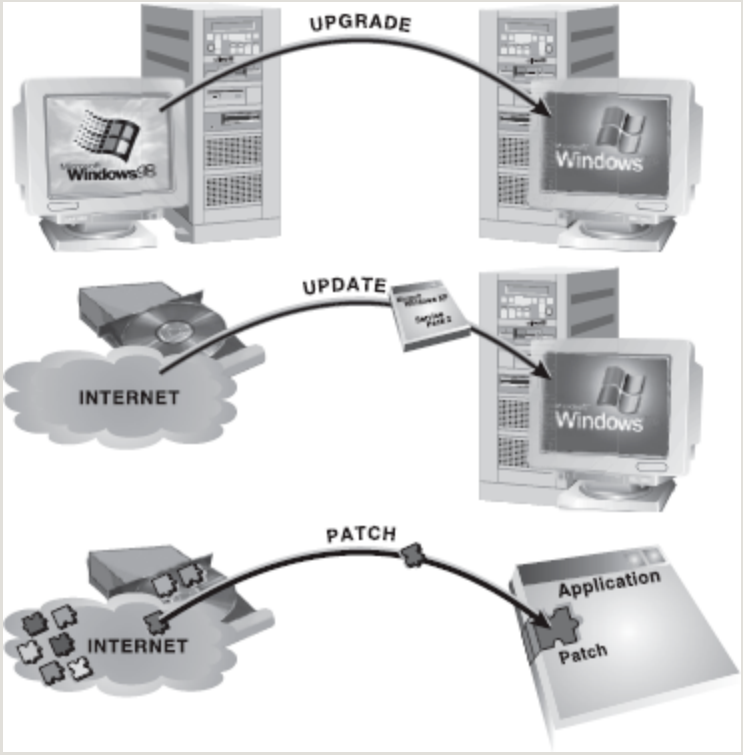


Figure 11.15 Software Upgrade, Update, and Patch

Sometimes software is installed just for testing or perhaps it has been superseded by a better software. Hence, unwanted software should be uninstalled, that is, cleanly removed. Uninstall program completely removes all files of that program and all associated files in other directories. Note that simply deleting the folder in which

the software is stored will not remove all the files related to that software; it needs to be uninstalled to remove all the linked files of the software. When a program is installed, its files are copied to the system disk, and changes are made to certain system files and folders to allow the program to work properly. Properly uninstalling a program means deleting all the files and undoing any changes made to system files.

11.5 SOFTWARE PIRACY

Software piracy is the unauthorized copying of an organization's internally developed software or the illegal duplication of commercially available software. While purchasing software, the user gets a licence to use the application but it cannot be copied on other machines. The use of the software must comply with the product licence agreement that is included with the software. A *software license* is a type of proprietary licence, which acts as a memorandum of contract between the producer and the user of computer software. It is sometimes also called an *end-user license agreement* (*EULA*), which specifies the permissions granted by the owner to the user.

11.5.1 Types of Software Piracy

There are six basic forms of software piracy, and all are harmful for the software publisher as well as the for the end user.

Softloading: *Softloading* (also known as *softlifting*) means sharing a program with someone who is not authorized by the licence agreement to use it. A common form of softloading involves purchasing a single licensed copy of software and then loading the software onto several computers, in violation of licensing terms. For example, sharing software with friends, co-workers, and others.

Software Counterfeiting: *Software counterfeiting* means producing fake copies of software and making it look authentic. This involves providing the packaging box, CDs or floppies, and manuals, all designed to look like the original product. Most commonly, a copy of a CD is made with a CD burner, and a photocopy of the manual is made. Counterfeit software is sold on street corners, and sometimes unknowingly sold even in retail stores. Counterfeit software is sold at prices far below the actual retail price.

Internet Downloading: It is the fastest growing form of piracy. With the growing number of users online and with the rapidly increasing connection speeds, the exchange of software on the Internet has attracted an extensive following. Currently, there are hundreds of thousands of "warez" sites providing unlimited downloads to any user. Often, the software provided through these warez sites is cracked to eliminate any copy protection schemes.

Renting: It involves someone renting out a copy of software for temporary use, without the permission of the copyright holder. The practice, similar to that of renting a video of a movie, violates the licence agreement of software.

Hard Disk Loading: This form of piracy involves installing an unauthorized copy of software onto a computer being sold to the end user. It is often committed by hardware dealers. This makes the deal more attractive to the buyer, at virtually no cost to the dealer. Dealer usually does not provide the buyer with manuals or the original CDs of the software.

Licence Misuse: Software is distributed under special discount licences to high-volume customers, computer manufacturers, or academic institutions. After getting a copy of software, manufacturers, customers, or

institutions redistribute the software to others who do not hold or qualify for these licences. This constitutes the misuse of licence. Some of the ways a licence can be misused are:

- **Original Equipment Manufacturer (OEM) Unbundling:** Often just called *unbundling*, it involves selling stand-alone software, which is originally meant to be included with a specific accompanying product. An example of this form of piracy is someone providing drivers with a specific printer without authorization.
- **Academic Product Misuse:** This form of licence misuse occurs when a product that has been specifically marked for distribution to educational institutions and students at reduced prices is diverted into normal commercial channels of distribution.
- **Not for Resale (NFR) Product Misuse:** This form of licence misuse occurs when a product that has been clearly marked *Not for Resale* is diverted into normal commercial channels of distribution. Such products are typically distributed as a promotional or sample product and are not licensed for normal commercial distribution and use.

11.5.2 Preventing Software Piracy

Software piracy all over the world has become a major problem. To avoid this, several licence and copyrights are issued to the software and the users. A *copyright* is a form of intellectual property that grants its holder the legal right to restrict the copying of an original, creative expression for a defined period. Nowadays, some software developers use software copyrights in order to ensure that the software they develop will remain free, using licences such as the *general public license (GPL)*. It is a good idea to put the copyright notice on the computer software both on the packaging and in the program as it serves as a warning to others that the software is protected by copyright. In addition to the copyright notice, warning against unauthorized use should also be mentioned. The general rule is that the creator of some work is the first owner of copyright on it. Here are some guidelines to prevent software piracy:

- Always buy software from authorized sellers and when in doubt, contact the software manufacturer for a list of authorized sellers.
- While purchasing software, consumers should make sure that they get authenticated software licences. In case of hardware purchase, do not

allow dealers to install unlicensed software onto the computer.

- Unless authorized in the licensing agreement, employees of an organization should not make personal copies of software used in an organization.
- One should not make copies of software other than those allowed by the licence agreement.
- Software renting or reselling must not be practised.
- The best way to prevent software piracy is to report it to the concerned copyright holder.

Piracy affects us all. It is not just the IT industry that loses out to software piracy; the user is often left in a vulnerable position. They will not be entitled to technical support or have access to information regarding upgrades. Illegal products may be incomplete, virus contaminated, or their functionality may be impaired. Hence, if software consumers have any information regarding the abuse or illegal sale of software, they should contact the software copyright holder.

11.6 SOFTWARE TERMINOLOGIES

Some of the common terms used in the field of software are given below:

Public Domain Software: *Public domain software* refers to any program that is not copyrighted. This software is free and can be used without restrictions, that is, the user can copy, distribute, and even modify the software without obtaining permissions from the software developer. Usually, public domain software is offered through websites, electronic bulletin boards, user groups, and other sources.

Freeware: The term *freeware* is commonly used for copyrighted software that is given free by its owner. Though it is available for free, the owner retains the copyright, which means that a user does not have the right to modify anything in the software. Thus, freeware software permits redistribution but not modification. Freeware usually carries a licence that permits redistribution but may have other restrictions such as

limitations on its commercial use. Freeware is an umbrella term that covers:

- **Donationware:** The authors of donationware ask that anyone using their software make a donation to the authors or to some third party such as a charity.
- **Postcardware:** It is freeware (no-charge software that is freely shared) that requires only that the user send the software provider a postcard as a form of payment.
- **Abandonware:** It is commercial software that has not been sold for a long time or whose copyright holder is defunct. The licences of most of such software forbid redistribution or require payment; hence distributing it violates the owner's copyright (even if the owner does not or cannot enforce it).
- **Adware:** It is distributed as freeware, but it requires the user to view advertisements to use the software. Note that adware is sometimes called *spyware*.

Shareware: *Shareware* is the software that is distributed free on a trial basis. It allows people to redistribute copies for a limited period. Anyone who continues to use a copy is required to pay a licence fee. Therefore, the free use of the software is usually limited to a certain period. It is distributed without payment ahead of time as is common for proprietary software. Typically shareware software is obtained free of charge by downloading, thus allowing one to try out the program ahead of time. It is accompanied by a request for payment, and often payment is required as per the terms of the licence past a set period. WinZip software is an example of shareware, which is distributed with a 30-day trial period.

Commercial Software: *Commercial software* represents the majority of software purchased from software publishers. This software comes pre-packaged, is available in software stores, and can be obtained through the Internet. It is developed by business organizations, which aim to earn profits from its use. It is always copyrighted. The licensing restrictions vary somewhat from vendor-to-vendor and product-to-product, and change frequently. It is generally more

expensive than competing shareware products, for example Microsoft Windows.

Open-Source Software: *Open-source software* is created by generous programmers and released into the public domain for public use. The underlying programming code is available to the users so that they may read it, make changes to it, and build new versions of the software incorporating their changes for software. Usually these softwares are distributed under an open-source licence—GPL. Such a licence requires that the source code should be distributed along with the software, and that the source code is freely modifiable, with least restrictions, such as the requirement to preserve the author's names and copyright statement in the code. The Linux operating system is an example of open-source software.

Proprietary Software: In general, *proprietary* describes a technology or product that is owned exclusively by a single company that carefully guards knowledge about the technology or the product's internal working. Some proprietary products can only function properly, if at all, when used with other products owned by the same company. Example of a proprietary product is Adobe Acrobat, whose portable document format (PDF) files can only be read with Acrobat Reader. Proprietary software is also called as closed-source software (CSS). Its use, redistribution, or modification are prohibited or restricted so much that the user effectively cannot use it freely.

Firmware: *Firmware* is a program or data that has been written onto read-only memory (ROM). For example, the BIOS (which is installed inside a computer on a chip) checks different parts of the system before loading the operating system into memory. Being firmware, it ensures that it will always be available and

will not be damaged in the case of a power failure. ROMs, PROMs, and EPROMs that have data or programs recorded on them are firmware (see Figure 11.16).

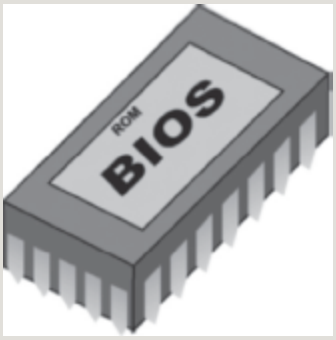


Figure 11.16 Firmware

LET US SUMMARIZE

1. *Software* is a generic term for organizing collection of computer data and instructions. Computer software can be divided into two major categories: *system software* and *application software*.
2. *System software* consists of several programs, which are directly responsible for controlling, integrating, and managing the individual hardware components of a computer system. Some examples of system software are *operating system, device driver, language translators,* and *system utility*.
3. *Application software* is used to accomplish specific tasks rather than just managing a computer system. Word processors, spreadsheets, image editors, database management systems, presentation software, and desktop publishing applications are some of the examples of application software.
4. *Software installation* is the process of copying software from CD-ROM (or floppy disks) to the hard disk and configuring it to work with the computer hardware.
5. Software upgrade is the process of adopting a newer version of existing software with an enhanced feature.
6. *Software update* is a new release (or version) of software that is generally understood to be an error-correction release and does not contain new functionality. It usually adds relatively minor new features to a product or corrects errors (bugs) found since the program has released.
7. A *software patch* is a collection of one or more files that correct flaws in the performance, reliability, or security of a specific software product.

8. *Software piracy* is an unauthorized copying of an organization's internally developed software or the illegal duplication of commercially available software. It includes illegal activities such as softloading, software counterfeiting, Internet downloading, renting, hard disk loading, and licence misuse.
9. *Public domain software* refers to any program that is not copyrighted. This software is free and can be used without restrictions, that is, the user can copy, distribute, and even modify the software without obtaining permissions from the software developer.
10. The term *freeware* is commonly used for copyrighted software that is given free by its owner. It permits redistribution but not modification and usually carries a licence that permits redistribution but may have other restrictions, such as limitations on its commercial use.
11. *Shareware* is the software that is distributed free on a trial basis. It allows people to redistribute copies for a certain period.
12. *Commercial software* represents the majority of software purchased from software publishers. This software comes pre-packaged, is available in software stores, and can be obtained through the Internet.
13. *Open-source software* is created by generous programmers and released into the public domain for public use. The underlying programming code is available to the users so that they may read it, make changes to it, and build new versions of the software incorporating their changes for software.
14. *Proprietary software* describes a technology or product that is owned exclusively by a single company that carefully guards knowledge about the technology or the product's internal working.
15. *Firmware* is a program or data that has been written onto ROM.

EXERCISES

Fill in the Blanks

1. .....is the set of coded instruction that brings the machinery to life.
2. Software is often divided into two major categories, namely, ... and .....
3. A.....acts like a translator between the device and programs that use the device.
4. The function of an operating system that allows creation, deletion, suspension, resumption, scheduling, and synchronization of processes is called.....
5. .....program scans the hard disk for viruses and removes them.
6. .....helps in converting programming language into machine language.
7. .....is the process of copying software from CD-ROM (or floppy disks) to the hard disk and configuring it to work with the computer hardware.
8. .....is the unauthorized copying of an organization's internally developed software or the illegal duplication of commercially available software.
9. .....is the software, which comes with permission for people to redistribute copies for a limited period.

10. Firmware is the program or data that has been written onto ……..

Multiple-choice Questions

1. Which of the following is system software?
    1. Microsoft Word
    2. Adobe Photoshop
    3. Microsoft PowerPoint
    4. Windows 98
2. ……..analyses and executes each line of source code in succession, without looking at the entire program.
    1. Compiler
    2. Operating System
    3. Interpreter
    4. Device Drivers
3. The assembly language program must be translated into machine code by a separate program called……..
    1. Assembler
    2. Interpreter
    3. Compiler
    4. Loader
4. Softloading is also known as ……..
    1. Softlifting
    2. Softwarecopy
    3. Both (a) and (b)
    4. None of these
5. If you change Windows 98 operating system to Windows XP, you are actually performing……..
    1. Update
    2. Upgrade
    3. Patch
    4. None of these
6. Which of the following is application software?
    1. Database Management System
    2. Image Editor
    3. Spreadsheets
    4. All of these
7. Usually, installation files have the extension……..
    1. .doc
    2. .gif
    3. .exe
    4. .jpeg
8. Which of the following activity can be termed as software piracy?
    1. Softloading
    2. Counterfeiting
    3. Renting
    4. All of these
9. Choose the odd one out
    1. Donationware
    2. Postcardware
    3. Shareware
    4. Adware
10. Linux is a type of …….. software.

1. Shareware
2. Open Source
3. Commercial
4. Proprietary

State True or False

1. Operating system is application software.
2. A device driver allows another program to interact with a hardware device.
3. Data recovery is the process of retrieving deleted or inaccessible data from failed electronic storage media.
4. Anti-virus is the program to protect from the piracy.
5. Installation files usually have the extension .EXE or .ZIP.
6. Software update and software upgrade perform the same function.
7. Software piracy is the authorized copying of an organization's internally developed software.
8. Public domain software refers to any program that is not copyrighted.
9. Firmware software is embedded in the hardware devices.
10. Microsoft Word is an image editing software.

Descriptive Questions

1. What do you understand by software? Discuss its types.
2. Explain the difference between hardware and software with their definition and uses.
3. Explain the concept of software installation and uninstallation.
4. How is software update different from software upgrade? Explain.
5. What is software piracy and how can it be avoided?

ANSWERS

Fill in the Blanks

1. Software
2. System Software, Application software
3. Device Drivers
4. Process Management
5. Anti-Virus
6. Language Translators
7. Software Installations
8. Software piracy
9. Shareware
10. Read-only memory (ROM)

Multiple-choice Questions

1. (d)
2. (c)
3. (a)
4. (a)
5. (b)

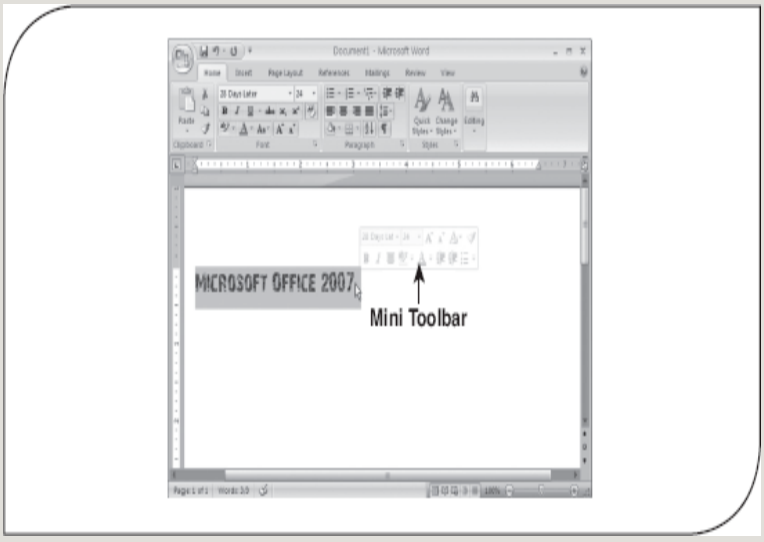
6. (d)
7. (c)
8. (d)
9. (c)
10. (b)

State True or False

1. False
2. True
3. True
4. False
5. True
6. False
7. False
8. True
9. True
10. False

Chapter 12

Introduction to Microsoft Office 2007



This chapter familiarizes the reader with the latest version of Microsoft Office named Microsoft Office 2007, which has entirely new look and feel from that of previous versions. The new user interface of Microsoft Office 2007 makes the most features you need readily accessible as well as you can find other useful features that you might not already know very easily. Thus, it helps in getting more work done, more easily and in a more efficient manner. The chapter begins with the introduction of new features and applications included in Office 2007. The discussion continues with highlights of new user interface of Office 2007 such as quick access toolbar, Ribbon, command tabs, command sets, and many more. Finally, it concludes by discussing some new and enhanced features included in common applications of Office 2007 such as MS Word 2007, MS Excel 2007, MS Access 2007 and MS PowerPoint 2007.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The user interface of Microsoft Office 2007
- The new features of Microsoft Office Word 2007
- The new features of Microsoft Office Excel 2007
- The new features of Microsoft Office Access 2007
- The new features of Microsoft Office PowerPoint 2007

12.1 INTRODUCTION

Office 2007 is the latest version of Microsoft's productivity suite. Other than the applications in previous versions, it includes new applications like Groove 2007, OneNote 2007, and Communicator 2007. Some new server-side tools like Office SharePoint Server 2007 have also been added in Office 2007. In addition, it offers a variety of new and exciting features. The most noticeable features that are common to the core applications in the Office 2007 include the new task-oriented graphical user interface, the new Office XML formats, and the ability to easily attach custom XML data to a file. In this chapter, we introduce the Office 2007 user interface and discuss some of the new features or enhancements made in its core applications.

12.2 INTRODUCTION TO OFFICE 2007 USER INTERFACE

Microsoft has made significant changes in the user interface of most of the applications including Word, Excel, PowerPoint, Access, and Outlook (in the composing and reading window) of Office 2007. In the earlier versions of MS Office, there were menus and toolbars using which a user could find the features and commands to get the desired task done. In Office 2007, a

new interface named as *Office Fluent* has been introduced in which these menus and toolbars have been replaced with the Ribbon and Office button. This new user interface is more intuitive and easier to navigate. Some of the highlights of this new interface are discussed in this section (Figure 12.1).

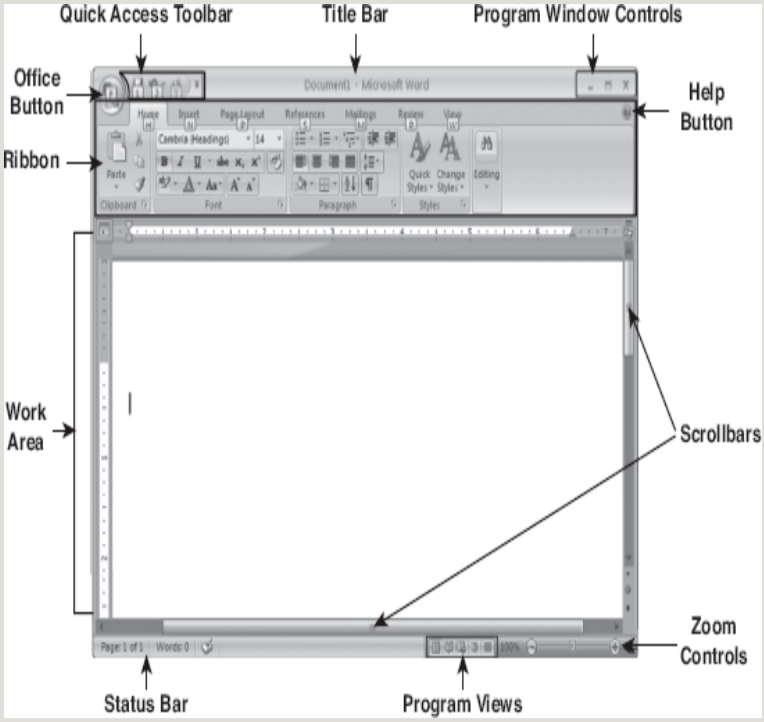


Figure 12.1 User Interface of Office Word 2007

Office Button: The *Office Button* ( ) is located at the top left corner of the Office program window. It provides quick access to the basic commands that are common to all Office programs like creating, opening, saving, printing and closing a file. Other than these basic commands, it has some more commands like **Prepare, Finish,** and **Publish** depending on the program being used. Using this button, you can also access a list of

documents you have most recently worked with and a set of various program options (Figure 12.2).

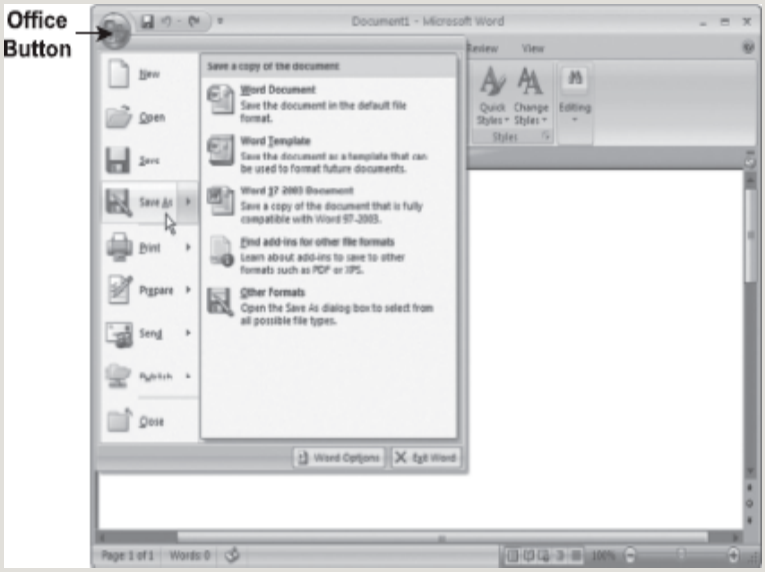


Figure 12.2 Office Button

Note: *In Microsoft Office Outlook 2007, the Office Button is displayed only when you read or create a message, task, contact, or calendar item.*

Quick Access Toolbar: The *Quick Access Toolbar* is a customizable toolbar located to the right of the Office button on the Office program window. By default, it contains **Save, Undo**, and **Redo** buttons. You cannot only add or remove the buttons to the Quick Access Toolbar but also move the Quick Access Toolbar below the ribbon (that is, near the work area).

You can customize the **Quick Access Toolbar** by using the **Customize Quick Access Toolbar** button ( ) which is present immediate right to the Quick Access Toolbar. When you click this button, a menu appears that contains a list of the commands. To add a command

in the Quick Access Toolbar, select it from the menu. A check mark beside the command name indicates that the command is added to the Quick Access Toolbar. To remove a command from the Quick Access Toolbar, deselect the name of the command in the menu. You can also search for more commands to add to the Quick Access Toolbar by clicking the **More Commands** button in the menu (Figure 12.3).

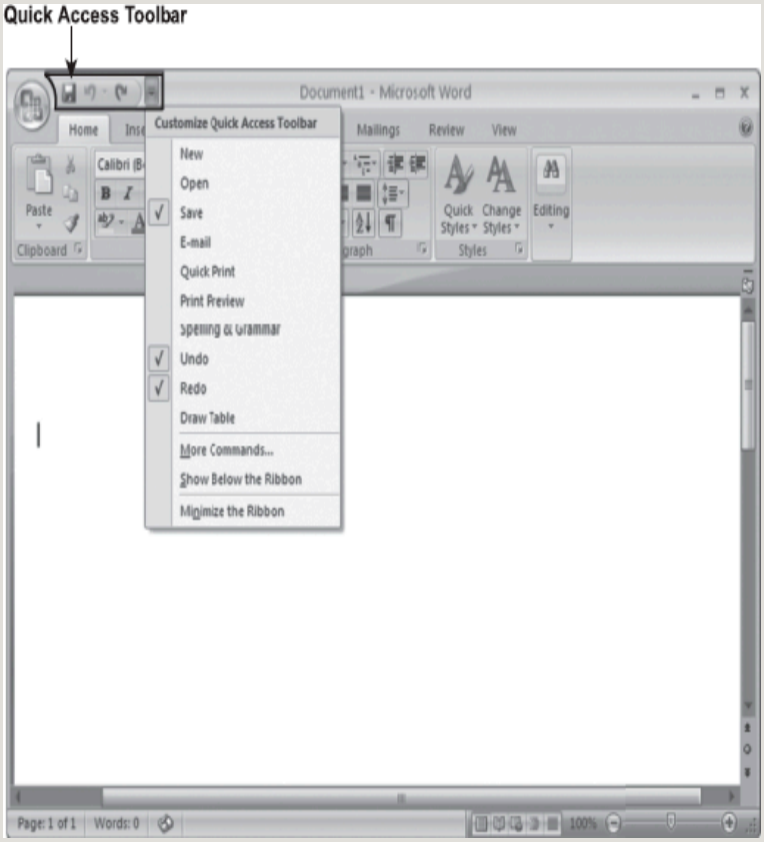


Figure 12.3 Customizing Quick Access Toolbar

Ribbon: As stated earlier, in some of the applications of Office 2007, the menus and toolbars have been replaced with the Ribbon. The *Ribbon* is located just below the title bar. It is designed specially to provide

quick access to commands that you need to complete a task. It consists of the commands which are organized into logical groups; these groups are further organized under tabs. The ribbon along with its components is shown in (Figure 12.4).

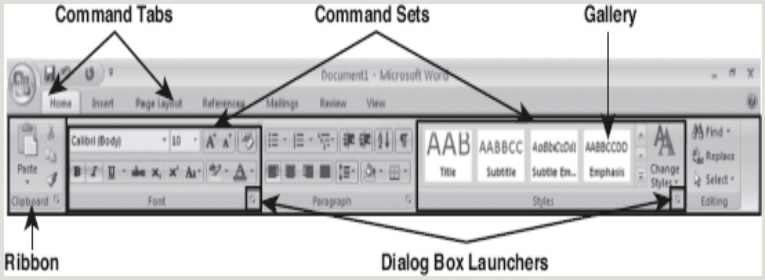


Figure 12.4 Ribbon

Note: *You can minimize or restore the ribbon by pressing* **Ctrl+F1.**

Command Tabs: A *command tab* or simply *tab* consists of commands that are related to a specific type of activity. For example, **Page Layout** tab consists of the commands related to the layout or formatting of a page or pages in your document. By default, the commands under the **Home** tab are displayed on the ribbon. You can view or use the commands under a different command tab by simply clicking on that command tab.

Command Sets: The commands in a command tab are organized under a group according to their functions. The group of commands related to the selected tab in the program window is known as *command set*.The commands specific to the formatting of font like Font, Font size, Bold, Italic, etc. are grouped under the **Font** command set. The name of the command set is displayed below the commands. The command sets under Home

tab are **Clipboard, Font, Paragraph, Styles**, and **Editing**.

Note: *In addition to commands, some command sets have galleries. A gallery provides you various options that you can apply on selected text or object.*

Dialog Launcher: MS Office 2007 also provides traditional style dialog boxes for some command sets on the Ribbon. The dialog boxes for the command sets are indicated by the presence of a small arrow ( ) at the bottom right corner of a command set (known as *dialog launcher*). Clicking this arrow, you can open the dialog box that allows you to access the commands of that command set. For example, if you click the dialog launcher at the bottom of the **Paragraph** command set under the **Home** tab, a Paragraph dialog box appears (Figure 12.5).

Figure 12.5 Paragraph Dialog Box

Contextual Tabs: All the available tabs do not appear on the ribbon; some tabs appear only when needed. Such

tabs are called *contextual tabs*. These tabs become visible when you create or select a specific object in your document. For example, when you insert or select a table, the **Table Tools** section with **Design** and **Layout** tab appears on the Ribbon (see Figure 12.6). The Design tab consists of a variety of command sets that allow you to format a table the way you want it to appear. The Layout tab provides command sets that allow you to position and organize data in the cells in the way you want. The purpose of the contextual tabs is to reduce cluttering of commands on the ribbon.

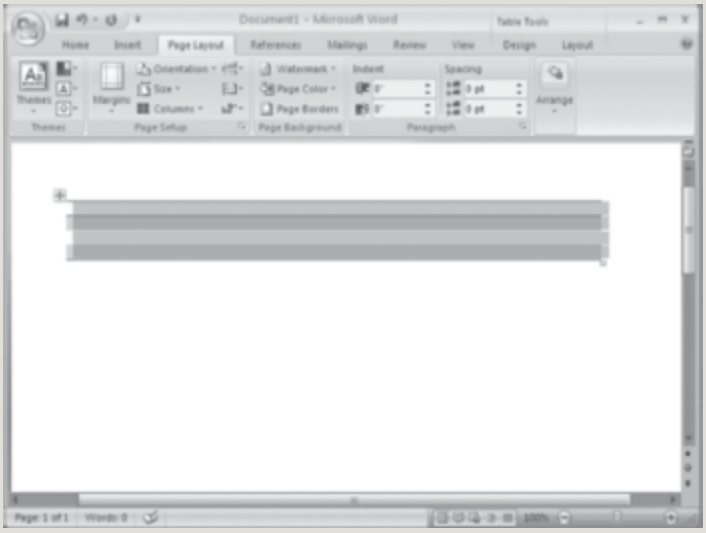


Figure 12.6 Contextual Tabs after Selecting Table

KeyTips: Like earlier versions of Office, Office 2007 also provides keyboard shortcuts to access the commands quickly. However, the users need not remember the keyboard shortcuts. Office 2007 provides you KeyTips that show a way to access the commands via the keyboard. KeyTips are nothing but the indicators that display the keyboard shortcuts for the commands. To activate the KeyTips, simply press **Alt** key. On pressing

the Alt key, the KeyTip will appear on the Office button, tab names, and all command buttons in the Quick access toolbar (see Figure 12.7). Note that a KeyTip can be a single letter, digit, or a combination of letter and numbers.

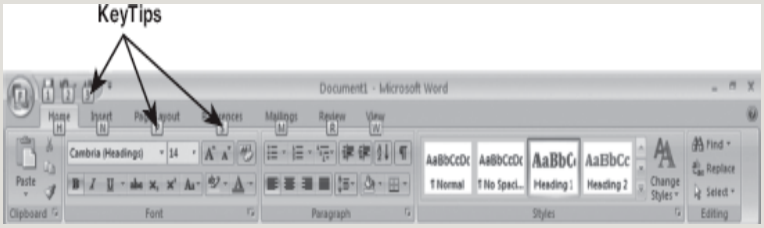


Figure 12.7 KeyTips

You can use the desired features by pressing the corresponding KeyTip. For example, if you want to insert a picture in your document, press **N** to activate the **Insert** tab. Doing this displays the KeyTip over each command under the Insert tab (see Figure 12.8). Now to insert picture, press **P**. Similarly, you can use any command under any tab.

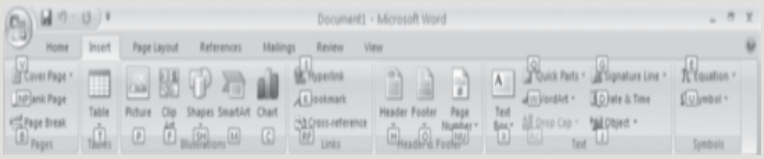


Figure 12.8 KeyTips under Insert Tab

Super ToolTips: A very convenient feature provided by the Office 2007 is the Super ToolTip. A Super ToolTip

is the small window that provides information about a command along with an image. It is displayed when the user rests the mouse pointer on the command, as shown in (Figure 12.9). It can be considered as the enhanced form of the ToolTip as it has larger window that displays more descriptive text than a ToolTip and can have a link to a Help topic. To hide the Super ToolTip, simply move the mouse away from the command.

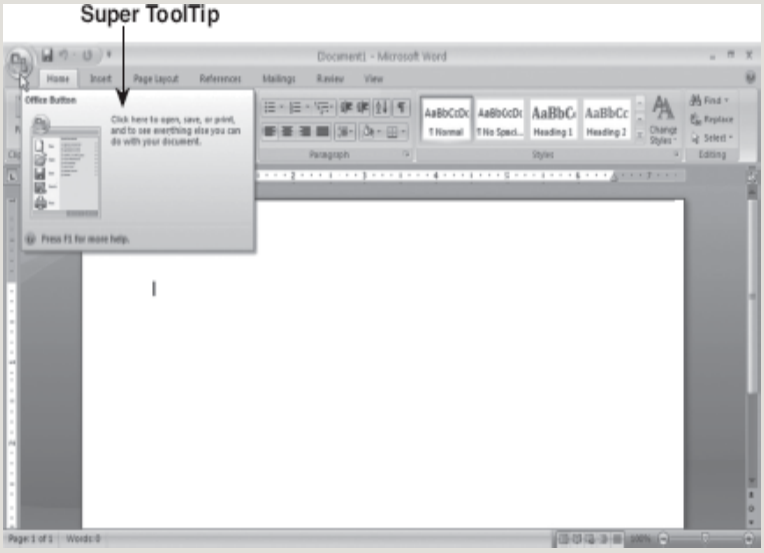


Figure 12.9 Super ToolTips

Note: *Whenever a user points the mouse over the dialog launcher, the Super ToolTip displays an image of the dialog box that will appear if you click that dialog launcher.*

Mini Tool bar: A mini toolbar is a handy, miniature, and semitransparent toolbar that appears when you select some text in an Office program. It contains the basic text formatting commands including Font, Font styles, Font size, Font Color, etc. It allows you to quickly access the commands that are frequently used while

working in an Office program. When you move the mouse pointer over some command in the mini toolbar, it becomes non-transparent. Mini toolbar also appears above/below the pop-up menu that appears when you right-click the selected text (Figure 12.10).

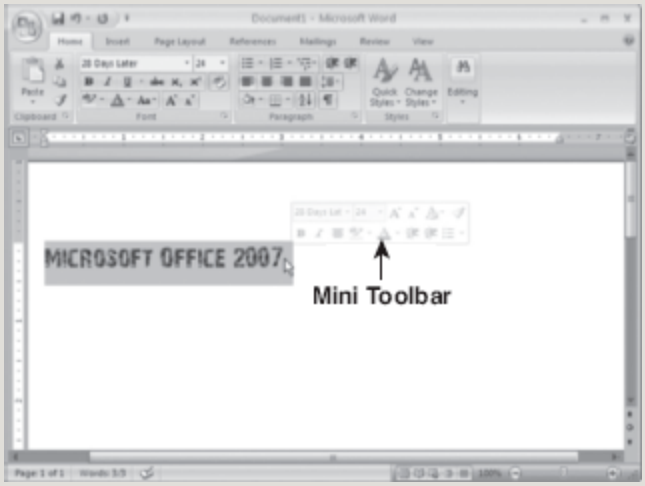


Figure 12.10 Mini Toolbar

Note: *You cannot customize the Mini toolbar.*

Live Preview: The *Live Preview* feature enables you to see the effect of the selected option from the drop-down menu or gallery on the selected object or current document before actually applying it. To use this feature, simply move the mouse pointer over an option. The effect of that command will be shown in the work area of application, thereby allowing users to try different options quickly and select the one that works best before actually applying it to the document. It saves time by avoiding the process of formatting, undoing, and redoing repeatedly to get the desired results.

Microsoft has not only made changes in the user interface of Office 2007 but has also added new features as well as enhanced some existing features in each program also. This section covers the enhancements made in the four main Office 2007 programs including Word, Excel, Access, and PowerPoint.

12.3.1 Microsoft Office Word 2007

Today, MS Word is the most popular word processor that helps you produce professional-looking documents. Using Word 2007, you can create and format your document effortlessly than the earlier versions. Some of the new features/enhancements made in Word 2007 are as follows:

- **Charting:** The new charting and diagrammatic features of Word 2007 allow you to create high-impact graphics, thereby producing an effective document. These features include transparency, drop shadows, contour effects, etc.
- **Document Comparison:** In Word 2007, you can compare two versions of a document quickly. It allows you to find what changes were made to a document by marking the deleted, inserted, and moved text in the third version of the document.
- **Digital Signature:** Word 2007 allows you to add digital signature to your document, thereby ensuring the authenticity, integrity, and origin of the document. Once a digital signature has been added to the document, it becomes read-only so that no modifications can be made to its content.
- **Mark as Final:** The Mark as Final feature of Word 2007 prevents readers from making changes in a document by making it read-only. When this command is applied to a document, all the commands related to the editing are disabled or turned off and a Mark as Final icon appears in the status bar to indicate its final status. Note that it only conveys to the readers that you are sharing a final version of a document. Anyone who receives the document that has been *marked as final* can edit that document by removing Mark as Final status from the document.
- **Conversion of Word Documents to PDF or XPS:** Word 2007 allows you to save files in portable document format (PDF) or XML paper specification (XPS) which are the common file formats for sharing documents. Both these formats ensure that formatting of document remains intact whenever the document is viewed online or printed and its contents cannot be modified easily.

12.3.2 Microsoft Office Excel 2007

Microsoft Office Excel 2007 allows you to analyse, manage, and share the data easier than the earlier versions. Some of the new features/enhancements made in Excel 2007 are as follows:

- **Increased Number of Rows and Columns:** To manage large amount of data in a worksheet, the number of rows and columns has been increased. Excel 2007 provides you more than 1 million rows and 16 thousand columns in a worksheet.
- **Conditional Formatting:** In Excel 2007, conditional formatting can contain up to 64 conditions instead of only three conditions as in earlier versions of Excel. Moreover, it also provides new conditional formatting types like colour scales, data bars, or icon sets.
- **Easy Formula Writing:** In Excel 2007, writing formulas has become easier because of several enhancements like resizable formula bar, Function AutoComplete, and structured references. The formula bar is resized on entering long, complex formulas, which prevents other data from being hidden by the formulas in your worksheet. Function AutoComplete feature allows you to quickly write the proper formula syntax. Excel 2007 also facilitates to use the structured references to reference named ranges and tables in a formula. This makes it easier to work with table data while using formulas that reference a table, either portions of a table or the entire table.
- **New OLAP Formulas and Cube Functions:** In Excel 2007, some new OLAP formulas and functions have been introduced to work with multidimensional databases (such as SQL server analysis services). Using OLAP formulas, you can build complex, free form, OLAP data-bound reports. New cube functions can be used to extract OLAP data from analysis services and display it in a cell.
- **Enhanced Sorting and Filtering:** Excel 2007 allows you to arrange the data in some specific order more quickly and easily than the earlier versions of Excel using enhanced filtering and sorting features. In Excel 2007, you can sort data up to 64 levels instead of only three levels. You can filter data by a specific format also, such as cell colour, font colour, or dates instead of text only.
- **Easy-to-use PivotTable and PivotCharts:** Excel 2007 allows you to create PivotTables with much ease. You can create a PivotTable by simply selecting the fields that are needed instead of dragging the data as required in earlier versions. Once a PivotTable is created, you can sort, group, and calculate totals for a large amount of data with ease using new and enhanced features. Like PivotTables, PivotCharts can be created easily to observe comparisons, patterns, and trends. In addition, Excel 2007 offers many tools and menus specific to the PivotChart that makes the analysis of chart data easy.
- **Saving to New File Formats:** Like Word 2007, Excel 2007 also allows you to save files as PDF or XPS.

12.3.3 Microsoft Office Access 2007

Microsoft Office Access 2007 provides a wide variety of professionally designed database templates and other useful features that increase the productivity of the user as well as make the user learn faster. Some of the new features/enhancements made in Office Access 2007 are as follows:

- **Layout View:** Access 2007 includes a new layout view that lets you make design changes in forms and reports while viewing them. You can add or remove a field or apply formatting without switching to the design view as in earlier versions.
- **Enhanced Quick Create by using the Create Tab:** The Create tab on the ribbon lets you create new objects such as forms, reports, tables, SharePoint lists, queries, macros, etc. quickly. When you create an object, Excel 2007 considers the active object. As a result, a new object can be created quickly based on the active object. For example, if a table is opened, a new form can be created based on that table in just two mouse-clicks.
- **Easy to Create Tables:** In Access 2007, designing a table has become faster and easier due to the improvement in Datasheet view. You can add or remove a field to/from the table in the Datasheet view, that is, without switching to Design view as required in earlier versions. Moreover, Access infers common data types such as text, number, date and time, etc. when the first data is entered in a new column. Thus, you are no longer required to specify the common data types.
- **Attachment Data Type:** Access 2007 provides a new Attachment data type that allows you to store all types of documents (pictures, graphics, Office files, binary files, etc.) in your database. You can store multiple attachments with a single record. Moreover, Access automatically compresses the attachments whenever required, thereby maximizing the memory usage.
- **Easy to Import and Export data:** In Access 2007, you can perform import and export operations efficiently that you use on a regular basis. It lets you save the details of import and export operation that is needed to repeat the operation as a specification. Moreover, while saving the specification, Access offers the option to schedule the operation.
- **Export to PDF and XPS:** Access 2007 also allows you to save files in PDF or XPS.

12.3.4 Microsoft Office PowerPoint 2007

Microsoft PowerPoint 2007 comes with new interface, new templates, and enhanced formatting options that enable you to give a professional look to your presentations. Now you can quickly and easily design an

impressive presentation that catches the attention of your audience. Some of the new features/enhancements made in PowerPoint 2007 are as follows:

- **SmartArt Graphics:** With SmartArt graphics feature, you can create designer-quality illustrations quickly and easily without the aid of a designer. PowerPoint 2007 lets you choose a type of SmartArt graphic while creating it. Each type contains several different layouts. After you choose a layout, you can easily switch layouts to try different layouts to find the best one that illustrates your information. Most of your text, text formatting, colours,styles, and effects are automatically applied to the new layout.
- **New Graphics Effects:** PowerPoint 2007 allows you to give professionally designed look to your presentation by adding 3D effects like reflection, shadow, glow, soft edges, rotation, etc. to shapes, SmartArt graphics, tables, text, and WordArt.
- **Presenter View:** One of the new exciting features offered by PowerPoint 2007 is the presenter view. Using this feature, you can deliver your presentation on two monitors; you can run your presentation on one monitor and the audience can view it on a second monitor. In this view, you can create a customized presentation for your audience by using thumbnails to select slides out of sequence.
- **Slide Libraries:** Using PowerPoint 2007 Slide Libraries, you can share and reuse the content of slides. You can store presentations as individual slides on a centralized storage (Slide Library) where you can publish your slides and reuse the slides on it.
- **Mark as Final:** Like Word 2007, in PowerPoint 2007 also, you can use the Mark as Final feature to prevent readers from making changes in a presentation by making it read-only.
- **Export to PDF and XPS:** PowerPoint 2007 allows you to save files in PDF or XPS.

Note: *Yo u can store, share, and reuse slides in a Slide Library only if Microsoft Office PowerPoint 2007 is running on your computer and is connected to a server running Microsoft Office SharePoint Server2007.*

LET US SUMMARIZE

1. Office 2007 is the latest version of Microsoft's productivity suite. Other than the applications in previous versions, it includes new applications like Groove 2007, OneNote 2007, and Communicator 2007. Some new server-side tools like Office SharePoint Server 2007 have also been added in Office 2007.
2. In Office 2007, a new interface named as *Office Fluent* has been introduced in which these menus and toolbars have been replaced with the Ribbon and Office button.
3. The Office Button ( ) is located at the top left corner of the Office program window. It provides quick access to the basic commands that are common to all Office programs like creating, opening, saving, printing, and closing a file.
4. The *Quick Access Toolbar* is a customizable toolbar located to the right of the Office button on the Office program window. By default, it contains Save, Undo, and Redo buttons. You can not

only add or remove the buttons to the Quick Access Toolbar but also move the Quick Access Toolbar below the ribbon (that is, near the work area).

5. The *Ribbon* is located just below the title bar. It is designed specially to provide quick access to commands that you need to complete a task. It consists of the commands which are organized into logical groups; these groups are further organized under tabs.
6. A *command tab* or simply *tab* consists of commands that are related to a specific type of activity. By default, the commands under the Home tab are displayed on the ribbon.
7. The commands in a command tab are organized under a group according to their functions. The group of commands related to the selected tab in the program window is known as command set. The name of the command set is displayed below the commands.
8. MS Office 2007 also provides traditional style dialog boxes for some command sets on the Ribbon. The dialog boxes for the command sets are indicated by the presence of a small arrow ( ) at the bottom right corner of a command set (known as *dialog launcher*).
9. All the available tabs do not appear on the ribbon; some tabs appear only when needed. Such tabs are called contextual tabs. These tabs become visible when you create or select a specific object in your document.
10. Office 2007 provides you KeyTips that show a way to access the commands via the keyboard. KeyTips are nothing but the indicators that display the keyboard shortcuts for the commands.
11. A *Super ToolTip* is the small window that provides information about a command along with an image. It can be considered as the enhanced form of the ToolTip as it has larger window that displays more descriptive text than a ToolTip and can have a link to a Help topic.
12. A mini toolbar is a handy, miniature, and semitransparent toolbar that appears when you select some text in an Office program. It allows you to quickly access the commands that are frequently used while working in an Office program.
13. The *Live Preview* feature enables you to see the effect of the selected option from the drop-down menu or gallery on the selected object or the current document before actually applying it.
14. Using Word 2007, you can create and format your document effortlessly than the earlier versions.
15. Microsoft Office Excel 2007 allows you to analyse, manage, and share the data easier than the earlier versions.
16. Microsoft Office Access 2007 provides a wide variety of professionally designed database templates and other useful features that increase the productivity of the user as well as make the user learn faster.
17. Microsoft PowerPoint 2007 comes with new interface, new templates, and enhanced formatting options that enable you to give a professional look to your presentations. Now you can quickly and easily design an impressive presentation that catches the attention of your audience.

Fill in the Blanks

1. In Office 2007, a new interface named as ..... has been introduced.
2. Using ....., you can also access a list of documents you have most recently worked with and a set of various program options.
3. The group of commands related to the selected tab in the program window is known as .....
4. ..... become visible when you create or select a specific object in your document.
5. The command sets under ..... tab are Clipboard, Font, Paragraph, Styles, and Editing.
6. ..... are the indicators that display the keyboard shortcuts for the commands.
7. The ..... feature enables you to see the effect of the selected option from the drop-down menu or gallery on the selected object or the current document before actually applying it.

Multiple-choice Questions

1. In Office 2007, a new interface named as *Office Fluent* has been introduced in which the menus and toolbars have been replaced with .....
    1. Ribbon
    2. Mini Toolbar
    3. Office Button
    4. Both (a) and (c)
2. Which of the following keys is used to activate the KeyTips?
    1. Ctrl
    2. Alt
    3. Shift
    4. Ctrl+Alt
3. Which of these programs do not support new file formats like PDF or XPS?
    1. Word 2007
    2. Excel 2007
    3. Access 2007
    4. None of these
4. Which of the following tab(s) appear when you insert or select a table?
    1. Design
    2. Layout
    3. Format
    4. Both (a) and (b)
5. Which of the following features makes a document read-only so that no modifications can be made to its content?
    1. Digital Signature
    2. Mark as Final
    3. Inspect Document
    4. Document Comparison

6. In Excel 2007, you can sort data up to ..... levels instead of only three levels.
    1. 10
    2. 32
    3. 64
    4. 24
7. Using ....., you can deliver your presentation on two monitors.
    1. Layout View
    2. Design View
    3. Presenter View
    4. Both (a) and (c)

State True or False

1. You cannot move the Quick Access Toolbar.
2. A mini toolbar is a handy, miniature, and semitransparent toolbar that appears when you select some text in an Office program.
3. KeyTip can be considered as the enhanced form of the ToolTip.
4. In Word 2007, you can compare two versions of a document quickly.
5. The document that has been marked as final cannot be edited.
6. In Excel 2007, you can filter data by text only.
7. Access infers common data types such as text, number, date and time, etc. when the first data is entered in a new column.

Descriptive Questions

1. How does the user interface of Office 2007 is different from the earlier versions?
2. Discuss the various components of ribbon.
3. How can you customize the Quick Access Toolbar?
4. What is the purpose of contextual tabs?
5. Why SuperToolTips are considered as the enhanced form of the ToolTip?
6. What are structured references? Explain their significance?
7. For which purpose, Presenter View is used?

ANSWERS

Fill in the Blanks

1. Office Fluent
2. Office Button
3. Command Set
4. Contextual Tabs
5. Home
6. KeyTips
7. Live preview

Multiple-choice Questions

1. (d)

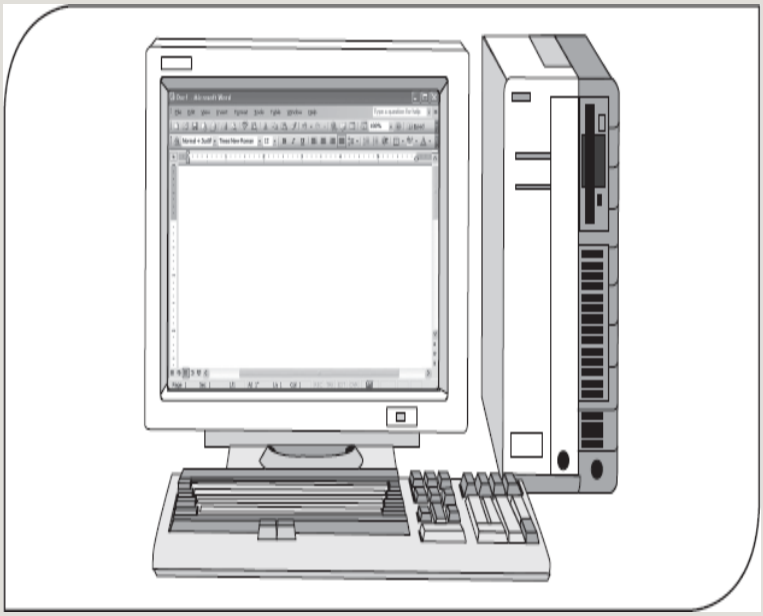
2. (b)
3. (d)
4. (d)
5. (a)
6. (c)
7. (c)

State True or False

1. False
2. True
3. False
4. True
5. False
6. False
7. True

Chapter 13

Microsoft Office Word 2003



This chapter is penned to provide a foundation to the reader to harness the power of word processing using Microsoft Office Word 2003. Word enables the user to effectively create documents quickly and efficiently. In this chapter, the reader will learn to create and format documents using various in-built features of Word. The chapter also covers various tools and features such as bullets and numbering, headers and footers, spelling and grammar checking, and many more through which text-based document can be enhanced. The chapter provides a broad discussion on inserting tables and graphics in a document. The chapter concludes with an explanation on previewing and printing a Word document.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The basic concepts of word processing with Microsoft Office Word 2003
- The steps required to create a word document from scratch
- How to select and format the text to make a document more attractive to a reader
- How tables present information in an aligned and organized manner by displaying the text in a grid of rows and columns
- How to check spelling and grammar in Microsoft Office Word 2003
- The types of graphics that Microsoft Office Word 2003 uses to add emphasis and visual impact on documents
- The different options of printing a document offered by Microsoft Office Word 2003

13.1 INTRODUCTION

Word processing is the manipulation of characters, words, text, numbers, sentences, and paragraphs in the document so that it is error-free and looks attractive. Initially, the work of preparing and formatting a document was done manually with the help of typewriter, which was time consuming and could not be modified or used again. These days, this task is performed with the help of computers, making it much easier. A word processor is an application software for word processing. It deals with character codes in the same objective way as a computer deals with number/codes. This results in reduced retyping time for error-free hard copy, facilitation of document revision, change before committing to paper, faster output speed, reduced proofreading, and elimination of stationary

wastage. Almost everyone has a need for this type of software because it can be used for common tasks.

With a word-processing program, you can create letters, tables, newsletters, and academic papers that are easy to revise and much more aesthetically pleasing than documents created by a typewriter. The various word-processor packages that are being used are Word Perfect, Word Star, and Microsoft Word, etc. Of all these word-processing packages, Microsoft Word has a wide variety of features that make it one of the most popular and widely used applications. Microsoft Word produces professional documents quickly and efficiently.

13.2 STARTING MICROSOFT OFFICE WORD 2003

To open Microsoft Word, perform *any* one of the following steps.

- Double-click the **Microsoft Word** icon located on the desktop.
- Click **start**, point to **All Programs**, point to **Microsoft Office,** and then select **Microsoft Office Word 2003** (Figure 13.1).

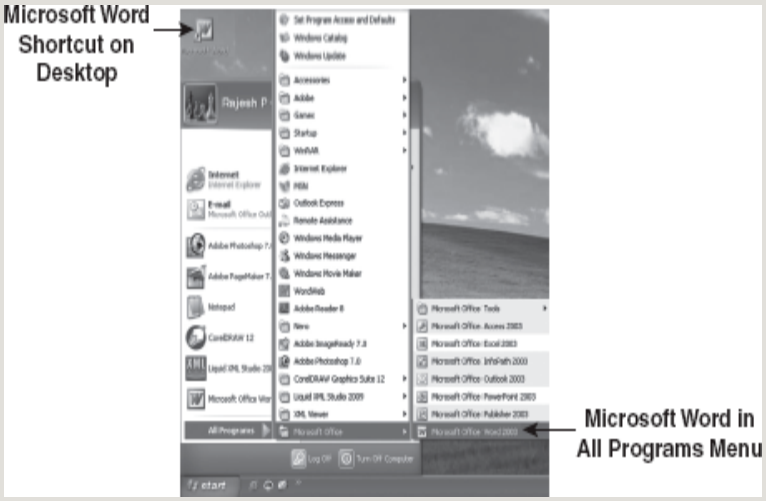


Figure 13.1 Starting Microsoft Office Word 2003

13.2.1 Microsoft Word Environment

When Microsoft Word is opened, the main screen of the program appears. This main window has different components, which are described below in detail.

Title Bar: The *title bar* is located at the top of the Word window and displays the names of the application and active document. It also tells you which Microsoft Word application is currently active. At the extreme upper-left corner of the title bar, there is a small **Word** icon ( ) (refferred as the **Control** menu). Clicking this icon displays a pull-down menu, which contains options to change the size of the window and placement of the document. On the right end of the title bar, there are **Minimize, Restore Down/Maximize,** and **Close** buttons (Figure 13.2).

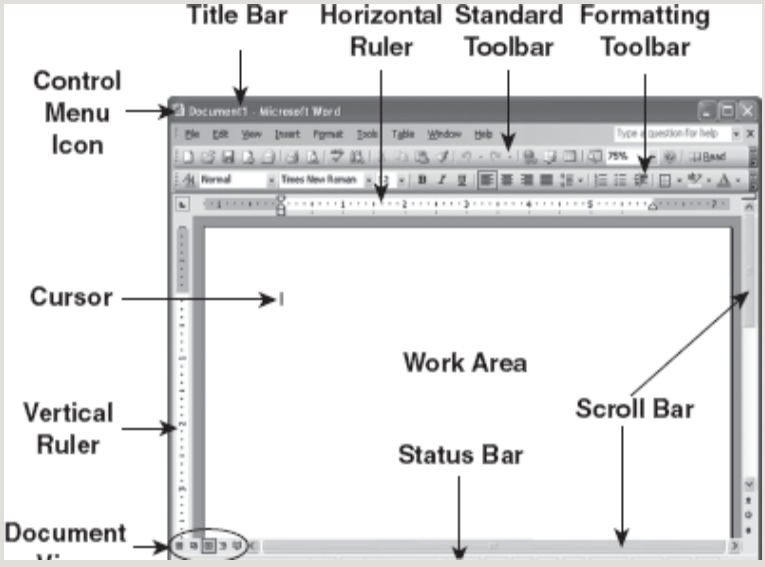


Figure 13.2 Main Microsoft Word Window

Menu Bar: The *menu bar* has many different menus, which control all options, functions, and commands for the entire Microsoft Word application. Each menu can be activated by pressing **Alt+** the underlined letter of the menu from the keyboard. By default, Microsoft Word 2003 contains nine menus, which have an associated pull-down menu. For example, the File menu contains commands to open, create, and print a document.

Toolbars: Generally, a *toolbar* is positioned just below the menu bar. Toolbar contains command buttons that provide single-click entry to carry out many text-processing tasks. Word 2003 has 21 toolbars, each having its own functions. By default, only Standard and Formatting toolbars are displayed in the Word environment. Additional toolbars like Drawing and Picture toolbars can be added by right-clicking on the menu bar and selecting the desired toolbar(s) from the pop-up menu, as shown in Figure 13.3. Alternatively, select **Toolbars** from the **View** menu, and then select the desired toolbar(s) from the pop-up menu.

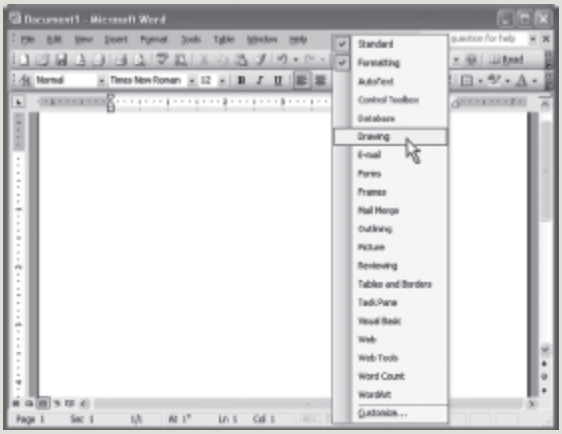


Figure 13.3 Various Toolbars

Ruler: Microsoft Word has two rulers, *Horizontal* and *Vertical*. These rulers are used to change the format of a document by adjusting the indents, margins, tab stops, and table columns. You can hide or view rulers by clicking the **Ruler** option from the **View** menu.

Status Bar: The *status bar* is located at the bottom of the Word window, and it provides information related to current cursor location, current page number, total number of pages, and program status notes (like Spelling and Grammar status, overtype mode indicators, etc.).

Work Area: The *work area* is the area in the middle of the Word window that allows you to type text, insert images, tables, and perform many more such actions. Within this area, small blinking vertical line (known as *cursor*) appears, which marks the position where the next typed character will appear in the work area.

Document View: Word can display the documents with varying levels of details through different views. The main difference between these views is your personal preference as to how you wish to work with the document. Microsoft Word provides the following views:

- **Normal View:** It displays text with character and paragraph formatting, but does not show headers and footers. This view separates pages with a dotted line, dividing the last line of one page from the first line of the next.
- **Web Layout View:** In this view, the document appears in a large font and with more space in between lines. The page layout is determined based on the computer screen rather than by the paper. This view is used to create documents for the Web.
- **Print Layout View:** This view displays a document in almost full WYSIWYG (what you see is what you get) preview mode. All character, paragraph, and document formatting options are displayed, along with headers and footers. The document is separated on-screen into what appear to be real pages.
- **Outline View:** This view displays a document as an outline with expandable and collapsible heading levels. It is used for rearranging entire sections of a document or for moving to specific section in a long document quickly.
- **Reading Layout View:** This view increases the readability of the content of the document. Pages represented in this view are designed

to fit well on the screen. It hides all toolbars except for the Reading Layout and Reviewing toolbars.

- **Document Map:** In this view, a separate pane (known as *Document Map*) that shows a list of headings in the document is displayed. When you click a heading in the document map, the corresponding content is displayed in the work area. This view helps to navigate through the document quickly and keep track of your location in it.
- **Thumbnails:** In this view, small renderings of each page in your document are displayed in a separate pane. Using this view, you can get a visual impression of the content of each page. To jump directly to a page, you can click the corresponding thumbnail image.

Note: *A pane is a part of the document window bounded by and separated from other parts by vertical or horizontal bars.*

13.2.2 Getting Help

Microsoft Office provides an exhaustive help system to provide assistance when you are experiencing difficulty using any Office application. To get help, follow *any* one of the steps given below:

- Press the **F1** key.
- Select **Microsoft Office Word Help** from the **Help** menu.
- Click the **Microsoft Office Word Help** button ( ) on the Standard toolbar.

When you perform any of the abovementioned actions, Word Help task pane appears (Figure 13.4).

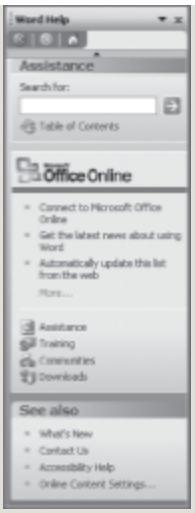


Figure 13.4 Word Help Task Pane

In this task pane, simply type your question or keyword relevant to your topic in **Search for** box and click the ( ) button. Word will search through its help database and list topics related to the question you entered. If one of the topics listed appears to be what you are looking for, click on that topic to read its contents. Word will display the help topic alongside the Word window where you can read the topic or access any of the other help information (see Figure 13.5).

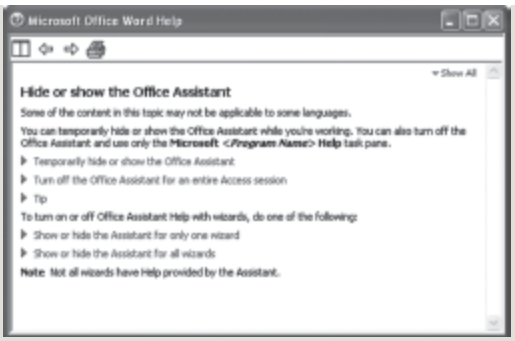


Figure 13.5 Getting Help

Once you have viewed the help information, click the **Close** button (located at the top-right corner) to close the help window and return to the document.

Note: *A task pane is a window within an office application that provides commonly used commands. You can use these commands while working on your document.*

ScreenTips: A very handy feature provided by the Office applications is the ScreenTips. A *ScreenTip* is a message that appears on the screen to provide information about a toolbar button. The tip is displayed when the user rests the mouse pointer Figure 13.6 ScreenTip on the button, as shown in Figure 13.6. To hide the ScreenTip,simply move the mouse away from the button.

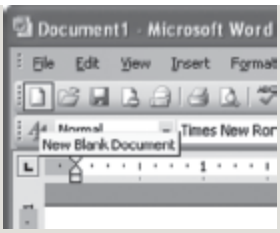


Figure 13.6 Screen Tip

13.3 WORKING WITH WORD 2003 DOCUMENTS

In this section, you will learn the steps required to perform the following in Word 2003.

- Creating a new document
- Opening an existing document
- Saving a document and making a backup copy
- Closing the document and Word application

13.3.1 Creating a New Document

Whenever you start Word, it opens a new untitled document window so that you can begin a new project. If Word is already running, and you want to create a new document, click on **New Blank Document** button ( ) on the Standard toolbar. A new document can also be opened by following the steps given below:

- Select **New** from the **File** menu to display the New Document task pane.
- Select **Blank Document** ( ) under the **New** section. Word opens a new document, which is sequentially numbered like *Document1, Document2,* and so on. Similarly, you can also create an XML document, a Web page, an E-mail message, and a document from the existing document by selecting appropriate options (Figure 13.7).

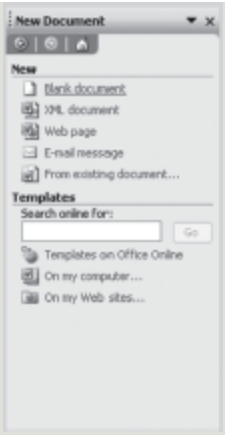


Figure 13.7 New Document Task Pane

13.3.2 Opening an Existing Document

To open an existing document, follow *any* of the steps given below:

- Select **Open** from the **File** menu to display the Open dialog box. You can also open the Open dialog box by clicking on the **Open** button ( ) on the Standard toolbar. The Open dialog box allows you to specify the name of the desired file in the **File name** drop-down box. If the desired file is not in the current location, you can locate it by navigating through **Look in** drop-down box. Once the file is found, select the file and click the **Open** button.
- Word document can also be opened by double-clicking the Word file icon, placed on the specified location (Figure 13.8).

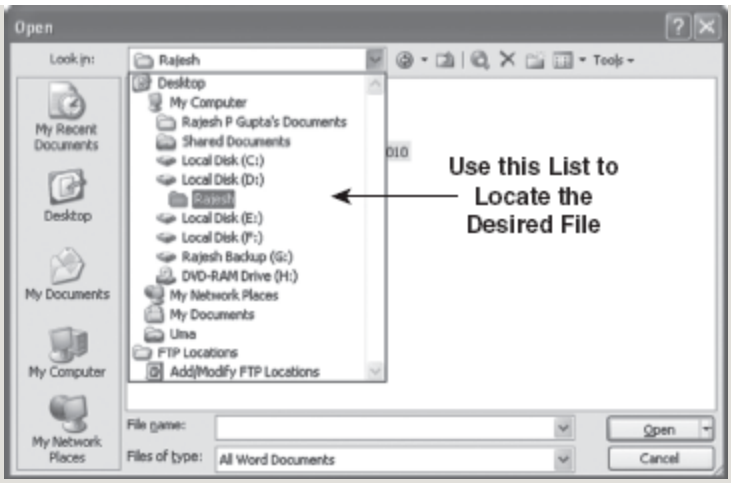


Figure 13.8 Open Dialog Box

THINGS TO REMEMBER

Saving a File

Once you have saved the new document, clicking on the **Save** button does not display any dialog box because Word assumes that you do not want to specify the name or the location of the document again. Word assumes that you just want to update the originally saved copy by overwriting it with this new version of the same file. If you do want to put a copy of this same file in a new location, create a new copy with a new name, or change the file type, you must go to the **File** menu and choose **Save As**. This way, you can save your file at different stages and keep multiple versions on your hard drive.

13.3.3 Saving a Document

Once a document is created, you can start typing the text and insert graphics and tables. When this is done, you must save the document for future references. To save the document, Microsoft Word provides two menu options, namely, **Save** and **Save As**. Although both these commands have similar functions, yet they differ in one aspect. When you create a new document and click the **Save** button ( ) on the Standard toolbar or choose **Save** from **File** menu, the Save As dialog box is displayed. This dialog box allows you to specify the name of the file, where to save the file, and the file type of the document (whether to save it as a web page or in rich text format) (Figure 13.9).

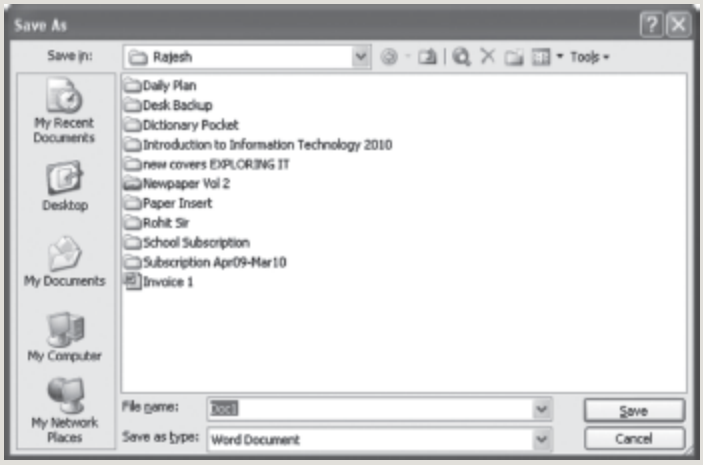


Figure 13.9 Save As Dialog Box

In the Save As dialog box, you can use the drop-down box at the top, labelled **Save in**, to select the location where the new copy is to be saved. You can also change the name of the copy by typing new name in **File name** box and then save it either in a new location or in the

same location as the original (because this new copy now has a different name). Furthermore, you can change the file type by selecting a file type from the **Save as type** drop-down box at the bottom of the Save As dialog box.

Note: *By default, Microsoft Word documents are saved with .DOC extension, although it can also be saved in other file formats like .TXT, .RTF, and .HTML*

13.3.4 Closing a Document and Exiting Microsoft Word

After finishing all the word-processing activities, you may want to close the document and exit Word. For this, follow *any* of the steps given below:

- Select **Close** from **File** menu to close the document and select **Exit** from **File** menu to close the Word program (Figure 13.10).

Figure 13.10 Closing a Document and Exiting Microsoft Word

- Click on the **Close Window** button at the upper-right corner of the Word window to close the current document. Note that when a single document is opened in Word, then there are two close buttons. One is located at the top of the application window, which closes the Word application while the second (lower) one present on the document closes only the current file. If a user makes changes to a file and has not saved those changes, Word will ask whether to save the changes before closing the file.

13.4 WORKING WITH TEXT

In the previous section, we discussed how to create and save a Word document. Now let us learn how to enter, select, and format text. Apart from these activities, you will also learn how to align and indent the text, copying and moving text, finding and replacing words, undo and redo operations, making lists, and inserting headers and footers.

FACT FILE

The tab is designed to make it easier to align text easily and precisely. Each time the **TAB** key is pressed, the cursor moves 0.5" along the line of text by default. When the **TAB** key is pressed, the text to the right of the cursor is moved to the right.

13.4.1 Inserting Text

Before Word can perform operations on text such as moving and copying, it is necessary that some text is present or typed in a document. To enter the text, just create or open a document and start typing. Note that the text appears where the blinking cursor (I) is located. If you want to type in capitals (like ITLESL instead of itlesl), activate the **Caps Lock** key by pressing it once. You can also use the **Shift** key to type the text in capitals.

In Word, there are two typing modes, namely, *insert mode* and *overtyping mode*. In the insert mode, the newly typed text does not replace the already written text, instead the existing text is pushed forward. On the other hand, in overtyping mode, the newly typed text replaces the already written text. To check whether you are in the insert mode or not, look at the status bar located at the bottom of the screen. If the letters **OVR** are in grey, you are in the insert mode and if they are in black, you are in the overtype mode (Figure 13.11).

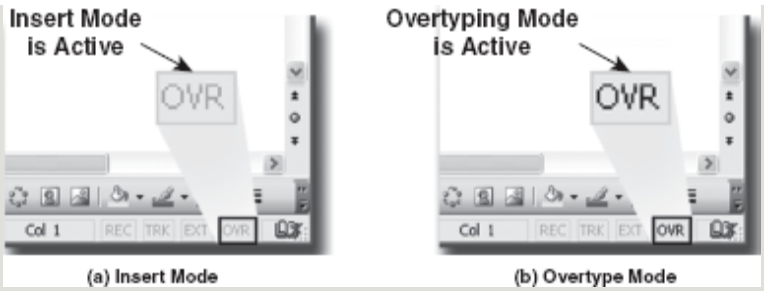


Figure 13.11 Typing Modes

By default, the insert mode is automatically activated. To change to the overtyping mode, double-click the letters **OVR** on the status bar. You can toggle between insert and overtype mode by double-clicking it again and again. Let us consider an example of inserting text in the overtyping mode. Suppose, you want to edit the content in a document from "Jenny has a large blue car" to "Jenny has a large pink car," follow the steps given below:

1. Type the text "Jenny has a large blue car."
2. Place the cursor before the letter "b" in "blue."
3. Enable the overtype mode.
4. Type the word pink.
5. The sentence should now read—"Jenny has a large pink car" (Figure 13.12).

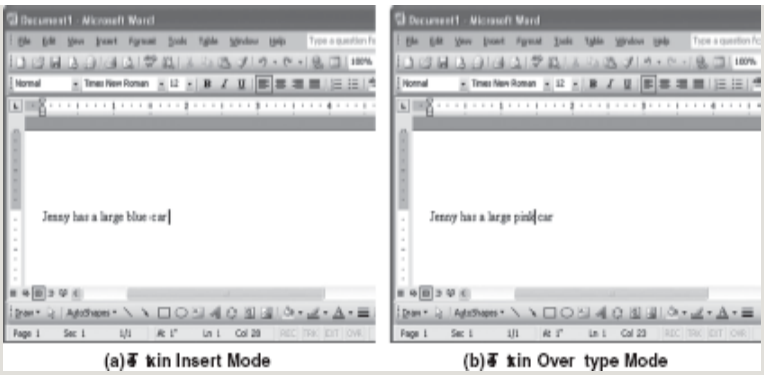
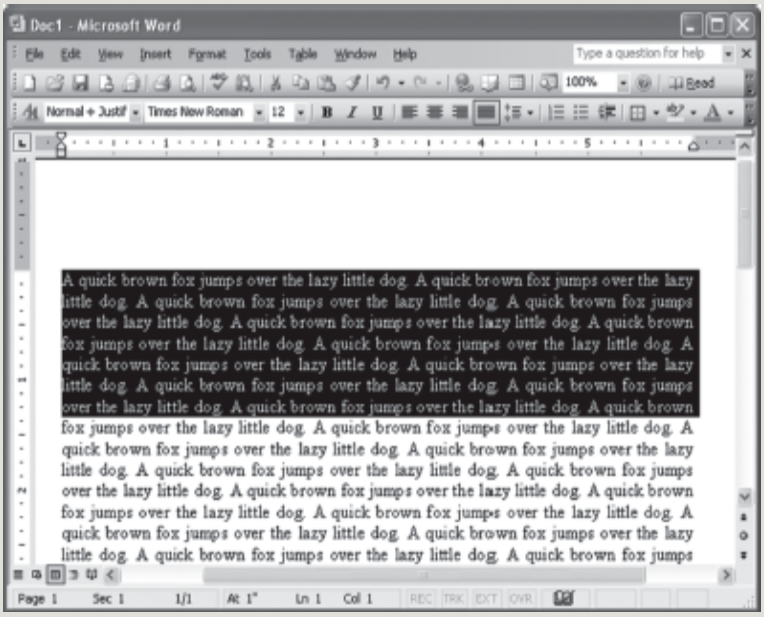


Figure 13.12 Using Overtyping Mode

13.4.2 Selecting Text

Before performing any operation on text such as moving, copying, formatting, etc., the text must be selected. Whenever you select a text, it gets highlighted (Figure 13.13).


Doc1 - Microsoft Word
File Edit View Insert Format Tools Table Window Help
Type a question for help
Normal + Justif
Times New Roman
12
100%
Read
A quick brown fox jumps over the lazy little dog. A quick brown fox jumps over the lazy little dog. A quick brown fox jumps over the lazy little dog.

Depending on your requirement, you can select a word, a line, or multiple lines, etc. Selection can be done either using a mouse or keyboard shortcuts as listed in Table 13.1.

Table 13.1 Selecting Text

To Select	Mouse Action	Keyboard Shortcuts
A Word	Double-click anywhere on the word	**Ctrl+Shift+Right** Arrow or **Ctrl+Shift+Left** Arrow
A Line	Click on the left margin next to the line. Note that the mouse cursor will change to arrow	**Shift+End or Shift+Home**
Multiple Lines	Click on the left margin and keep on dragging till the desired lines are selected	**Shift+Up** Arrow or **Shift+Down** Arrow
Paragraph	Triple-click anywhere in the paragraph	**Ctrl+Shift+Up** Arrow or **Ctrl+Shift+Down** Arrow
Entire Document	Triple-click in the left margin	**Ctrl+A**
Rectangular Area	Press **Alt** key and then drag the mouse	**Ctrl+Shift+F8** and use Arrow keys

13.4.3 Formatting the Text

As you enter and edit a text, you can change its appearance to add emphasis and make the document easier to read. This is known as *formatting*. It is the process of determining how the typed text will appear on a printed page. Formatting is applied by altering the appearance of text by setting the typeface (font), size, line spacing, and colour. Formatting can be done by using the Formatting toolbar or by using the **Format**

menu (Figure 13.14). To format a text, follow the steps given below:

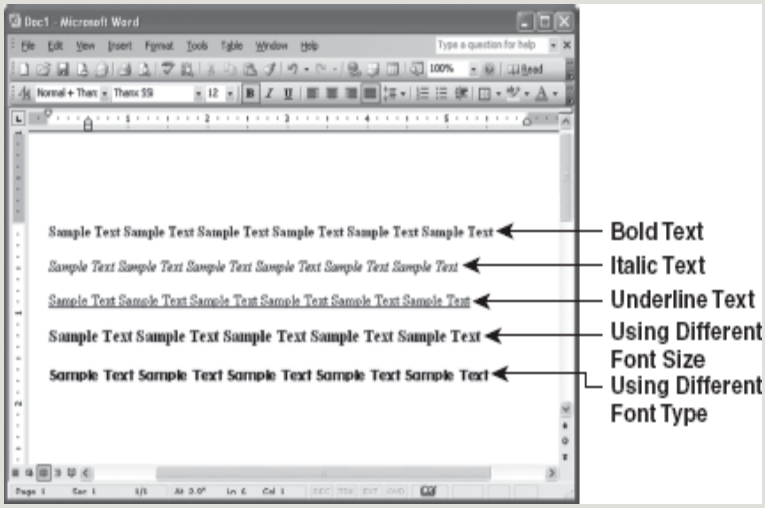


Figure 13.14 Formatting Text

- Select the text to be formatted.
- Use buttons on the Formatting toolbar to format the selected text. Some of the buttons used in formatting text are given in Table 13.2.

Table 13.2 Formatting Text

Command	Button	Description
Font		Changes the font of the selected text
Font Size		Changes the font size of the selected text
Bold		Bolds the selected text
Italic		Italicizes the selected text
Underline		Underlines the selected text
Highlight		Highlights the selected text
Font Color		Changes the font colour of the selected text

Text can also be formatted using the Font dialog box. For this, follow the steps given below:

1. Select the text to be formatted.
2. Select **Font** from the **Format** menu to menu to display the Font dialog box.
3. From this dialog box, the font, font style, and text size can be specified. This dialog box also allows you to apply underline, colour, and other effects to the selected text.
4. Note that formatting is displayed in the **Preview** box at the bottom of the Font dialog box. Once formatting is done, click **OK** to apply the formatting changes to the selected text (Figure 13.15).

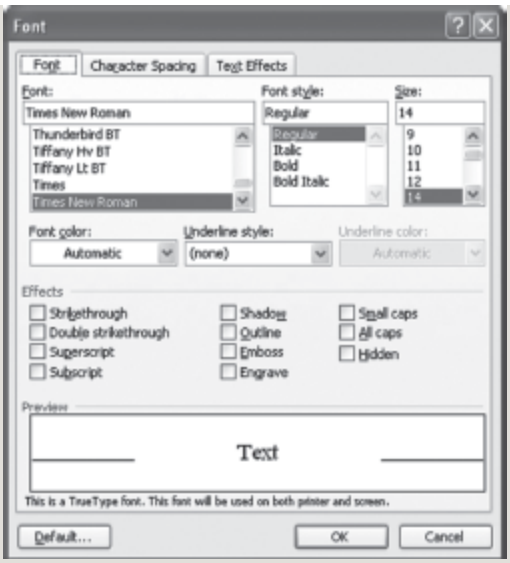


Figure 13.15 Font Dialog Box

Highlighting and Changing Text Colours: While working in Microsoft Office applications like Word, Excel, and PowerPoint, you may wish to change the attributes of some important phrases or words. This can be done by changing the font size, making them appear bold, italic, or underlined. Apart from these formatting techniques, Microsoft Office also allows you to emphasize important words or sentences by changing their font and/or background colour.

To highlight the text, just as you would do with a highlighter pen in the case of textbooks, Microsoft Office provides you with a *Highlight* tool. To highlight the existing text, follow the steps given below:

1. Click on the down arrow beside the **Highlight** button ( ) on the Formatting toolbar (Figure 13.16).

Figure 13.16 Highlight Colours

2. Select the desired colour from the drop-down list to highlight text.
3. When you move on to the document, the mouse pointer changes to ( ). Now, click and drag the mouse to select the text to be highlighted.
4. After finishing with text highlighting, click on the **Highlight** button again to turn it off, or simply press the **Esc** key.

By default, Microsoft Word displays and prints the text in black. In case you want to change the colour of the text, follow the steps given below:

1. Select the text for which you want to change the colour.
2. Click on the down arrow beside the **Font Color** button ( ) from the Formatting toolbar.
3. Select the desired colour from the pull-down list for the selected text (Figure 13.17).

Figure 13.17 Font Colours

Text Manipulation with Scientific Notation: You might have observed various mathematical expressions like $5x^3 + 4x^2 + 7x + 6$ and chemical formulas like CO_2. In the expression $5x^3 + 4x^2 + 7x + 6$, the digits 3 and 2

appears slightly above the normal baseline, and in the formula CO_2, the digit 2 appears slightly below the baseline. A character appearing above the baseline is known as a *superscript* and a character appearing below the baseline is known as a *subscript*. Subscript and superscript formats can be easily applied to the text in the Word.

To apply a subscript or a superscript format to a character (or text) in a document, follow these steps:

1. Select the character (or text) on which subscript or superscript format is to be applied.
2. Select **Font** from the **Format** menu to display the Font dialog box (see Figure 13.18).

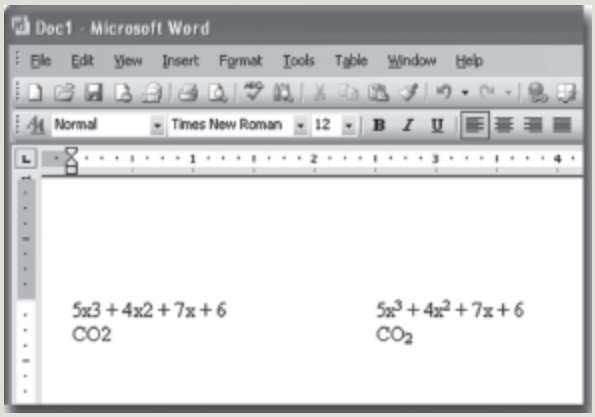


Figure 13.18 Changing Text to Subscript and Superscript

3. In the **Effects** section, select the **Superscript** checkbox or the **Subscript** checkbox according to the requirement.
4. Click **OK** to close the Font dialog box and apply the formatting changes to the selected text.

Inserting Symbols: Special characters that generally cannot be typed from the keyboard are known as *symbols*. These are the characters like summation symbol (Σ), alpha (α), pi (π), etc. that are mostly used in mathematical and scientifc formulas. Word also allows

you to insert many more symbols other than mathematical and scientifc symbols in a document.

To insert a symbol in a document, follow these steps:

1. Select **Symbol** from the **Insert** menu to display the Symbol dialog box in which **Symbols** is the active tab (Figure 13.19).

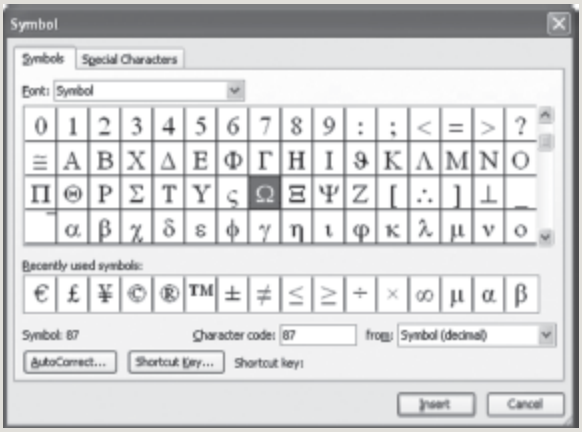


Figure 13.19 Inserting a Symbol

2. Select the desired symbol from the **Font** drop-down box and click **Insert**. The selected symbol is inserted in your document. After insertion, the **Cancel** button changes to **Close** button.
3. Click **Close** to close the Symbol dialog box.

13.4.4 Aligning and Indenting Text

Alignment and indentation operations are a part of paragraph formatting. These operations are performed to give a well-defined "structure" to the document. This gives the document a professional look (Figure 13.20). To align or indent a paragraph, follow the steps given below:

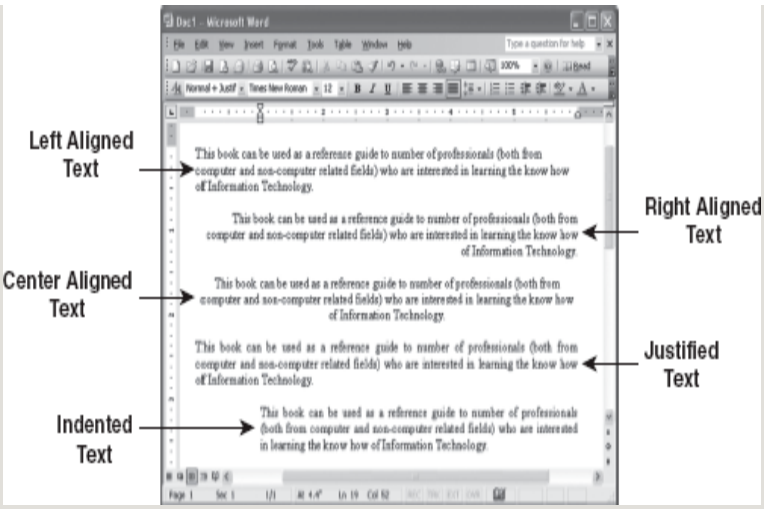


Figure 13.20 Paragraph Formatting

1. Select the paragraph to be aligned or indented.
2. Use buttons on the Formatting toolbar to align or indent the paragraph. Some of the buttons used in aligning or indenting paragraph are given in Table 13.3.

Table 13.3 Paragraph Formatting

Command	Button	Description
Align Left		Left aligns the paragraph or selected text
Center		Centre aligns the paragraph or selected text
Align Right		Right aligns the paragraph or the selected text
Justify		Justifies the paragraph or selected text
Decrease Indent		Decreases the indent of the selected paragraph
Increase Indent		Increases the indent of the selected paragraph

A paragraph can also be formatted using the Paragraph dialog box. For this, follow these steps:

1. Select the paragraph to be formatted.
2. Select **Paragraph** from the **Format** menu to display the Paragraph dialog box (Figure 13.21).

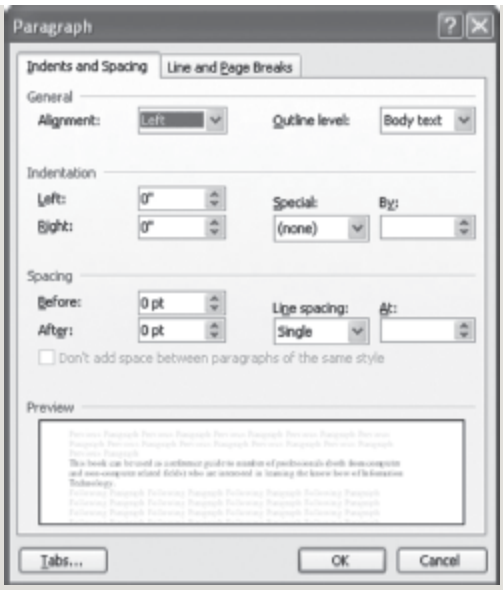


Figure 13.21 Paragraph Dialog Box

3. Specify the desired alignment and indentation for the selected paragraph.
4. Note that the formatting is displayed in the **Preview** box at the bottom of the Paragraph dialog box. Once formatting is completed, click **OK** to apply the formatting changes to the selected paragraph.

FACT FILE

Style

A *style* is a set of formatting specifications that can automatically be applied to the text. In smaller documents, you can alter the appearance of text by changing the font size and colour, adjusting paragraph settings, aligning and indenting text. These settings can be applied one at a time by clicking toolbar buttons and using the keyboard or dialog boxes. Instead of applying these settings one at a time, a combination of different character and paragraph formats can be incorporated into a style.

13.4.5 Formatting Style

Formatting the text using various menu options and toolbar buttons is viable for small documents with only a few pages. However, when a document contains many pages, or when consistent formatting is crucial, you should use *Styles* instead.

Styles can significantly reduce the time it takes to format a document. For example, if a document is divided into sections with headings (like this book has), it is a good idea to use heading styles. It is much easier to maintain consistency in a document if styles are used. For example, if a heading style is used in a document and you want to change the appearance of the heading, just alter the definition of the style. This change will be cascaded through all headings in the documents that were created using the style, so you do not have to select each one and make the formatting changes individually.

To apply a style, follow the steps given below:

1. Place the cursor in the paragraph where the style is to be applied.
2. Click the **Style** drop-down box on the Formatting toolbar and select a style by clicking on it (Figure 13.22).

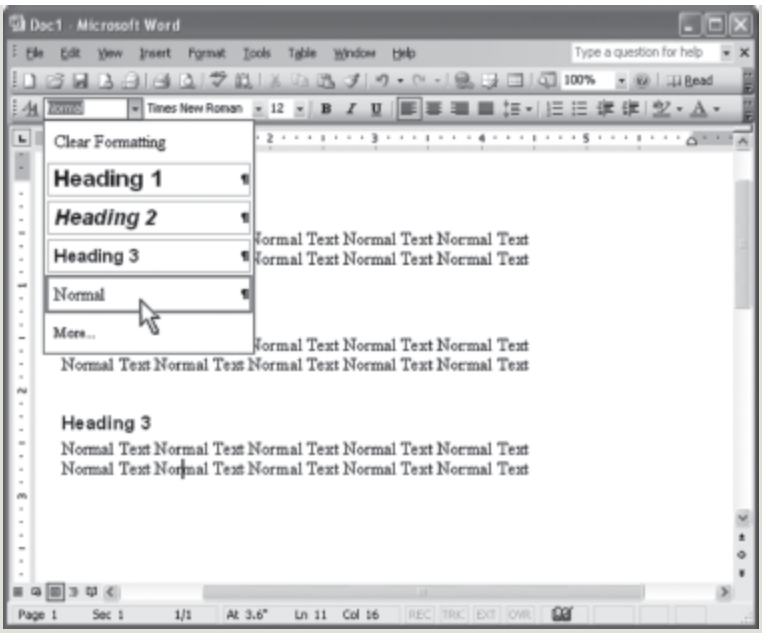


Figure 13.22 Style Box

Accessing All Available Styles: For simplicity, Microsoft Word 2003 only displays a small number of the built-in styles in the Style box of a new document. If you want any of the other styles to be available, you have to open the Style dialog box. This can be done by following the steps given below:

1. Select **Styles and Formatting** from the **Format** menu to display the Styles and Formatting task pane.
2. Select **All styles** from **Show** drop-down box to display all the styles (Figure 13.23).

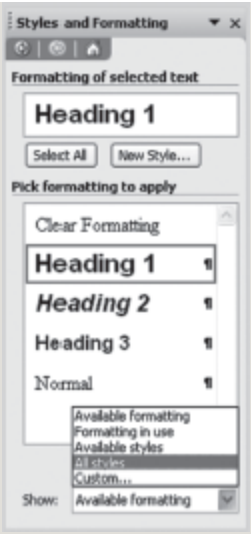


Figure 13.23 Styles and Formatting Task Pane

3. Click on the style you want to apply on the text.

13.4.6 Format Painter

A handy feature for formatting text is the *Format Painter*. For example, if you are formatting a paragraph heading with a certain font face, size, and style, and you want to format another heading the same way, you need not manually add each attribute to the new headline. Instead, you can use the Format Painter, which is a time-saving feature for quickly copying the format of a block of text. This can be done by following the steps given below:

1. Select the text containing the formatting you wish to copy.
2. To copy formatting to a single location, click once on **Format Painter** button ( ) on the Standard toolbar. To copy it to multiple locations, double-click on **Format Painter** button. Notice that the mouse pointer changes to resemble a paintbrush ( ).
3. Now click and drag the mouse to select the text you wish to format, and then release the mouse.
4. If the format painter is copied by double-clicking, press the **Esc** key to come out of painter mode.

13.4.7 Using Cut, Copy, and Paste in Word

When you are working on a document, sometimes you may decide to move a certain portion of text (or even images), or copy the same text or image into some other area or in some other document. If the text is to be moved within the same document or to another document, it should be first cut and then pasted. When you choose to cut the text, Word removes it from the current document and places it in the clipboard, a temporary storage area in computer memory. Now to place the text stored in the clipboard, you will have to perform the paste action. Copying the text is similar to the cut operation, except that copy retains the text at the original place, but in case of cut, the text is removed from the original location. To perform the cut, copy, and paste, follow the steps given below:

1. Select the text that is to be moved or copied.
2. In case you want to move the text, choose **Cut** from the **Edit** menu or click the **Cut** button ( ) on the Standard toolbar. If you want to copy the text, choose **Copy** from the **Edit** menu or click the **Copy** button ( ) on the Standard toolbar.
3. Click the mouse on the place in the document where you want to insert the text.
4. Choose **Paste** from the **Edit** menu or click the **Paste** button ( ) on the Standard toolbar. The text that you copied to the clipboard is pasted to the place where the mouse is clicked.

13.4.8 Finding and Replacing Text

Find and replace option in Microsoft Word is used to locate a particular word, phrase, or set of characters in a document, and replace it with something else. Although you could manually scan through the document for find and replace purposes, it takes time and there is always the possibility that you might miss the text in question on a first or second pass. Microsoft Word's Find and Replace feature swiftly and unerringly locates anything you are looking for, and once the desired text is located, it can automatically be replaced by the new text. For example, if you have used word, say “January 2004” in

the document and later want it to be replaced with "January 2005," Word finds all instances of "January 2004" and replaces it with "January 2005." Microsoft Word also finds and replaces all grammatically wrong word(s). For example, if you want to replace the word "narrow" with the word "broad" throughout a document, Word intelligently changes not only "narrow" to "broad," but also words like "narrower" to "broader."

To find and replace a text within the document, follow the steps given below:

1. Select **Find** from the **Edit** menu to display the Find and Replace dialog box.
2. Click on the **Find** tab if you want to find the desired word in the document, or click **Replace** tab to find a word and replace it with new word.
3. Type the text that you want to search in the **Find what** box.
4. Click **Find Next** button to find the first occurrence of the text. If you want to replace the text, then select the **Replace** tab and type the replacement text in the **Replace with** box (Figure 13.24).

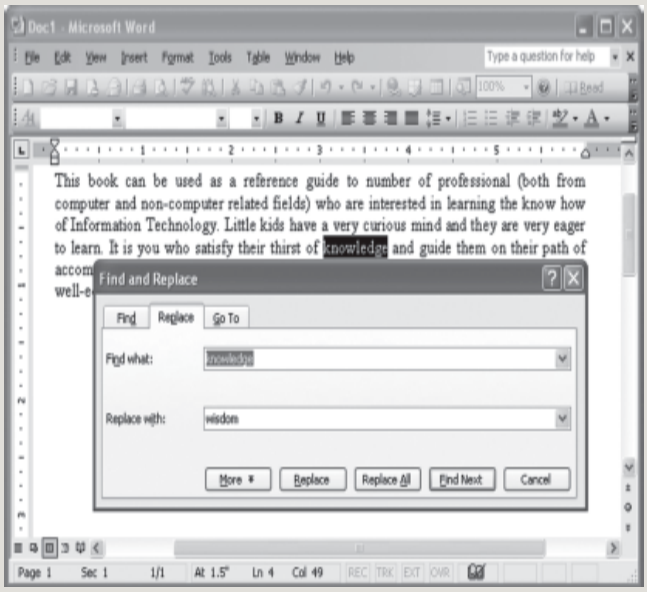


Figure 13.24 Find and Replace Dialog Box

5. Select **Replace** button to replace each occurrence of the word individually or click **Replace All** to replace all occurrences of the word at once. Word will display a message when it has replaced all occurrences.
6. Keep on pressing the **Find Next** and **Replace** button until you are finished with finding and replacing the text.
7. To close the Find and Replace dialog box, click the **Cancel** button or press the **Esc** key.

13.4.9 Undo and Redo

Microsoft Office is a very powerful suite of applications. These applications remember all the actions that you performed on a file. If a portion of a document is deleted or changed, you can revert to the original state by using the **Undo** command. This feature instructs the application to ignore the last action (whether it was deleting, copying, or moving). Word 2003 has virtually unlimited capacity of performing undo action, as long as you do not close the document. However, certain actions like saving a file or printing a page cannot be undone.

To undo the last action(s), click the **Undo** button ( ) on the Standard toolbar, or select **Undo** from the **Edit** menu. If you want to undo a number of actions at the same time, then click the down arrow ( ) beside the **Undo** button to display a list of actions that can be undone [see Figure 13.25(a)].

Figure 13.25 Undo and Redo Actions

If an undo action is set and then you want to reverse it, then **Redo** command can be used to reverse the undo action. To redo the last undo action, click on the **Redo** button ( ) on the Standard toolbar, or select **Redo** from the **Edit** menu. If you want to redo a number of actions at the same time, then click the down arrow ( ) beside the **Redo** button to display a list of actions that can be redone [see Figure 13.25(b)].

13.4.10 Bullets and Numbering

In many situations, you need to prepare a list of actions. For example, while writing an algorithm, it is always advisable to use bullets and numbering to list all the steps. If these steps are written in a paragraph form, separated only by colons or commas, it may become difficult to understand the text. Microsoft Word 2003 supports list-makers with its *bullets and numbering* features. To use bullets and numbering, first select the text you want to format. Next, use the **Bullets** ( ) or **Numbering** ( ) button on the Formatting toolbar to apply bullets or numbering (Figure 13.26).

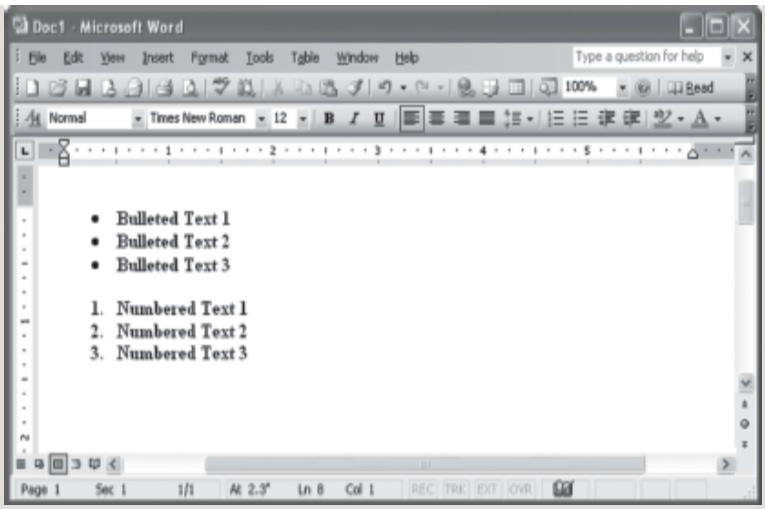


Figure 13.26 Bulleted and Numbered Text

Bullets and numbering can also be applied to text by selecting **Bullets and Numbering** from the **Format** menu. This will display the Bullets and Numbering dialog box (see Figure 13.27). To apply bullets, use the **Bulleted** tab, and if a numbered list is to be made, use the **Numbered** tab. There are seven default bullet formats (and seven default numbering formats) available in the dialog box. You can modify any of these to create your own formats.

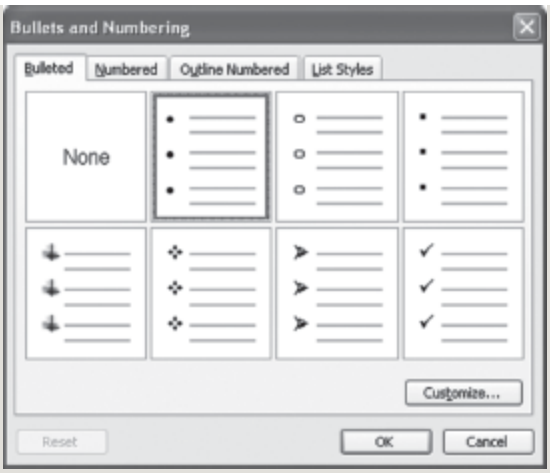


Figure 13.27 Bullets and Numbering Dialog Box

If you want to remove the bullets or numbering, select he text and click the **Bullets** or **Numbering** button on the Formatting toolbar.

13.4.11 Header and Footer

Headers are text that appears at the top of every page, and *footers* contain text that appears at the bottom of each page. Every Word document contains header and footer areas where you can put a document title, author name, page number, graphics, and any other essentials that you want to repeat over several (or all) pages in the document. If nothing is inserted in the headers and footers, they still exist, but are left blank and take up no space on the page. What you put in the header or footer appears on every page of the document.

To create a header and footer in a document, follow the steps given below:

1. Select **Header and Footer** from the **View** menu. This displays the Header and Footer toolbar with a dotted box for header at the top of the page and a dotted box for footer at the bottom of the page (Figure 13.28).

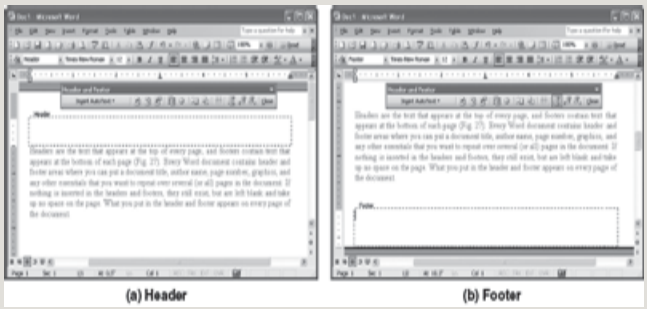


Figure 13.28 Header and Footer

2. Type the text in Header and/or Footer box. Note that headers and footers can be formatted (like changing font colour, size, etc.) just like any other text.
3. If necessary, use appropriate options from the Header and Footer toolbar to add time, insert page number, date, etc., as listed in Table 13.4.

Table 13.4 Header and Footer Toolbar Options

Command	Button	Description
Insert Page Number		Inserts page numbers that automatically update when pages are added or deleted
Insert Number of Pages		Prints the total number of pages in the active document
Format Page Number		Formats the page numbers in the current section
Insert Date		Inserts date that automatically updates so that the current date is displayed
Insert Time		Inserts time that automatically updates so that the current time is displayed
Switch Between Header and Footer		Switches the insertion point between the header and footer areasorMoves the insertion point from header to footer area or vice versa

4. After the text or graphic has been entered, click **Close** button on the Header and Footer toolbar or double-click in the work area to return to work there.

13.5 WORKING WITH TABLES

Tables are one of the most powerful features of Microsoft Word. They provide a means of presenting information in an organized and attractive manner that displays text in a grid of rows and columns of cells.

In Microsoft Word 2003, tables can be created in number of ways.

- Select **Insert** from the **Table** menu and then choose **Table** from the pop-up menu to display the Insert Table dialog box, from where you can choose the number of columns and rows required in the document. Note that you can always add or delete columns and rows at a later stage (Figure 13.29).

Figure 13.29 Insert Table Dialog Box

- Another way of creating a table is to click the **Insert Table** button ( ) on the Standard toolbar. Drag the mouse along the grid, highlighting the number of rows and columns for the table (Figure 13.30).

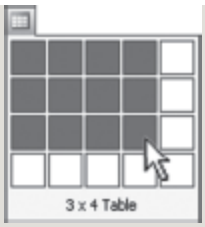


Figure 13.30 Inserting Table by Dragging

When you select any of the above two methods, the table will be inserted into the document as shown in the Figure 13.31.

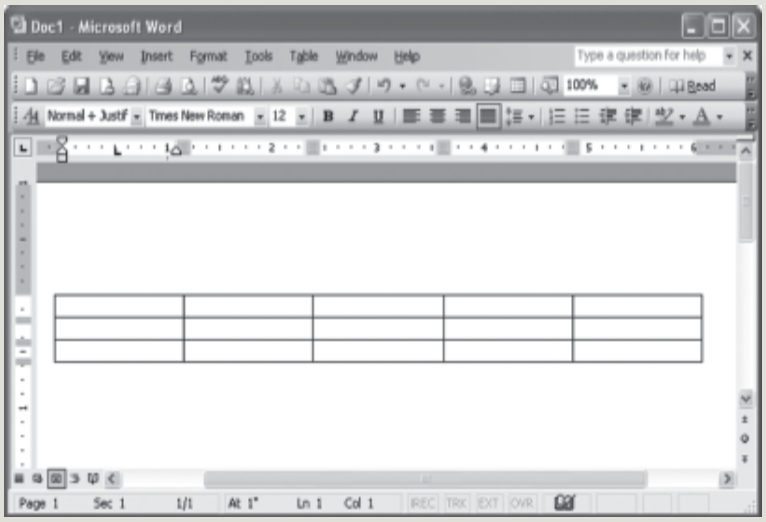


Figure 13.31 Creating a Table

- A table can also be drawn onto the document. To draw a table, select **Draw Table** from the **Table** menu or click the **Tables and Borders** button ( ) on the Standard toolbar. The cursor is now the image of a pencil ( ) and the Tables and Borders toolbar has appeared. Draw the

cells of the table with the mouse. Figure 13.32 displays an example of creating a table using the drawing method.

After creating the table, enter the text into the table by pointing to the cell. Press the **Tab** key to move to the next cell. When you reach the last cell of the table you can press the **Tab** key again to create a new row in the table. To move to a previous cell press **Shift+Tab.**

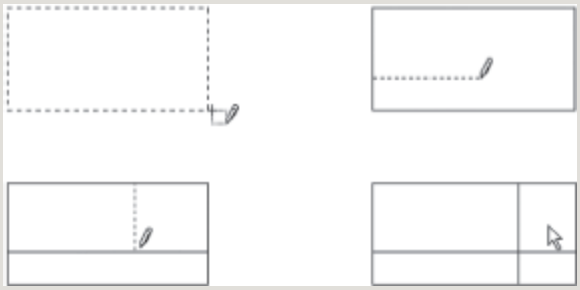


Figure 13.32 Drawing a Table

Tables and Borders Toolbar: The *Tables and Borders* toolbar allows you to add border styles, shading, text effects, alignment, and more options to your table (Figure 13.33).

Figure 13.33 Tables and Borders Toolbar

Table 13.5 lists all the available buttons on the Tables and Borders toolbar.

Table 13.5 Tables and Borders Toolbar Options

Command	Button	Description
Draw Table		Inserts a table by dragging onto the document
Eraser		"Removes any line in the table. This is mainly used to merge cells by removing the dividers between them
Line Style		Defines the style and thickness of the table border
Line Weight		Defines the width for table's border
Border Color		Defines the colour of the table's border
Outside Border		Applies line style, line weight, and border colour settings to the outside border of the cell with the insertion point or to any selected part of the table
Shading Color		Opens a palette of colours from where you can select a colour to apply it as a background shading to the cell containing the insertion point or to any selected part of the table
Insert Table		Displays the Insert Table dialog box
Merge Cells		It is only available if two or more cells are selected. Use it to remove the dividing borders of the cells and merge them into a single cell
Split Cells		Splits a cell into more than one cell
Align Top Left		Defines the horizontal and vertical alignment of the text within a cell
Distribute Rows Evenly		Makes the selected rows or cells to equal row height
Distribute Columns Evenly		Makes the selected columns or cells to equal column width
Table AutoFormat		Opens the Table AutoFormat dialog box from where you can choose from predetermined table styles

Command	Button	Description
Change Text Direction		Toggles the direction of text in a cell between vertical and horizontal
Sort Ascending		Sorts the rows in a table in ascending order
Sort Descending		Sorts the rows in a table in descending order
AutoSum		Adds the content of a range of cells

Selecting Parts of a Table: For working with the tables, one must have to select the different parts of the table. Selection allows formatting or deleting different parts of the table. Moreover, it also allows adding extra rows and columns in the table.

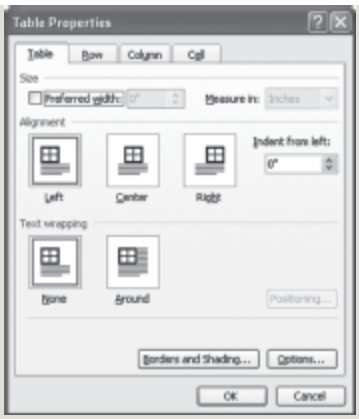
Different parts of a table can be selected either by using mouse or by using the shortcuts listed in Table 13.6.

Table 13.6 Shortcuts Used to Select a Table

For Selecting	Using Table Menu	Using Mouse
Single Cell	Position the cursor in the cell. Click the **Table** menu, point to **Select**, and then click **Cell**	Click the bottom-left corner of the cell when a black arrow appears
Single Row	Position the cursor in the row. Click the **Table** menu, point to **Select** and then click **Row**	Click outside the table to the left of the row when a white arrow appears
Single Column	Position the cursor in the column. Click the **Table** menu, point to **Select** and then click **Column**	Click outside the table above the column when a black down arrow appears
Multiple Rows	None	Click outside the table to the left of the row and drag the mouse down
Multiple Columns	None	Click outside the table above the column and drag the mouse to select the desired columns
Whole Table	Position the cursor in the table. Click the **Table** menu, point to **Select** and then click **Table**	Hold down the **Alt** key. Double-click inside the table when the mouse looks like an I bar

13.5.1 Table Properties

The *Table Properties* dialog box is used to modify the alignment of the table with the body text and the text within the table. This dialog box can be displayed by right-clicking in any cell and choosing **Table Properties** from the pop-up menu. This dialog box shows the current settings for the selected table. These settings appear as options in four different tabs: *Table, Row, Column,* and *Cell* (Figure 13.34).



w

Figure 13.34 Table Properties Dialog Box

Click on **Table** tab to format size, alignment, and text wrapping. The **Row** tab is used to change the cell's height and alignment. To format the columns, just click on the **Column** tab. Here you can set the width of each column according to the requirement. Every column can be of the same width, or of different width. Note that the formatting will be applied only to the sections that have been selected. When you finish making changes to the table, click on **OK** button to close the dialog box.

13.5.2 Moving and Resizing Table

When you place the mouse over the table, a move handle and a resize handle will appear on the top-left and bottom-right corners of the table respectively. Click and drag the move handle to move the table and release the mouse button when the table is positioned where you want it. Click and drag the resize handle to resize the table. You can also change the column width and row height. Move the mouse pointer onto the column line at the right of the column whose width you wish to alter. The pointer will change to ( ). Click and drag this to the right to increase the column width or to the left to

decrease the column width. Similarly, you can also change the height of cells on a row or rows within the table by dragging the row resizer ( ) in upward or downward direction (Figure 13.35).

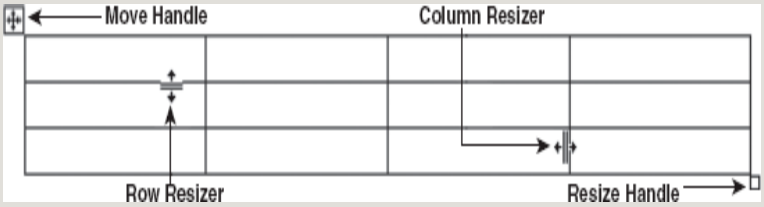


Figure 13.35 Moving and Resizing Table

13.5.3 Formatting Text in Table

Like word paragraphs, text stored in table cells can also be formatted. All the usual formatting options, such as text alignment, bullets and numbering, indentation, spacing, tab stops, font colors, font styles, borders and shading, etc., work well within tables and cells. To make the text stored in a row of the table bold and italic, follow the steps given below:

1. Select the row to be formatted (Figure 13.36).

Figure 13.36 Row Selected in a Table

2. Click appropriate buttons on the Formatting toolbar to format the text in the selected row (Figure 13.37).

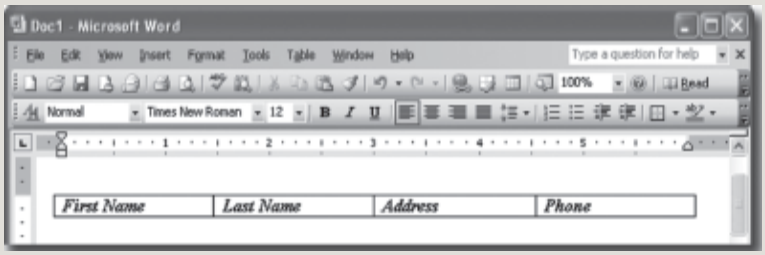


Figure 13.37 Formatting the Table

Aligning Text: Whenever you create a new table and enter a text, by default, it is left aligned. However, you can change the text alignment if required. The text in the table can be justified and aligned to the centre or right side. To change the alignment of the text, follow the steps given below:

1. Select a cell or a series of cells or an entire row or column.
2. Click the appropriate alignment button on the Formatting toolbar to align the selected text (Figure 13.38).

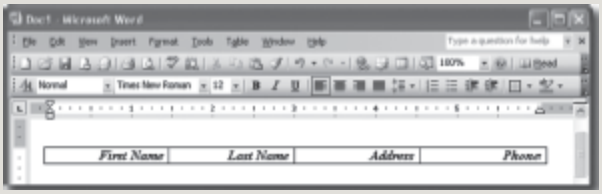


Figure 13.38 Aligning Text in Table

Positioning Text: When you enter a text in tables, by default it is positioned horizontally in the table cells. However, Microsoft Word allows you to change the text orientation in table cells so that it can be positioned in a

vertical direction. To perform this, follow the steps given below:

1. Place the cursor in the cell where the text is to be typed.
2. Select **Text Direction** from the **Format** menu to display the Text Direction dialog box.
3. Select the desired text style and click **OK.** Now, if you type a text in the cell, it appears in the same orientation as you selected (Figure 13.39).

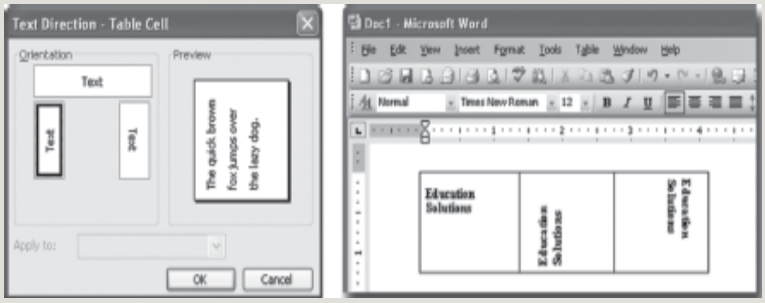


Figure 13.39 Changing the Orientation of the Text in the Table

Using Table AutoFormat: To make a table more presentable, Microsoft Word provides *Table AutoFormat* feature that allows one to apply different formats, such as borders, shading, special fonts and colour, etc., to a table in a quick and easy manner. To apply an AutoFormat to a table, follow the steps given below:

1. Place the cursor anywhere in the table.
2. Select **Table AutoFormat** from the **Table** menu or click **Table AutoFormat** ( ) button on the Tables and Borders toolbar to display the Table AutoFormat dialog box (Figure 13.40).

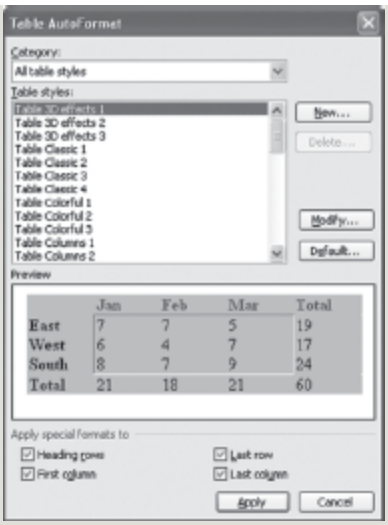


Figure 13.40 Table AutoFormat Dialog Box

3. Click on the desired format from the **Table styles** section. Note that you can preview the selected format in the Preview box.
4. To apply special formats, click the checkboxes under the **Apply special formats to** section. The options included in this section indicate where AutoFormat will place distinctive formatting for the headings or for the entire text in the table.
5. Click **Apply** to apply the formatting and close the Table AutoFormat dialog box (Figure 13.41).

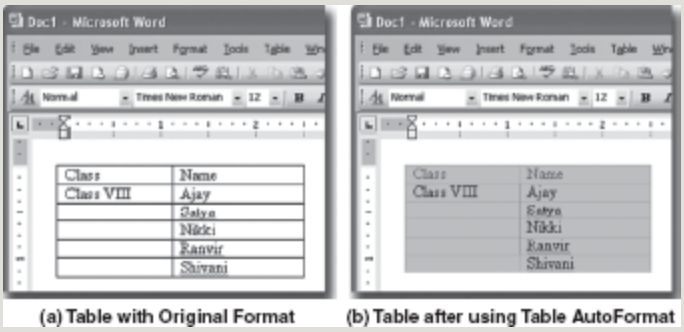


Figure 13.41 Using Table AutoFormat

13.5.4 Adding and Deleting Rows and Columns

New rows can be added to an existing table. The simplest way to do this is to move the cursor to the last column of the last row and press the **Tab** key. You can also insert rows by using the **Table** menu. On the **Table** menu, point to **Insert** and then choose **Rows Above** or Rows Below. Similarly, columns can be added by selecting the **Columns to the Left** or **Columns to the Right.**

For deleting rows and columns from the table, click inside the row or column that you wish to delete or select several rows or columns to delete. On the **Table** menu, point to **Delete,** and then click **Rows** or **Columns.**

13.5.5 Converting Text to Table

You can convert any text, whether it is a tabbed list or a series of paragraphs, into a table using the *Convert Text to Table* feature of the Word. To perform this, follow the steps given below:

1. Select the text to be converted into table.
2. Select **Convert** from the **Table** menu and choose **Text to Table** option to display the Convert Text to Table dialog box.
3. Enter the required number of columns and rows in the **Table size** section.
4. Click the desired options in the **AutoFit behavior** section to specify the columns' sizing.
5. Select the desired delimiter or separator in the **Separate text at** section in order to specify what text to be added in an individual cell (Figure 13.42).

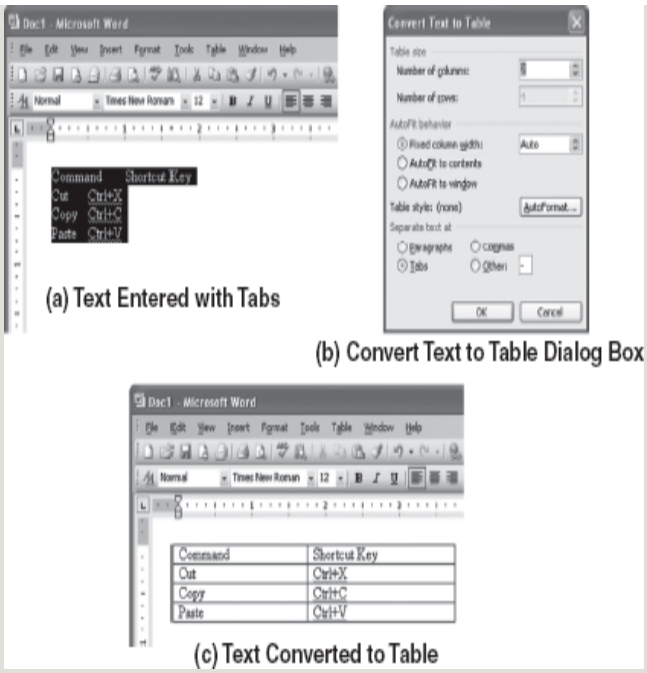


Figure 13.42 Converting Text to Table

13.5.6 Converting Table to Text

You can also convert an existing table to text using the *Convert Table to Text* feature of the Word. To convert a table to text, follow the steps given below:

1. Select the text in table.
2. Select **Convert** from the **Table** menu and choose **Table to Text** option to display the Convert Table to Text dialog box.
3. Select the desired delimiter or the separator in the **Separate text with** section that will separate the different columns (Figure 13.43).

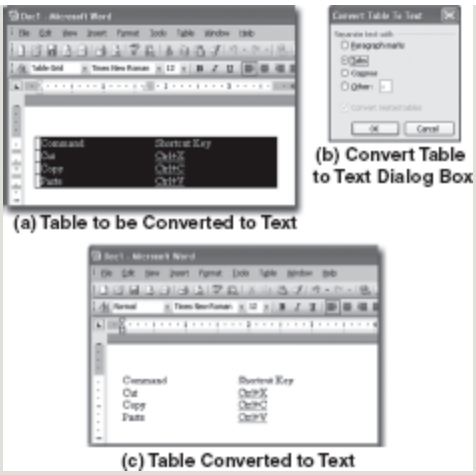


Figure 13.43 Converting Table to Text

13.6 CHECKING SPELLING AND GRAMMAR

Microsoft Word allows you to check spelling and grammar in the desired text in a document. Before printing or reading a document, you should always check for typing errors. Spelling and grammar errors are highlighted in the document with a red and green underlines, respectively, and may provide corrective suggestions. It can thus help you become better spellers, even broaden your vocabulary if you are curious enough to check out the meanings of the many words you have never come across before.

13.6.1 Using the Auto Spell and Auto Grammar Features

Microsoft Word automatically checks the spelling and grammar as you type the text and provides visual clues about the spelling and grammar unless this feature is turned off. Misspelled words are underlined in red, while sentences with grammatical flaws are underlined in green. Note that these lines are for viewing purposes only and will not be printed.

Although spelling and grammatical errors can be corrected manually, Word provides easy-to-use assistance when you right-click on the error, as shown in Figure 13.44. When you right-click on a spelling or grammatical error, Word displays a popup menu with suggestive corrections. Select the desired correction and Word will automatically make the correction.

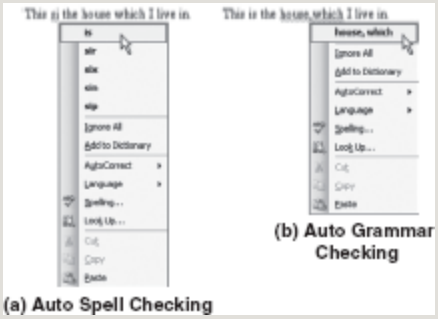


Figure 13.44 Auto Spell and Grammar Checking

13.6.2 Using Spelling and Grammar Dialog Box

To start the spell check process, perform *any* of the following steps:

- Click the **Spelling and Grammar** button ( ) on the Standard toolbar.
- Select **Spelling and Grammar** from the **Tools** menu.

When you perform any of the above actions, Word 2003 will stop at words that it does not recognize, as well as repeated words, such as, “this this is a mistake”, in the Spelling and Grammar dialog box. You can choose from the choices Microsoft Word offers, type in your own change, or ask to ignore that instance or all instances of the word in the document (see Figure 13.45).

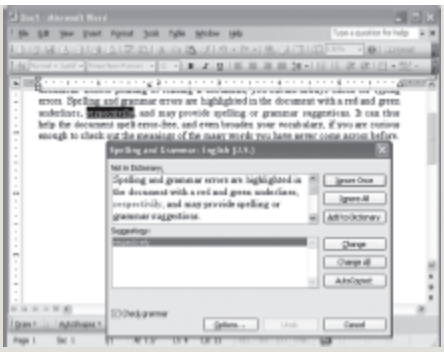


Figure 13.45 Spelling and Grammar Dialog Box

The top section of the dialog box contains the first unrecognized word—shown in red with its accompanying sentence so you can see how you have used the word within the document. The bottom section contains suggestions for correcting the word. If the word should remain as it is, select the **Ignore Once** button. Word also offers the option of **Ignore All** if the word in question appears throughout the document. If one of the suggestions is correct, double-click on the correct spelling or highlight the word and choose the **Change** button. If you are afraid you misspelled a word more than once, click on the **Change All** button. If the word should be added to the custom dictionary, click the **Add to Dictionary** button. You can use the **AutoCorrect** button to add the word to the *AutoCorrect* list. In the future, when you misspell this word while typing, Word will automatically correct it— without you having to access the spell checker. When you click the just mentioned buttons, Word will continue the spelling and grammar checking until it finds the next mistake or when there are no mistakes left in the document or the selected text.

13.6.3 Using AutoCorrect Options

The *AutoCorrect Options* feature automatically detects common errors like misspelled words, grammatical errors, and incorrect capitalization and replaces them with the correct text. For example, “grammer” can be rectified as “grammar.” You can also use AutoCorrect to quickly insert text, graphics, or symbols. For example, type (c) to insert ©, or type *asap* to insert *as soon as possible*. There are several words available in the AutoCorrect dictionary; however, you can also enter a word that is commonly misspelled in the dictionary by selecting **AutoCorrect Options** from the **Tools** menu. This will display the AutoCorrect dialog box. In this dialog box, type in the misspelling in the **Replace** textbox and the corrected spelling in the **With** textbox. After that click the Add button and then close the dialog box by clicking the **OK** button (Figure 13.46).

Figure 13.46 AutoCorrect Dialog Box

13.6.4 Using Synonyms and Thesaurus

Sometimes you may wish to replace a word with its synonym. Microsoft Word has the *Thesaurus* feature, which does just the same. To use this feature, follow the steps given below:

1. Place the cursor on the word that is to be looked up.
2. On the **Tools** menu, select **Language** and then select **Thesaurus** to display the **Research** task pane.
3. A listing of similar words is provided under the **Thesaurus: English (U.S.)** section. To use a word from the list, point to it, click the down arrow, and then click **Insert** or **Copy**. You can also view the similar words of any word in the list by clicking that word (Figure 13.47).

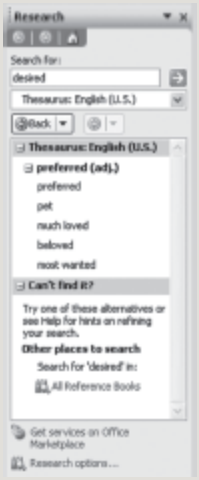


Figure 13.47 Research Task Pane

You can also quickly look up a word directly from within your document. To do this, select the word in contention, right-click it, and then choose **Synonyms** from the pop-up menu. From the list of suggested words, choose any one you like or select the **Thesaurus** option to display the Research task pane (Figure 13.48).

Figure 13.48 Finding a Synonym

13.7 ADDING GRAPHICS TO DOCUMENT

Graphics are frequently used with text to add emphasis and visual impact in Word documents. Microsoft Word 2003 uses two basic types of graphics: *pictures* and *drawing objects*.

13.7.1 Pictures

Pictures are used in the document to enhance its readability and understandability. These include bitmaps, scanned pictures, photographs, and clip art. You can use any image editing software to create photographs, line drawings, and other graphics, as long as you can save the file in a format that Word 2003 recognizes. Some common formats are *.BMP, .GIF, .JPG,* and *.WMF*.

Inserting Picture from a File: Word allows you to insert pictures or scanned photographs from other programs and locations into the Word document. To add photographs or graphics to an existing file, follow the steps given below:

1. Click on the desired location, where the picture is to be inserted.
2. On the **Insert** menu, click **Picture** and select **From File**. This opens the Insert Picture dialog box as shown in Figure 13.49.

Figure 13.49 Insert Picture Dialog Box

3. Use the **Look in** drop-down box to find the image to be inserted.
4. Select the file name from the list and click the Insert button.

Clip Art: Microsoft Word comes with its own set of pictures in the Clip Gallery. The *Clip Gallery* includes a wide variety of clip art that makes it easy for you to enhance your documents with professionally designed images. Most clip art images are in metafile format, which means you must ungroup an image and convert it to a drawing object in order to enhance it. To add a Clip Art from the Clip Gallery, follow the steps given below:

1. Click on the desired location, where the clip image is to be inserted.
2. On the **Insert** menu, click **Picture** and select **Clip Art** to display the Clip Art task pane. You can also open Clip Art Gallery by clicking the **Insert Clip Art** button ( ) on the Drawing toolbar (Figure 13.50).

Figure 13.50 Insert Clip Art Task Pane

3. To find an image, click in the box following **Search for** and enter keywords describing the image you want to use and Click on **Go** button.
4. Move the mouse on the image to be inserted and click drop-down arrow alongside the image. This opens the pop-up menu, which is shown in Figure 13.51.

Figure 13.51 Inserting Clip Art

5. Click **Insert** from the menu to insert the image. You can also drag a picture or other clip from the Clip Gallery to your document.
6. To add multiple images, continue selecting images to be added to the document. When you are finished using the Clip Gallery,

click the **Close** button on the Clip Gallery title bar to close Clip Art task pane.

13.7.2 Editing a Picture

Once an image is inserted, it can be changed and enhanced by using the Picture toolbar. To edit an image, follow the steps given below:

- Select the graphic by clicking on it so the selection handles appear on the sides and corners of the picture. Click and drag these handles to resize the image. To resize the object, place the mouse pointer over one of the selection handles until it becomes a double-headed arrow ( ). The picture can be resized in three ways: *horizontally, vertically,* or in *proportion*.

    1. To stretch horizontally, drag either of the side handles.
    2. To stretch vertically, drag the handle to either the top or the bottom.
    3. To resize proportionally, drag one of the corner handles.

Image effects can also be changed by using the Picture toolbar. When you select a picture, the Picture toolbar appears with tools that you can use to crop the picture, add a border to it, or adjust its brightness and contrast. If the Picture toolbar does not appear, right-click the picture and select **Show Picture Toolbar** from the shortcut menu (Figure 13.52).

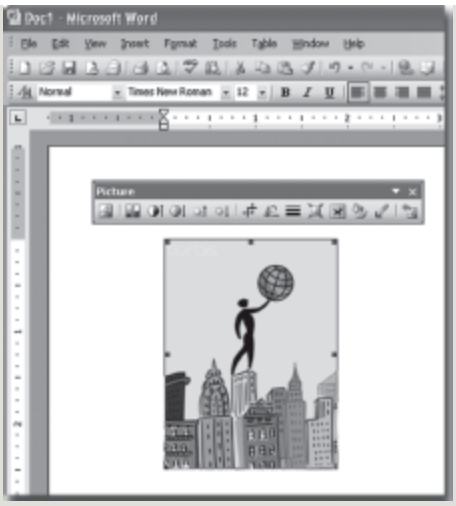


Figure 13.52 Selecting Picture for Editing

Table 13.7 lists all the available buttons on the Picture toolbar.

Table 13.7 Picture Toolbar Options

Command	Button	Description
Insert Picture		Displays the image selection window and allows the user to change the image
Color		Allows making the image in greyscale, black and white, or a watermark
More/less Contrast		Modifies the contrast between the colours of the image
More/less Brightness		Darkens or brightens the image
Rotate Left 90°		Rotates the picture 90° to the left
Crop		Deletes outer portions of the image
Line Style		Adds a variety of borders to the graphic
Compress Pictures		Reduces the picture's file size
Text Wrapping		Modifies the way the document text wraps around the graphic
Format Picture		Displays all the image properties in a separate window
Set Transparent Color		Sets transparent colour on the object
Reset Picture		Deletes all the modifications made to the image

13.7.3 Drawing Objects

Word allows you to add objects such as lines, curves, and WordArt. The drawing objects include:

- **AutoShapes**: It comprises lines, curves, basic shapes, and many such objects.

- **WordArt:** It comprises text written using special effects.

AutoShapes: Microsoft Word comes with a set of ready-made AutoShapes containing several categories of shapes, including lines, basic shapes, flowchart elements, stars and banners, and callouts that you can use in your documents. You can resize, rotate, flip, colour, and combine the shapes with other shapes such as circles and squares to make shapes that are more complex. To use AutoShapes, follow the steps given below:

1. On the **Insert** menu, click **Picture** and then select **AutoShapes.**

   Or

   On the **View** menu, click **Toolbars** and then select **Drawing** to display the Drawing toolbar. This toolbar contains the **AutoShapes** menu (Figure 13.53).

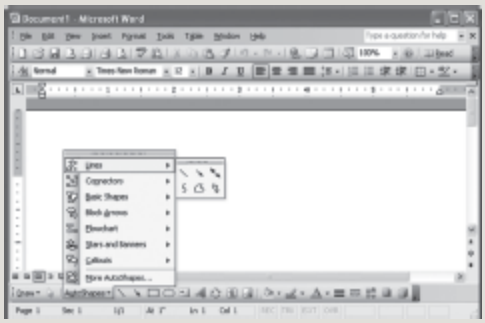


Figure 13.53 AutoShapes Menu

Table 13.8 lists all the available buttons on the AutoShapes menu.

Table 13.8 AutoShapes Menu Options

Command	Button	Description
Lines		Draws a straight line, arrow, or double-headed arrow
Connectors		Connects the shapes and keeps them connected
Basic Shapes		Draws many two- and three-dimensional shapes, braces, and brackets
Block Arrows		Draws block arrows from many types of two- and three-dimensional arrows
Flowchart		Draws flowchart
Stars and Banners		Draws stars, bursts, banners, and scrolls
Callouts		Draws callouts which are used to label some text on the document
More AutoShapes		Used to choose more AutoShapes that are available

2. Select the required shape on the toolbar. Use the drag-and-drop method to draw the shape in a document (Figure 13.54).

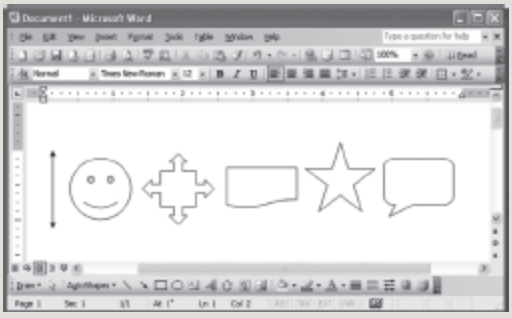


Figure 13.54 Inserting AutoShapes

3. When the shape has been made, it can be resized using the sizing handle, and other adjustments specific to each shape can be modified using the yellow diamond handles. Each AutoShape can be rotated by clicking and dragging the green handles around the image to rotate it (Figure 13.55).

Figure 13.55 Resizing and Rotating AutoShapes

WordArt: Using *WordArt Gallery*, you can create many text effects. You can create shadowed, skewed, rotated, and stretched text, as well as text that can be fitted to pre-defined shapes. To insert a WordArt drawing object, follow the steps given below:

1. Click **Insert WordArt** button ( ) on the Drawing toolbar. Alternatively, select **Picture** from the **Insert** menu and then select **WordArt**. When you perform any of these actions, the WordArt Gallery dialog box is displayed (Figure 13.56).

Figure 13.56 WordArt Gallery Dialog Box

2. Select a WordArt style to apply to the document and click **OK** to close the dialog box. This opens Edit WordArt Text dialog box, where the text is entered for the chosen style (Figure 13.57).

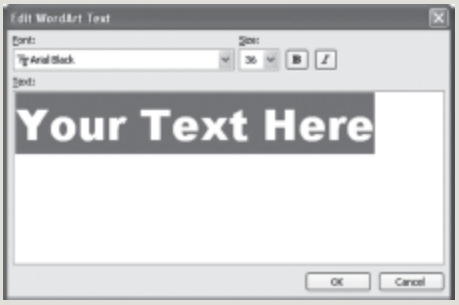


Figure 13.57 Edit WordArt Text Dialog Box

3. Select a font by clicking the down arrow next to the **Font** list. The font size can also be selected by clicking the down arrow next to the **Size** box. To make the text bold, click the ( ) button, and to make the text italics, click the ( ) button.
4. When text has been entered and font selection made, click **OK**. The WordArt now appears in the document. With the selection handles visible on the WordArt, you can resize, move, and rotate the WordArt (Figure 13.58).

Figure 13.58 WordArt in Document

To add or change effects to the text, use the buttons on the WordArt and Drawing toolbars. The WordArt toolbar appears when you click the WordArt special text. Table 13.9 lists all the available buttons on the WordArt toolbar.

Table 13.9 WordArt Toolbar Options

Command	Button	Description
Insert WordArt		Opens the WordArt Gallery dialog box, allowing you to create a new WordArt object
Edit Text		Opens the Edit WordArt Text dialog box, allowing you to edit and format WordArt text
WordArt Gallery		Opens the WordArt Gallery dialog box, allowing you to select a different WordArt style
Format Object		Opens the Format WordArt dialog box, allowing you to format a WordArt object so that you can add colours and lines (borders) and adjust the size and layout
WordArt Shape		Displays a palette of WordArt shapes, allowing you to select a different object shape such as changing the slopes, curves, and arches
Text Wrapping		Displays a palette of wrapping styles that affect how you can move the object on a page
WordArt Same Letter Heights		Changes all the letters in WordArt text to the same height within the defined shape
WordArt Vertical Text		Changes the appearance of the text by making it either vertical (up/down) or horizontal (left/right)
WordArt Alignment		Displays a menu from which you can select a desired alignment option
WordArt Character Spacing		Displays a menu from which you can select the desired character spacing between letters such as loose, tight, and normal

Drawing Toolbar: The *Drawing* toolbar is used to change and enhance the drawing objects with colours, patterns, borders, and other effects. The drawing toolbar includes the following options Table 13.10:

Table 13.10 Drawing Toolbar Options

Command	Button	Description
Draw		Gives you options to group images, move items forward or back, move or rotate images, and set AutoShape defaults
Select Objects		Used to select objects in the document. To select a single object, click the object with the arrow. To select one or more objects, drag the arrow over the objects you want to select
AutoShapes		Displays the in-built shapes
Line		Draws a straight line where you click or drag in the document
Arrow		Draws an arrow-headed line where you click or drag in the document
Rectangle		Draws a rectangle or a square where you click or drag in the document
Oval		Draws an oval or a circle where you click or drag in the document
Text Box		Draws a text box where you click or drag in the document
Insert WordArt		Activates the WordArt Gallery using which you can insert WordArt objects
Insert Diagram or Organization Chart		Allows you to insert a diagram or organization chart from the Diagram Gallery
Insert Clip Art		Opens the Clip Art task pane using which you can insert the desired clip art image
Insert Picture		Lets you select and insert the picture at the specified location
Fill Color		Adds, modifies, or removes the fill colour or fill effect from the selected object
Line Color		Adds, modifies, or removes the line colour from the selected object

Command	Button	Description
Font Color		Adds, modifies, or removes the text colour from the selected text
Line Style		Lets you choose the thickness and style of the line
Dash Style		Lets you make various dashed line styles
Arrow Style		Determines the look and type of arrows
Shadow Style		Allows you to give images a shadow and format that shadow
3-D Style		Allows you to give images a three-dimensional (3-D) look and format it

13.8 MAIL MERGE

Mail merge feature of Word allows you to merge generic information which is to be repeated in each letter, mailing label, or, envelope, etc., with varying information like, a list of names and addresses of employees. This allows having almost identical documents in which the content of the documents is the same for each employee; however, each of it can be customized for a particular individual in the list. For example, if one wants to send a form letter to five or five hundred different people, Word can be used to personalize each one and create envelopes or mailing labels.

13.8.1 Components of a Mail Merge

The components of the mail merge are as follows:

1. The *main document* comprises the generic information in the form of text and graphics. It also determines the format of the resultant merged document. The five types of main documents are:
    1. **Letters:** To create letters that can be sent to a group of people.
    2. **E-mail messages:** To create e-mail messages that can be sent to a group of people.

3. **Envelopes:** To create addressed envelopes for a group mailing.
4. **Labels:** To create address for a group mailing.
5. **Directory:** To create a directory containing a printed list of addresses or a catalogue.

2. A *data source* comprises varying information in which records representing individual items are stored. Each record consists of fields, for example, a record of an employee may consist of name and address, where a name may consist of fields, FirstName and LastName and the address may contain fields HouseNo, City, and State. A data source contains multiple records, representing information for numerous employees.
3. The *merged document* comprises the combination of generic text and variable information. In the main document, you can insert merge fields indicating the place where you want to place the variable information. When the information in the data source is merged with the main document, the merge fields are replaced with the variable information representing the data for each individual.

13.8.2 Mail Merge Steps

To perform mail merge, follow the steps given below:

1. Click **New** ( ) on the Standard toolbar to create a document and type the desired text. For example, enter the text in the new file as shown in Figure 13.59 and save it. An already existing file can also be used as the main document. If you already have a document containing the desired text, open it.

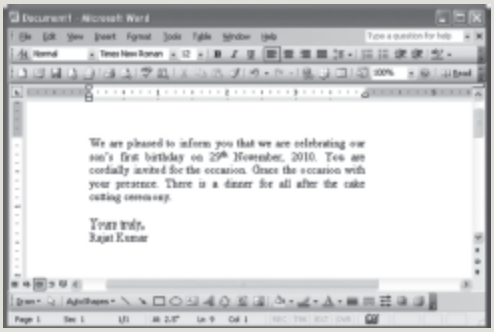


Figure 13.59 Sample Text

2. Click the **Tools** menu, point to **Letters and Mailings**, and then click **Mail Merge** to display the Mail Merge task pane (Figure 13.60).

Figure 13.60 Step 1: Select Document Type

3. Select the appropriate option under the **Select document type** section. For example, select the option **Letters** to create a letter.
4. Click **Next: Starting document** button to perform the next step of mail merge (Figure 13.61).

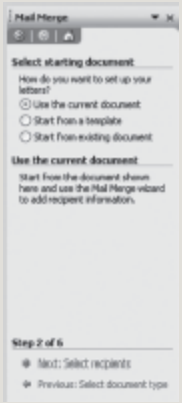


Figure 13.61 Step 2: Select Starting Document

5. Click **Use the current document** option under the **Select starting document** section to use the currently opened document as the main document.
6. Click **Next: Select recipients** to perform the next step of the mail merge. You can click **Previous: Select document type** to return to previous step and change the selection options (Figure 13.62).

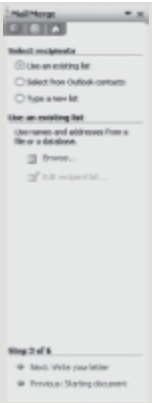


Figure 13.62 Step 3: Select Recipients

7. Select **Type a new list** under the **Select recipients** section to create the address list (Figure 13.63).

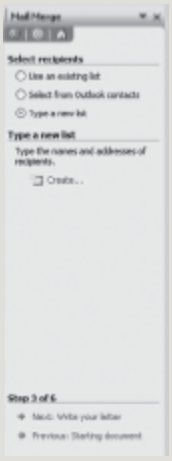


Figure 13.63 Step 3: Type a New List Option Selected

8. Click **Create** button to display the New Address List dialog box (Figure 13.64).

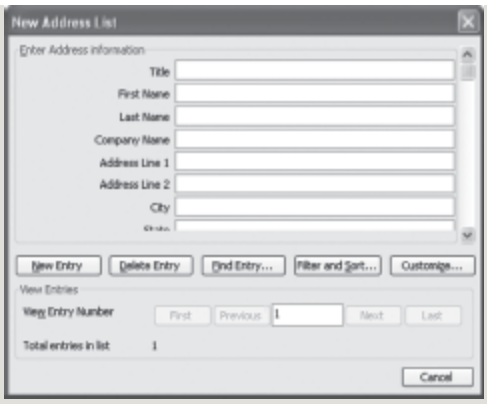


Figure 13.64 New Address List Dialog Box

9. In the New Address List dialog box, type the appropriate information you want to include for the first entry in the boxes under **Enter Address information** section. Note that as you start entering the information, the **Cancel** button changes to **Close** button.
10. Click **New Entry** button to complete the first entry and enter a new entry.
11. Repeat steps 9 and 10 until you have added all the address entries and click the **Close** button to display the Save Address List dialog box (Figure 13.65).

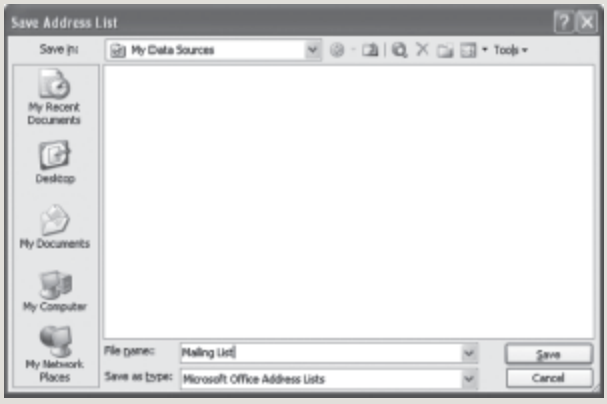


Figure 13.65 Save Address List Dialog Box

12. Specify the name of the file, the location of the file, and the file type.
13. Click **Save** button to display the Mail Merge Recipients dialog box (Figure 13.66).

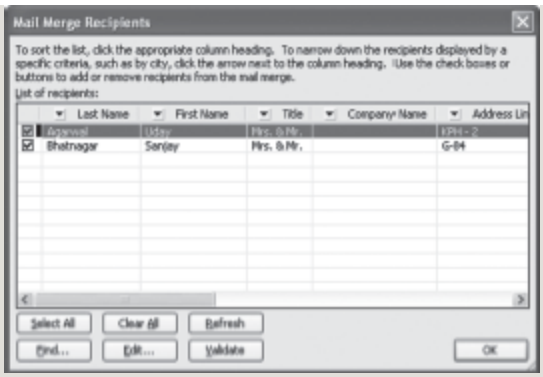


Figure 13.66 Mail Merge Recipients Dialog Box

14. In the Mail Merge Recipients dialog box, click the heading of the column you want to use to sort the list. You can also add or remove the recipients from the mail merge list by clicking the checkbox beside the name of the recipient. Click **OK** button to close the Mail Merge Recipients dialog box and return to the Mail Merge task pane.
15. Click **Next: Write your letter** to perform next step of the mail merge and to display the mail merge toolbar (Figure 13.67). Some of the tools are listed in Table 13.11.

Table 13.11 Mail Merge Toolbar Options

Command	Button	Description
Main Document Setup		Creates a main document
Open Data Source		Opens the existing list from which recipients are selected
Mail Merge Recipients		Allows to add or remove the recipients from the mail merge list
Insert Address Block		Inserts a merge field for the address in the main document at the specified location. This merge field will be replaced with an address from the mailing list
Insert Greeting Line		Inserts a merge field for the greeting line in the main document at the specified location
Insert Merge Fields		Inserts a merged field at the insertion point
Insert Word Field		Inserts a word field at the insertion point
View Merged Data		Switches between merge fields and the data that appear in them
Highlight Merge Fields		Highlights the merge fields
First Record		Shows the first record in the data store
Previous Record		Shows the previous record in the data store
Go to Record		Moves to the record specified in the box
Next Record		Shows the next record in the data store
Last Record		Shows the last record in the data store

Command	Button	Description
Find Entry		Finds a specified entry
Check for Errors		Checks and reports errors
Merge to New Document		Merges the main document and data
Merge to Printer		Merges the document and prints it

Figure 13.67 Step 4: of Mail Merge

16. Place the cursor at the location in the main document where you want to insert the merge fields, that is, where you want to merge name, addresses, etc. from the data source.
17. Click the **Address block** under the **Write your letter** section to display the Insert Address Block dialog box (Figure 13.68).

W

Figure 13.68 Insert Address Block

18. Select the desired format and click **OK** button. A merge field for the address appears in the letter. This merge field will be replaced with an address from the address list (Figure 13.69).

Figure 13.69 Merge Field for Address

19. Similarly, you can add **Greeting Line** if you want to add a greeting in the letter.
20. Click **Next: Preview your letters** to perform the next step of the mail merge (Figure 13.70).

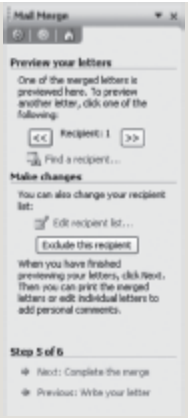


Figure 13.70 Step 5: of Mail Merge

21. Click **Next: Complete the merge** to finish the mail merge (Figure 13.71).

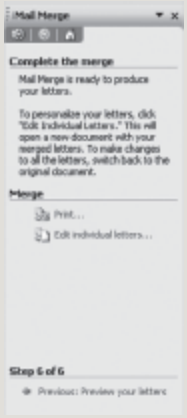


Figure 13.71 Step 6: of Mail Merge

22. Click **Edit individual letters** under the **Merge** section to merge the main document with the information from the mailing list. The Merge to New Document dialog box appears (Figure 13.72).

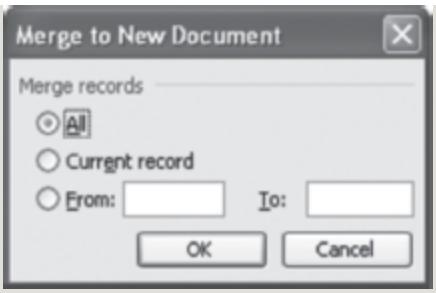


Figure 13.72 Merge to New Document Dialog Box

23. Click the appropriate option to specify the records you want to merge.
24. Click **OK** to create letters (Figure 13.73).

Figure 13.73 Merged Document

13.9 PRINTING A DOCUMENT

A major part of any word-processing software is to print a document. Microsoft Word allows you to choose what part of the document to print (such as the current page, multiple pages, or the entire file). In addition, you can specify which printer to use and how many copies to print.

13.9.1 Setting Page Margins

Page margins determine the distance between the text and the edge of the paper. To specify the margins in the document, choose **Page Setup** from the **File** menu to display the Page Setup dialog box. As shown in Figure 13.74, Microsoft Word 2003's default margins are 1.25 inches on the left and right and 1 inch on the top and bottom. If required, change the margin from the top, bottom, left, and right. To change the orientation (whether to print the document in landscape or portrait mode), select the option from the **Orientation** section. If you plan to print on both sides of the page, select the **Mirror margins** option from the **Multiple pages** drop-down list under **Pages** sections so that the margins on facing pages mirror each other. The **Preview** box at the right of the Page Setup dialog box displays thumbnail-sized representation of a sample page in your document. Be sure that your margins are appropriate for your document.

Figure 13.74 Page Setup Dialog Box

The Page Setup dialog box also allows you to format the size and orientation of the paper. To change the size of the paper, click the **Paper** tab and select the desired size (letter, A4, etc.) from the **Paper_size** list.

13.9.2 Print Preview

Print preview is a way to view the appearance of the printed document on-screen before printing the final output. Previewing the document is a great way to catch formatting errors, such as incorrect margins, overlapped text, boldfaced text, and other text enhancements. This helps in saving costly printer paper, ink, and time. To view the document in print preview mode, choose **Print Preview** from **File** menu or click the **Print Preview** button ( ) on the Standard toolbar. In the print preview mode, the document appears as a reduced image in the Print Preview screen. The Zoom feature can be used to magnify the image for readability. To magnify the text for a larger view, click the **Magnifier** button ( ) and place the cursor over the text you wish to enlarge. When the cursor resembles a magnifying glass encircling a + sign ( ), click the text and the text will be zoomed in. You can magnify any portion of the page. To return to the original view size, position the cursor over the text again and when the cursor resembles a magnifying glass encircling a – sign ( ), click to reduce the text to a smaller magnification. Click **Close** to return to the document (Figure 13.75).

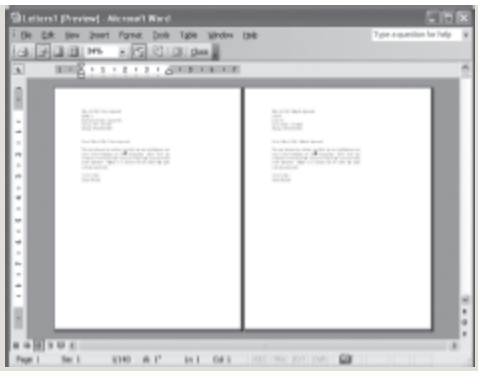


Figure 13.75 Print Preview

13.9.3 Printing the Document

Once you are ready with all the formatting and editing, you can print the document. To print a document, click the Print button ( ) on the Standard toolbar, or select **Print** from the **File** menu to display the Print dialog box (see Figure 13.76). This dialog box allows you to select the number of copies and how many pages of the document you want to print. You can also specify the range of the document to print by choosing:

Figure 13.76 Print Dialog Box

- **All:** To print all the pages of the document.

- **Current page:** To print the current page in the document.
- **Pages:** To print a certain number of pages.
- **Selection:** To print the selected portion of the document.

The **Number of copies** list allows you to enter the number of copies you want to print. If you want to print multiple copies, select the **Collate** checkbox to print a complete copy of the document before the first page of the next copy is printed. Now, the copies will come out in the order page 1, 2, 3; page 1, 2, 3, etc. The **Pages per sheet** list offers you to select the number of pages in the document that you want to print on each sheet of paper. Once all the printer settings have been set, click *OK* and Word will begin printing the document (Figure 13.76).

Important Keyboard Shortcuts

Document Actions

Action	Shortcut Key
Open a Document	CTRL+O
Create a New Document	CTRL+N
Close a Document	CTRL+W
Save As	F12
Save	CTRL+S
Help	F1
Print Preview	CTRL+F2
Print	CTRL+P
Close Word Application	ALT+F4

Text Formatting Actions

Action	Shortcut Key
Font Face	CTRL+SHIFT+F
Font Size	CTRL+SHIFT+P
Style	CTRL+SHIFT+S
Bold	CTRL+B
Italic	CTRL+I
Underline	CTRL+U
Change Case	SHIFT+F3
Font Dialog Box	CTRL+D
Left Alignment	CTRL+L
Center Alignment	CTRL+E
Right Alignment	CTRL+R
Justified Alignment	CTRL+J
Indent	CTRL+M
Apply Bulleted List	CTRL+SHIFT+L
Undo	CTRL+Z
Redo	CTRL+Y
Cut	CTRL+X
Copy	CTRL+C
Paste	CTRL+V

Text Formatting Actions

Action	Shortcut Key
Delete Previous Word	CTRL+Backspace
Find	CTRL+F
Replace	CTRL+H
Go To	CTRL+G

Table Actions

Action	Shortcut Key
Next Cell	Tab
Previous Cell	SHIFT+Tab
Beginning of Column	ALT+PageUp
End of Column	ALT+PageDown
Select Table	ALT+ (Num 5) (keeping NumLock OFF)

Miscellaneous Actions

Action	Shortcut Key
Thesaurus	SHIFT+F7
Spellchecker	F7

LET US SUMMARIZE

1. Word processing is the manipulation of characters, words, text, numbers, sentences, and paragraphs in the document so that they are error-free and looks attractive.
2. To open Microsoft Word, click start, point to All Programs, and then select Microsoft Word.
3. To get help, click the Microsoft Word Help button ( ) on the Standard toolbar. Alternatively, select Microsoft Office Word Help from the Help menu.
4. To create a new document, click on New Blank Document button ( ) on the Standard toolbar. Alternatively, select New from the File menu.
5. To open an existing document, click on the Open button ( ) on the Standard toolbar. Alternatively, select Open from the File menu.
6. To save the document, Microsoft Word provides two menu options, namely, Save and Save As. By default, Microsoft Word document files are saved with .DOC extension, although it can also be saved in other file formats like .TXT, .RTF, and .HTML.
7. To close the document, select Close from the File menu.
8. Formatting is the process of determining how the typed text will appear on a printed page. It is applied by altering the appearance of text by setting the font type and font size, line spacing, and colour. To format the text, first select the text to be formatted and then use various buttons on the Formatting toolbar. Text can also be formatted by using the Font dialog box, which can be displayed by selecting Font from the Format menu.
9. To highlight the text use the Highlight button ( ) from the Formatting toolbar. In case you want to change the colour of the

text, use the Font Color button ( ) from the Formatting toolbar.

10. To align or indent a paragraph, use the appropriate buttons on the Formatting toolbar. A paragraph can also be formatted using the Paragraph dialog box, which can be displayed by selecting Paragraph from the Format menu.
11. A style is a set of formatting specifications that can automatically be applied to the text. To apply a style, place the cursor in the paragraph where the style is to be applied. Then click the Style drop-down box on the Formatting toolbar and select a style by clicking on it.
12. The Format Painter is a time-saving feature for quickly copying the format of a block of text. This tool can be used by clicking the Format Painter button ( ) on the Standard toolbar.
13. To move the text, choose Cut from the Edit menu. To copy the text, choose Copy from the Edit menu. To paste the cut or copied text or image, choose Paste from the Edit menu.
14. To find and replace text within the document use the Find and Replace dialog box, which can be displayed by selecting Find from Edit menu.
15. To undo the last action, select Undo from the Edit menu. To redo the last undo action, select Redo from the Edit menu. You can also use the Undo button ( ) and Redo button ( ) from the Standard toolbar.
16. To use bullets and numbering, first select the text you want to format. Next, use the Bullets ( ) or Numbering ( ) button on the Formatting toolbar.
17. Headers are text that appears at the top of every page, and footers contain text that appears at the bottom of each page. To create a header and footer in a document, select Header and Footer from the View menu, or double-click on the header area of the document.
18. Tables can be created by using the Insert Table dialog box, which can be displayed by selecting Insert and then Table from the Table menu. Another way of creating a table is to click the Insert Table button ( ) on the Standard toolbar and then dragging the mouse along the grid to highlight the number of rows and columns for the table.
19. Microsoft Word automatically checks the spelling and grammar as you type the text and provides visual clues about the spelling and grammar. Misspelled words are underlined in red, while sentences with grammatical flaws are underlined in green.
20. Microsoft Word has the Thesaurus feature to replace a word with its synonym or antonym. This feature can be used by selecting Language and then Thesaurus from the Tools menu.
21. To add a picture from an existing file, use the Insert Picture dialog box, which can be displayed by selecting Picture and then From File from the Insert menu.
22. Microsoft Word comes with its own set of pictures in the Clip Gallery. This gallery includes a wide variety of clip art that makes it easy for you to enhance your documents with professionally designed images. To add a Clip Art from the Clip Gallery use the Insert ClipArt dialog box, which can be displayed by selecting Picture and then Clip Art from the Insert menu.
23. To use AutoShapes, select Picture and then AutoShapes from the Insert menu. Alternatively, use the AutoShapes menu from the Drawing toolbar.

24. To use WordArt, select Picture and then WordArt from the Insert menu. Alternatively, use the WordArt button ( ) from the Drawing toolbar.
25. Page margins determine the distance between the text and the edge of the paper. To specify the margins in the document use the Page Setup dialog box, which can be displayed by selecting Page Setup from the File menu.
26. Print preview is a way to review the appearance of the printed document on-screen before printing the final output. To view the document in print preview mode, choose Print Preview from File menu or click the Print Preview button ( ) on the Standard toolbar.
27. To print a document, click the **Print** button ( ) on the Standard toolbar. Alternatively, use the Print dialog box, which can be displayed by selecting Print from the File menu.

EXERCISES

Fill in the Blanks

1. .....button from the Standard toolbar is used to see how a document looks before it is printed.
2. A red line under a word in a document means that there is a ..... error.
3. To select a complete document, use the ..... keyboard shortcut.
4. To cancel the last text modified, ... button is used.
5. ..... bar displays the name of the document.
6. ..... and ..... are two basic toolbars in Word.
7. A ..... is the message that appears on the screen to provide information about toolbar button.
8. To save the document, Microsoft Word provides two options, namely, ... and .....
9. ..... and ..... option in Microsoft Word is used to locate a particular word, phrase, or set of characters in a document and replace it with something else.
10. The Print button appears on ..... toolbar.
11. Every Word document contains .....and ..... areas where you can put a document title, author name, page number, and graphics.
12. ..... provide means of presenting information in an organized and attractive manner that displays text in grid of rows and columns of cells.
13. ..... contains command buttons that provides single-click entry to carry out many text-processing tasks.
14. ..... is a time-saving feature for quickly copying the format of a block of text.
15. ..... toolbar allows you to add border styles, shading, text effects, alignment, and more option to a table.

Multiple-choice Questions

1. The toolbar that contains buttons for formatting a text and paragraph is .....

1. Drawing
2. Formatting
3. Standard
4. Drawing

2. To move data from one part of the document to another, following combination is used:.....

    1. Cut and Paste
    2. Cut and Delete
    3. Copy and Undo
    4. Cut and Insert

3. To save a document for the first time ..... option is used.

    1. Save As
    2. Cut
    3. Save
    4. Paste

4. To see the document before the printout is taken, the following option should be used:.....

    1. Insert Table
    2. Print Preview
    3. Paste
    4. Format Painter

5. Two basic types of graphics used in Word 2003 documents are .....

    1. Drawing Objects and Pictures
    2. AutoShapes and Clip Art
    3. Header and Footer
    4. Spelling and Grammar

6. Tables can be created in a number of ways, one of the most commonly used method is to create table by clicking on the ..... menu.

    1. Format
    2. Window
    3. Help
    4. Table

7. A user can use the ..... and ..... commands to search for and correct words in a document.

    1. Find and Replace
    2. Print and Print Preview
    3. Header and Footer
    4. Spelling and Grammar

8. The feature which automatically detects common errors, that is, misspelled words, grammatical errors is called .....

    1. Goto
    2. Thesaurus
    3. AutoCorrect
    4. Spelling and Grammer

9. Which of the following options is used to display information such as page numbers, title of the document, and page number in Word?

    1. Insert Table
    2. Header and Footer
    3. AutoCorrect
    4. Thesaurus

10. A user can specify the margins (top, bottom, left, and right margins) in a document, by choosing Page Setup from the ..... menu.

    1. Format

2. Table
3. AutoCorrect
4. File

11. Which type of justification aligns text on both the right and left margins of a document?

    1. Justify
    2. Center
    3. Right
    4. Balanced

12. To advance to the next cell in a table press the ..... key.

    1. Enter
    2. Tab
    3. End
    4. Insert

13. When you cut and paste something, it is temporarily transferred to: .....

    1. Another partition on the hard drive
    2. Virtual Memory
    3. Swap File
    4. Clipboard

14. A misspelled word is marked by .....

    1. Enclosing the word in a grey bar
    2. Changing the display to a blinking word
    3. Underlining the word with a solid black line
    4. Underlining the word with a wavy red line

15. The Print Preview icon .....

    1. Prints the document without page breaks
    2. Displays statistics about the document
    3. Saves the document in a print format
    4. None of these

State True or False

1. Microsoft Word is a part of Microsoft Office suite.
2. Word facilitates working in more than one document at the same time.
3. Office Assistant helps the user to write text in Word 2003.
4. The status bar shows information like page number, total number of pages, and so on.
5. The green line under the text indicates a grammatical error.
6. Underlining a word is an example of text formatting.
7. When a document is saved, it cannot be modified.
8. The printed copy of a document is called a hardcopy.
9. Headings are used to provide style to the font.
10. Bullets and Numbering buttons are available on the Standard toolbar.
11. Microsoft Word 2003 is an operating system.
12. The Office Assistant button is on the Standard toolbar.
13. The insertion point indicates where text will be inserted as you type.
14. To centre selected paragraphs, click the Center button on the formatting toolbar.
15. Word processor developed by Microsoft is Word Perfect.

Descriptive Questions

1. What are the components of Print Preview toolbar?
2. What do you understand by character and paragraph format?
3. Give description of View menu.
4. Differentiate between Save and Save As option.
5. Differentiate between Undo and Redo command.
6. Give the steps to use Thesaurus.
7. Write short notes on:
    1. Document view
    2. Redo command
    3. Thesaurus
8. Give the steps involved in inserting a picture into a file.
9. What are Headers and Footers in Word?
10. Describe table and its parts briefly.
11. Give steps for creating an invitation letter for your birthday party.

ANSWERS

Fill in the Blanks

1. Print Preview
2. Spelling
3. Ctrl+A
4. Undo
5. Title
6. Standard and Formatting
7. ScreenTip
8. Save and Save As
9. Find and Replace
10. Standard
11. Headers and Footers
12. Tables
13. Toolbar
14. Format Painter
15. Tables and Borders

Multiple-choice Questions

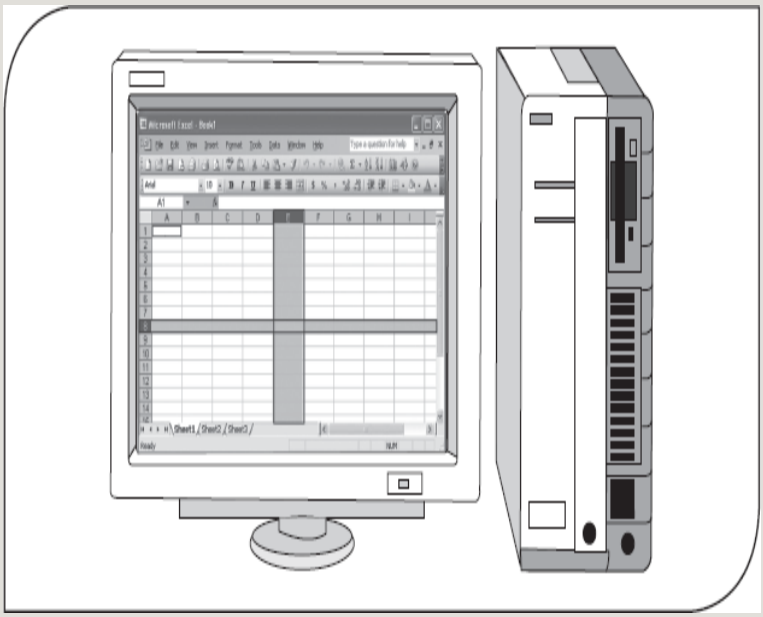
1. (b)
2. (a)
3. (a)
4. (b)
5. (a)
6. (d)
7. (a)
8. (c)
9. (b)
10. (d)
11. (a)
12. (b)
13. (d)
14. (d)
15. (d)

State True or False

1. True
2. True
3. True
4. True
5. True
6. True
7. False
8. True
9. True
10. False
11. False
12. True
13. True
14. True
15. False

Chapter 14

Microsoft Office Excel 2003



This chapter provides a foundation to harness the power of spreadsheet program using Microsoft Office Excel 2003. It enables the user to effectively create worksheets quickly and efficiently. The reader will learn to create and format workbook using various in-built features of Excel. The chapter also covers various tools and features such as formulas and functions, inserting charts, and formatting cells. The chapter goes on to discuss insert formulas and graphics in a worksheet. The chapter concludes with an explanation on how to preview and print an Excel workbook.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The basic concepts of spreadsheets using Microsoft Excel 2003
- The steps required to create, open, save, and close a workbook
- How to enter different types of data in a worksheet
- How to edit the cell contents and the different ways to format a worksheet
- The distinctive feature of Microsoft Excel—formulas and functions to dynamically perform mathematical operations from the data present in worksheets
- How Excel converts data into different types of charts for graphical presentation to provide more visual clarity and impact
- Different options of printing a workbook

14.1 INTRODUCTION

Microsoft Excel is a spreadsheet application that allows you to perform various calculations, estimations, and formulations with data. It is the electronic counterpart of a paper ledger sheet, which consists of a grid of columns (designated by letters) and rows (designated by numbers). Spreadsheets are popular because they represent a better alternative to manually computing mathematical calculations and are more accurate and time saving.

It also provides various facilities like inserting charts, creating graphs, analysing situations, and helps in decision-making. It is one of the best management tools. Excel provides flexibility to the user to manipulate the data without worrying about the size of the data for general applications. Excel 2003 permits a wide selection of options to be used in the creation of worksheets and

allows you to create an impressive spreadsheet presentation.

14.2 STARTING MICROSOFT OFFICE EXCEL 2003

To open Microsoft Excel, perform *any* one of the following steps:

- Double-click the **Microsoft Office Excel 2003** icon located on the desktop.
- Click **start,** point to **All Programs,** then point to **Microsoft Office,** and then select **Microsoft Office Excel 2003** (Figure 14.1).

Figure 14.1 Starting Microsoft Excel

14.2.1 Microsoft Excel Environment

When Microsoft Excel is opened, the main screen of the application appears. This main window contains many parts; these parts are described below in detail (Figure 14.2).

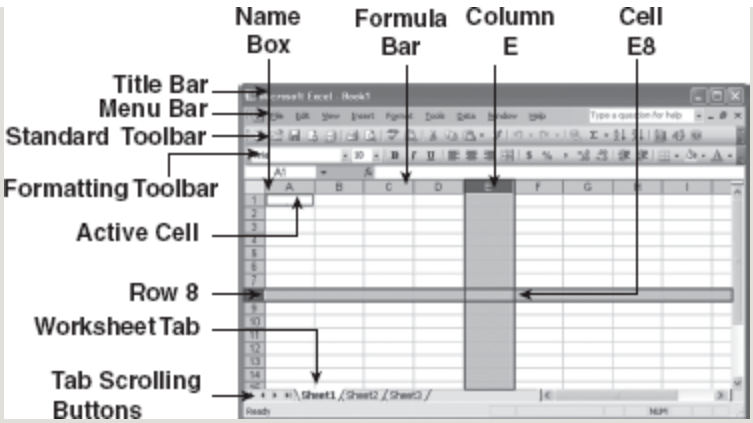
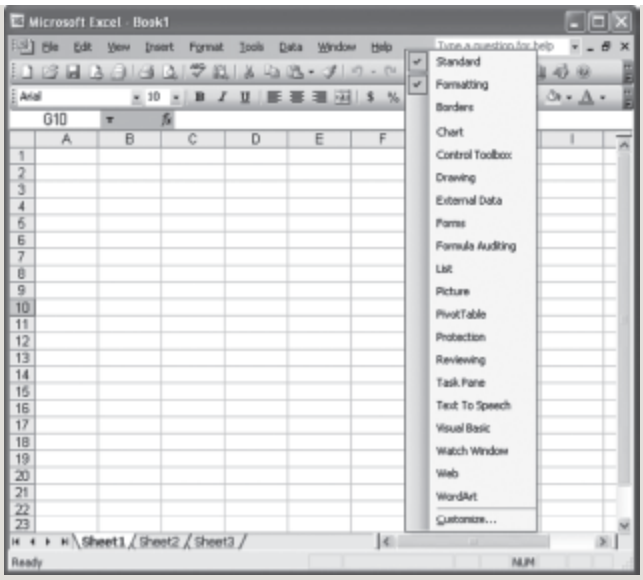


Figure 14.2 Main Microsoft Excel Window

Just like other office applications, Microsoft Excel has a *title bar* located at the top of the application window, which displays the name of the application and the active workbook. Below this bar is the *menu bar,* which contains different menus that have all the options, functions, and commands for the entire Microsoft Excel application. By default, Microsoft Office Excel 2003 contains nine menus, each of which has an associated pull-down menu. For example, the **File** menu contains commands to open, create, and print a workbook.

Generally, a *toolbar* is positioned just below the menu bar. Toolbar contains buttons, menus, or a combination of both that allows you to perform many tasks in a single-click. By default, Standard and Formatting toolbars are displayed in the Excel environment. Additional toolbars like Formula Auditing, and Chart toolbars can be added by right-clicking on the menu bar and selecting the desired toolbar(s) from the pop-up menu, as shown in Figure 14.3.



w

Figure 14.3 Various Toolbars

Name Box and Formula Bar: Generally, the Name Box and Formula Bar are located below the toolbars. The Name Box displays the name of the active cell or selected range and can be used to name a range of cells. The drop-down menu next to the Name Box is used to locate previously named regions. The Formula Bar displays the contents of the selected cell in the worksheet. It includes text, numbers, formulae, and functions. You can also edit the contents of a cell from the formula bar.

Worksheet Tabs and Tab Scrolling Buttons: Worksheet tabs appear at the bottom of the Excel window. Each tab represents a single sheet. Worksheet tabs allow to move from one worksheet to another within the same workbook. The scrolling buttons that appear to the left of the worksheet tab allow you to scroll more quickly and easily through the sheets. You can rename a sheet by right-clicking on the worksheet and selecting **Rename** from the pop-up menu that appears.

Status Bar: The status bar is located at the bottom of the Excel window and displays information about a selected command or an operation in progress. When a command is selected, the left side of the status bar briefly describes the command. It also indicates operations in progress, such as opening or saving a file, copying cells, or recording a macro. The right side of the status bar shows whether keys such as CAPS LOCK, SCROLL LOCK, or NUM LOCK are turned on.

Worksheet and Workbook: Worksheet is the area on a spreadsheet in which you do all the work. A *worksheet* is made up of horizontal rows and vertical columns. An Excel worksheet contains 65,536 rows and 256 columns, and each intersection of a column and row forms a cell.

Individual worksheets are linked to form a *workbook*. Fundamentally, a workbook can be related to a very sophisticated ledger. It allows the user to work and store various kinds of related data within a single Excel file. By default, each workbook contains *three* worksheets (Sheet1, Sheet2, and Sheet3) that can be accessed by sheet tabs. If you need more than three worksheets, choose **Worksheet** from the **Insert** menu and to delete any worksheet, choose **Delete Sheet** from the **Edit** menu.

Columns, Rows, and Cells: In Microsoft Excel, the *column* is defined as the vertical space from top to the bottom of the window. There are 256 columns in a single worksheet which are labelled in alphabetical order (A, B, C…Z, AA, AB, AC…IV). The maximum number of characters, which can be inserted in a column, is 255. A *row* is defined as the horizontal space across the window. The rows are labelled in a numbered order (1, 2, 3, …, 65,536). The maximum limit of the height of a row is 409 points (547 pixels).

The basic unit in spreadsheet is a *cell*. A cell is formed by the intersection of a row and a column. This intersection gives the cell a unique address through which a cell can be referenced through its column name followed by its row number. For example, the intersection of row 2 and column B gives the cell B2. The cell that denotes the current position of the insertion point is known as *active cell* and has a dark border around it. Data written in a cell include text, decimal numbers, date and time, currency, percentage, and scientific notation.

14.3 WORKING WITH EXCEL WORKBOOK

In this section, you will learn the steps required to produce Excel workbook from scratch. This includes:

- Creating a new workbook
- Opening an existing workbook
- Saving a workbook and making a backup copy
- Closing the workbook and Excel application

14.3.1 Creating an Excel Workbook

Whenever you start Excel, it opens a new untitled spreadsheet window so that you can begin a new task. If Excel is already running and you want to create a new workbook, then click on **New** button ( ) on the Standard toolbar. A new workbook can also be opened by following the steps given below:

1. Select **New** from the **File** menu to display the New Workbook task pane (see Figure 14.4).

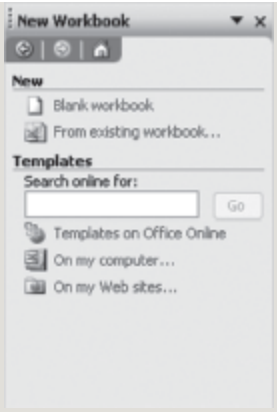


Figure 14.4 New Workbook Task Pane

2. Now select **Blank workbook** under the **New** section. Excel by default opens a new workbook, which is sequentially numbered like *Book1, Book2,* and so on.

14.3.2 Opening an Existing Workbook

To open an existing workbook, follow *any* of the steps given below:

- Select **Open** from the **File** menu to display the Open dialog box (see Figure 14.5). You can also display the Open dialog box by clicking the **Open** button ( ) on the Standard toolbar. The Open dialog box allows you to specify the name of the desired file in the **File name** drop-down box. If the desired file is not in the current location, you can locate it by navigating through **Look in** list box. Once the file is found, select the file and click the **Open** button.

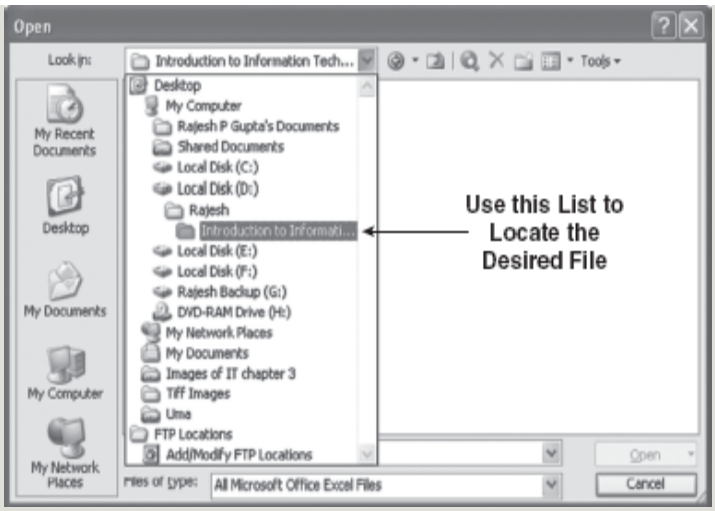


Figure 14.5 open Dialog Box

- Excel workbook can also be opened by double-clicking the Excel file icon, placed on the specified location.

14.3.3 Saving Workbook

Once a workbook is created, you can start typing the text, inserting graphics and charts, and so on. When this is done, you must save the workbook for future reference. To save the workbook, Microsoft Excel provides two menu options, namely, **Save** and **Save As**. If you are working on a untitled workbook, which has never been saved then using any of the option (**Save** and **Save As**) will have the same affect, it displays a Save As dialog box using which you can save the Excel file at the desired location (Figure 14.6).

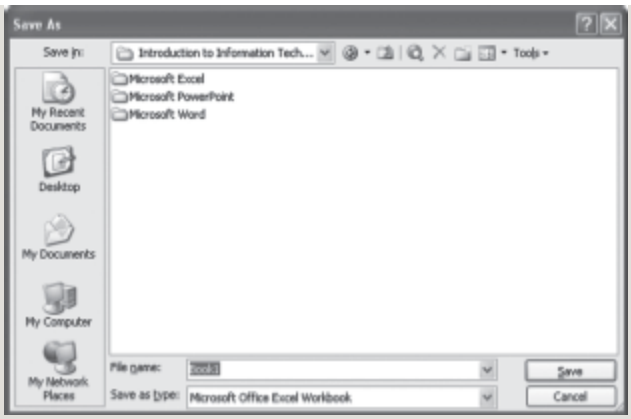


Figure 14.6 Save As Dialog Box

Once you have saved the new workbook, you can use the **Save** option to update changes made to a file. The **Save As** option is used to make multiple copies of the same file.

14.3.4 Closing a Workbook

After finishing all the work in the Excel spreadsheet, you may want to close the workbook. For this, follow *any* of the steps given below:

1. Select **Close** from the **File** menu.
2. Click at the **Close Window** button at the upper-right corner of the Excel window to close the current workbook. Note that there are two close buttons. One is located at the top of the application window, which closes the Excel application, while the second (lower) one present on the workbook closes only the current file (Figure 14.7).

Figure 14.7 Closing a Workbook

Note: *If a user makes changes to a file and has not saved those changes, Excel will ask whether to save the changes before closing the file.*

14.4 WORKING WITH WORKSHEET

In the previous section, we discussed how to create and save an Excel workbook. Now let us learn how to enter data in cells, select, and format cells. Apart from these activities, you will also learn how to insert and delete worksheets, copy and move data, and perform undo and redo operations.

14.4.1 Entering Data in Cells

In Excel, you can enter three types of data in a worksheet: *labels* (text), *values* (number), and *formulae*. Cells can also contain date and time. By default, labels are left aligned while values, time, and date are right aligned. Note that a label can include letters, spaces, punctuation, and numbers. Formulae are used to perform calculations on the values stored in a cell or range of cells. While using values, avoid commas and dollar signs as they are assigned special meaning in Excel (Figure 14.8).

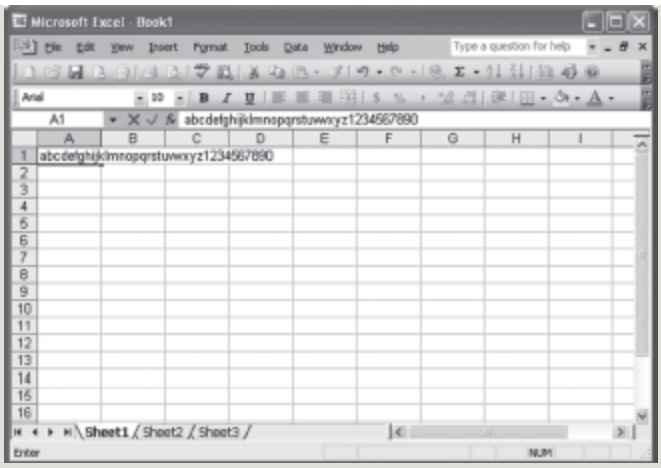


Figure 14.8 Entering Data

To enter data into a cell, first select the cell and then type the data. As you type, notice that the data appear simultaneously in the *Formula Bar* and in the *selected cell*. The data is placed into the cell when you press the **Enter** key, the **Tab** key, or any of the **Arrow** keys. Data can also be entered in the Formula Bar, by selecting the cell and writing the data directly in the Formula Bar. If you enter more data than the cell can display, Excel will either truncate the display of the label or extend it over into the next cell. (In case of values, Excel displays a row of number signs #####.) This happens because the column is not wide enough to show the full content of the cell. To see the full content of the cell, follow any of the steps given below:

- Point the mouse pointer at the border of the cell (boundary line between two columns). When the pointer turns into a double-headed arrow, drag the width of the column to resize the cell.
- Double-click on the cell's right border to automatically adjust the column's width in order to best fit the contents (Figure 14.9).

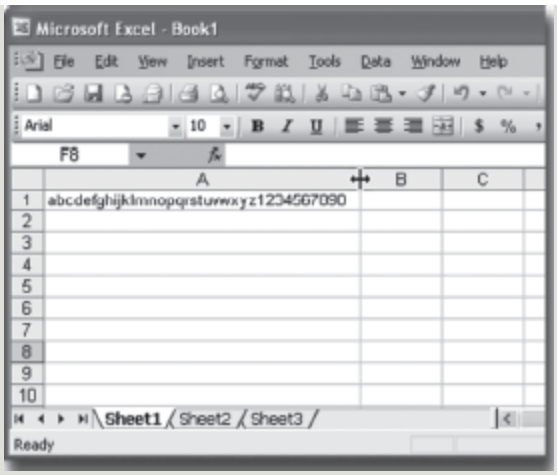


Figure 14.9 Resizing a Cell

14.4.2 Navigating through Cells

You can use either the mouse or the keyboard to navigate the Excel spreadsheet. To use the mouse, simply click in the desired cell. The keyboard offers a wider range of options for jumping to a particular location. The various shortcut keys used to move from one cell to another are listed in the Table 14.1.

Table 14.1 List of Keys Used for Moving in the Worksheet

Keys	Description
Up Arrow Key	Moves one cell up
Down Arrow Key or Enter Key	Moves one cell down
Left Arrow Key	Moves one cell left
Right Arrow Key or Tab Key	Moves one cell right
Ctrl+Right Arrow Key	Goes to the end of the row
Ctrl+Left Arrow Key	Goes to the beginning of the row
Ctrl+Up Arrow Key	Goes to the top row of the sheet
Ctrl+Down Arrow Key	Goes to the bottom row of the sheet
Ctrl+Home	Goes to the top of the worksheet
Page Up	Moves active cell up one screen
Page Down	Moves active cell down one screen

14.4.3 Naming of a Range of Cells

A range of cells is formed by selecting a group of adjacent cells in a worksheet. Naming a cell or a range of cells adds clarity and speeds up productivity especially when dealing with a very large spreadsheet. To name a range of cells, follow the steps given below:

1. Select the cell or the range of cells.
2. Click inside the **Name Box** to highlight the existing name of the cell.
3. Type in the name for the selected cell or the range of cells and press the **Enter** key (Figure 14.10).

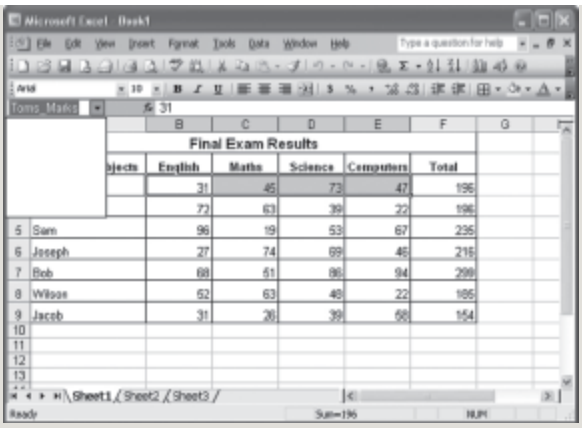


Figure 14.10 Naming a Range

14.4.4 Editing a Worksheet

While working with worksheets you need to edit the cell contents, add, or delete cells, rows, and columns, and so on. Excel provides a number of ways to format a spreadsheet. This includes changing fonts, colours, borders, etc.

Selecting Cells, Rows, and Columns: Whenever you want to make a change to a cell or a set of cells, you must first select it. To select an area, click on one cell, hold down the mouse button, and drag across the cells you want to include in the selection. Notice that as you move the cursor next to or onto a border, it changes from a white cross (✚) to an arrow-pointer. If you wish to select a large area of *adjacent* cells, use the **Shift** key to extend the selection. Click on the first cell of the range you want to select; then, while holding down the **Shift** key, click on the last cell in the range you want to select (Figure 14.11).

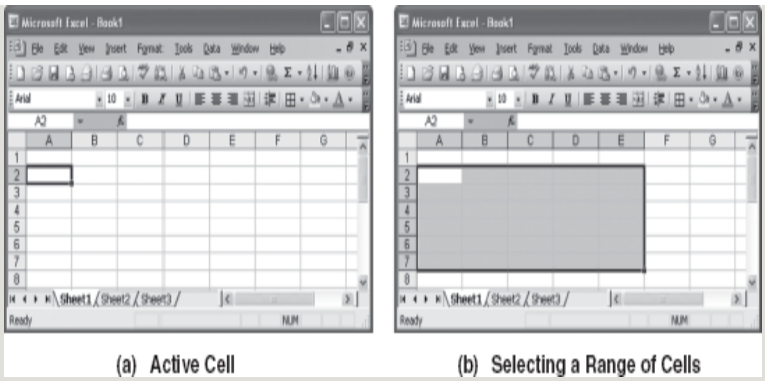


Figure 14.11 Selecting a Cell

An entire row or column can be selected by clicking on the desired row or column. For example, to select row "5" or the column "D," click on the row heading or the column heading (see Figure 14.12).

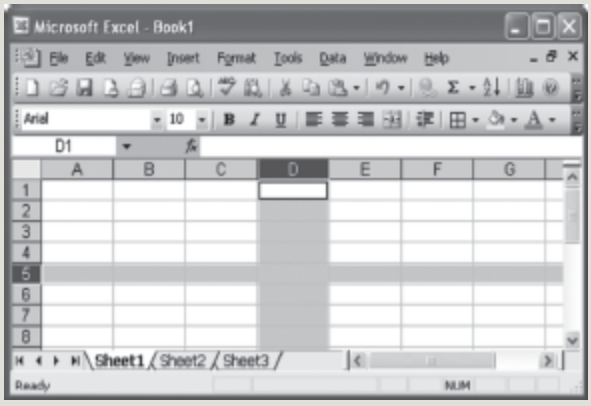


Figure 14.12 Selecting a Row and Column

Excel 2003 also allows for multiple selections and more than one range may be selected at the same time, that is,

non-adjacent range of cells can also be selected. For this, select the first range of cells; then, while holding down the **Ctrl** key, select the second range of cells. If a user wants to select the entire worksheet, then this can be done by clicking on the grey square at the top-left of the spreadsheet (Figure 14.13).

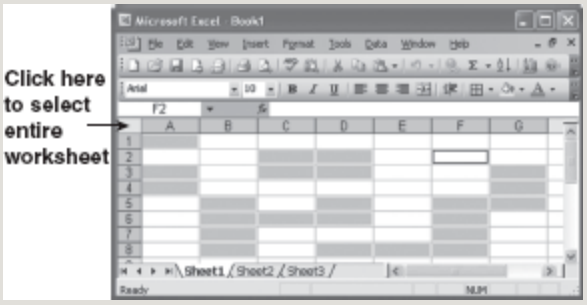


Figure 14.13 Selecting Non-Adjacent Range of Cells

Inserting Cells, Rows, and Columns: While working in Excel, users may be required to insert cells, rows, or columns to add new formulae or data. To insert cells, rows, or columns, select the cell and follow *any* of the steps given below:

- Select **Cells** from the **Insert** menu.
- Right-click on the cell and select **Insert** from the pop-up menu that appears.

When you perform any of the above-mentioned action, Microsoft Excel will display the Insert dialog box (see Figure 14.14).

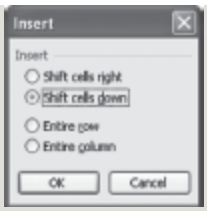


Figure 14.14 Insert Dialog Box

Choose any one option from the list and click **OK**. The various options present in this dialog box are listed in Table 14.2.

Table 14.2 Options of Insert Dialog Box

Options	Description
Shift Cells Right	Moves the active cell to the right and inserts a cell in its place
Shift Cells Down	Moves the active cell down and inserts a cell in its place
Entire Row	Moves all cells of the row containing the active cell down and inserts a row
Entire Column	Moves all cells of the column containing the active cell right and inserts a column

Deleting Cells, Rows, and Columns: Deletion removes the entire row or column from the spreadsheet. This action is similar to removing a Rows and Columns Inserting: Rows and columns can also be added by selecting Rows and Columns, respectively, from the **Insert** menu. The new row will be inserted above the row you have selected, and the new column will be inserted to the left of the selected column. The row numbers and column letters will change accordingly.

THINGS TO REMEMBER

Rows and Columns

Inserting: Rows and columns can also be added by selecting **Rows** and **Columns**, respectively, from the Insert menu. The new row will be inserted above the row you have selected, and the new column will be inserted to the left of the selected column. The row numbers and column letters will change accordingly.

Deleting: Multiple rows and columns can be deleted by first selecting one row or column and then holding down the **Shift** key and using arrow keys to select multiple rows and columns and then select **Delete** from the **Edit** menu or right-click on the cell and select **Delete** from the pop-up menu that appears.

Deleting Cells, Rows, and Columns: Deletion removes the entire row or a column from a Word table. To delete cells, rows, or columns, first select the cell and then follow *any* of the steps given below:

- Select **Delete** from the **Edit** menu.
- Right-click on the cell and select **Delete** from the pop-up menu that appears.

When you perform any of the above-mentioned action, Microsoft Excel will display the Delete dialog box (see Figure 14.15).

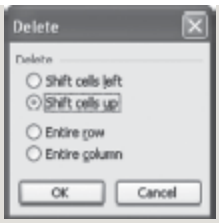


Figure 14.15 Delete Dialog Box

Choose from the list of options available and click **OK**. The various options present in this dialog box are listed in Table 14.3.

Table 14.3 Options of Delete Dialog Box

Options	Description
Shift Cells Left	Deletes the active cell and moves all the remaining cells to the left
Shift Cells Up	Deletes the active cell and moves all the remaining cells up
Entire Row	Deletes the entire row containing the active cell
Entire Column	Deletes the entire column containing the active cell

Editing Cell Contents: If you want to change the contents of the cell then, follow any of the steps given below:

- Single-click the cell to re-type the contents of the cell.
- Press **F2** or double-click the cell to modify the contents of the cell.
- Use the arrow keys to select the cell and edit the contents of the cell.
- Select the cell to display the contents of the cell in the Formula Bar. Now, you can edit the contents of the cell directly in the Formula Bar.

When you perform any of the above-mentioned actions, press **Enter** key or click anywhere in the worksheet to

accept the editing, or press **Esc** to cancel it.

Deleting Cell Contents: If you want to delete the contents of a cell, select the cell and follow *any* of the steps given below:

- Select **Clear** from the **Edit** menu and choose **Contents** option from the pop-up menu.
- Press **Delete** key.

Formatting Cells: Formatting data in Excel 2003 is similar to doing formatting in other Microsoft Office applications. Formatting is applied to the cells in order to change the appearance of the data stored in those cells. It is applied by altering the appearance of the cell by setting the alignment, typeface (font), size, style, and colour. Formatting can be done by using the Formatting toolbar or by using the **Format Cells** dialog box. To format the cell, follow the steps given below:

1. Select the cell to be formatted.
2. Use buttons on the Formatting toolbar to format the cell. Some of the buttons used in formatting cell are given in Table 14.4.

Table 14.4 Formatting Cell

Command	Button	Description
Font		Changes the font of cell
Font Size		Changes the size of the font
Bold		Selected text appears in boldface
Italic		Italicizes the selected text
Underline		Underlines the selected text
Align Left		Left aligns the paragraph or the selected text
Center		Centre aligns the paragraph or the selected text
Align Right		Right aligns the paragraph or the selected text
Merge and Center		Merges and centres the contents of two or more cells
Decrease Indent		Decreases the indent of the selected paragraph
Increase Indent		Increases the indent of the selected paragraph
Fill Color		Changes the colour of the cell
Font Color		Changes the font colour of the selected text
Currency Style		Applies currency style to the selected cell(s)
Percent Style		Applies percent style to the selected cell(s)

Command	Button	Description
Comma Style		Applies comma style to the selected cell(s)
Increase Decimal		Increases the number of digits after the decimal point
Decrease Decimal		Decreases the number of digits after the decimal point
Borders		Applies borders to the selected cell(s)

Cells can also be formatted using the Format Cells dialog box. For this, follow the steps given below:

1. Select the cell(s) to be formatted.
2. Right-click on the cell and select **Format Cells** from the shortcut menu. This displays the Format Cells dialog box (see Figure 14.16).

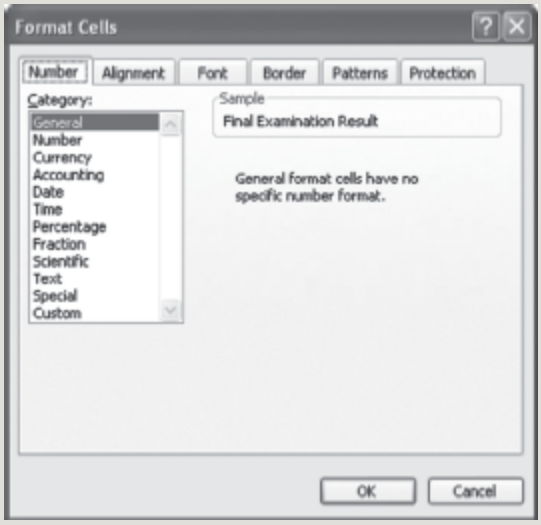


Figure 14.16 Format Cells Dialog Box

This dialog box contains a number of tabs to format the cell. These are *Number, Alignment, Font, Border, Patterns,* and *Protection* tabs. By default, this dialog box displays Number tab when opened.

Number tab Excel permits numbers to be formatted in many ways without changing the value of the number in a cell; number formats allow numbers to be represented so that they can be used in different kinds of scenarios. Different number formats are available under this tab. Some of the commonly used number formats are *General, Number, Currency, Accounting, Percentage*, and *Fraction*.

Alignment tab The options present in this tab allow you to change the position and alignment of the data within the cell. Data can be aligned with any or all four sides of a cell. Text can be aligned horizontally by selecting the various options such as **Left (Indent), Center, Right (Indent), Fill, Justify, Center Across Selection**, and **Distributed (Indent)**. The selected cells can also be aligned vertically. This can be done by selecting the **Top, Center, Bottom, Justify**, or Distributed option in the **Vertical** dropdown box in the Alignment tab.

Note: *The default alignment in Microsoft Excel 2003 aligns text to the left, numbers to the right, and centres logical and error values.*

You can also change the orientation of text in a cell (from top to bottom, bottom to top, left to right, etc.). This can be done by entering either the exact amount of rotation required into the **Degrees** text box or drag the **Text** dial to give the desired level of rotation under the **Orientation** section.

Font tab The options present in this tab allow you to change font attributes such as font, size, style, etc. The

Font tab can be used to set the font colour, underline text, and apply effects like, subscript, superscript, etc.

Border and patterns tabs The options present in border tab allow you to apply borders to the selected cells. There are different types of styles available to add border to selected cells. Click on the **Outline** button to add a basic border outside the selected cell range, or click on the **Inside** button to add borders between cells within the selected cell range. The options **present** in the Patterns tab allow you to apply shading to the selected cells.

Protection tab The options present in this tab allow you to lock the selected cell. However, it will have no effect unless the worksheet is protected. For this, select Protection from the **Tools** menu and choose **Protect Sheet** (Figure 14.17).

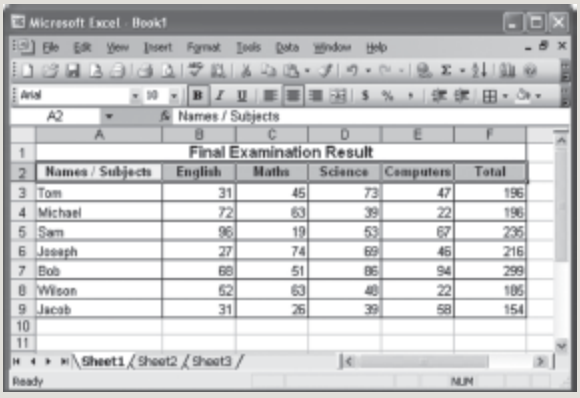


Figure 14.17 Formating cells

Using AutoFormat: Excel provides a feature known as *AutoFormat* that enables you to apply pre-defined layouts to the selected tables in the worksheet. These layouts do not alter the position of your data but apply colour backgrounds and rearrange borders with some

attractive effects. To AutoFormat a table, select a range of the cells and follow the steps given below:

1. Select **AutoFormat** from the **Format** menu to display the AutoFormat dialog box (Figure 14.18).

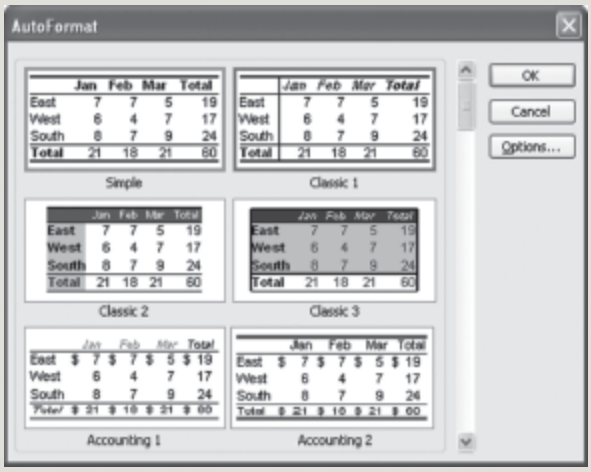


Figure 14.18 Using AutoFormat

2. From this dialog box, choose the appropriate option and click **OK** to format the table accordingly (see Figure 14.19).

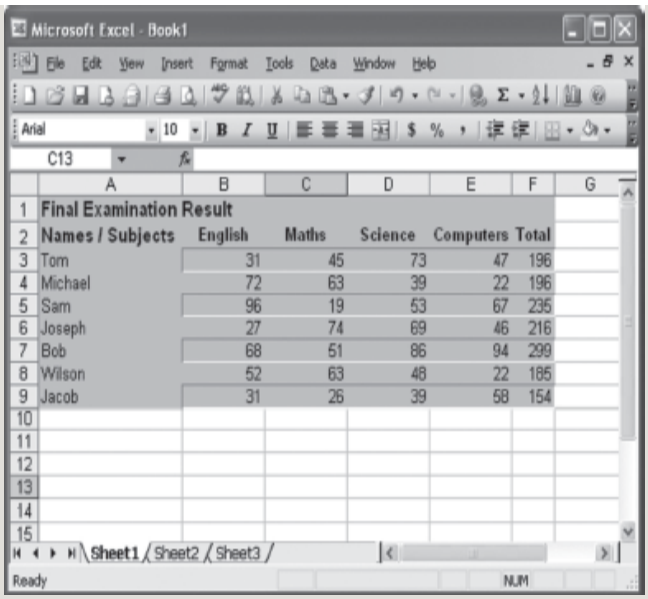


Figure 14.19 Cells After AutoFormat

If you want to revert to the original appearance of the cell, select **Undo AutoFormat** from the **Edit** menu.

14.4.5 Using Cut, Copy, and Paste in Excel

When you are working in a worksheet, you may decide to move the contents of one cell or a range of cells to some other part of the spreadsheet. If the text is to be moved within the same worksheet or to another worksheet, it should be first cut and then pasted. When you choose cut or copy, the cells are surrounded by a flashing dotted line, and their contents are not actually moved until you click in the cell where you want to paste. The contents of the selected cells do not disappear as they do in Word 2003. Copying the text is similar to the cut operation, except that copy retains the text at the original place but in case of cut, the text is removed from the original cell.

To perform the cut, copy, and paste operations, follow the steps given below:

1. Select the cell or the range of cells to be moved or copied.
2. In case you want to move the cell contents, choose **Cut** from the **Edit** menu or click the **Cut** button ( ) on the Standard toolbar. If you want to copy the contents, choose Copy from the Edit menu or click the **Copy** button ( ) on the Standard toolbar. Note that a dotted/fickering line surrounds the area that is cut or copied.
3. Click the mouse on the place in the workbook where you want to insert the text.
4. Choose **Paste** from the **Edit** menu, or click the **Paste** button ( ) on the Standard toolbar. The text that you copied to the clipboard is pasted to the location where the mouse is clicked (Figure 14.20).

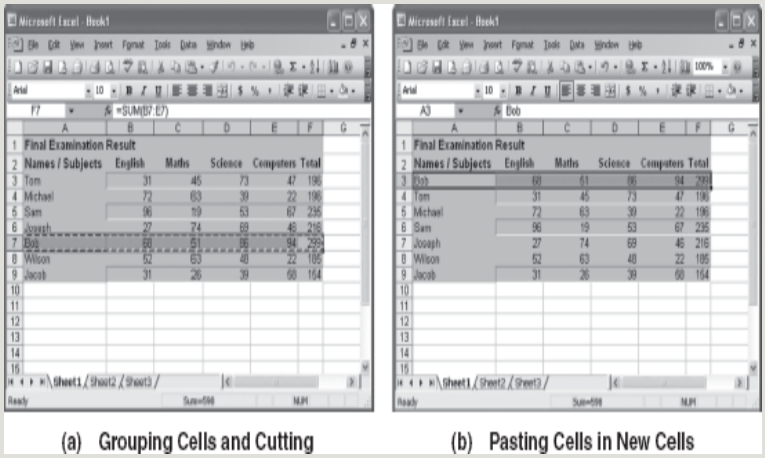


Figure 14.20 Using Cut and paste

Copying Cells Using Fill Handle: In Excel, the *Fill Handle* provides an easy method of copying contents of the selected cells to the adjacent cells in a column or in a row. Fill handle appears on the worksheet when you move your mouse over the right bottom corner of the active cell. To copy a cell to the adjacent cells using the fill handle, follow the steps given below:

1. Select the cell or range of cells to be copied.
2. Point at the fill handle. The mouse pointer changes to black plus sign ( ).
3. Drag the fill handle in the direction of the copy until the faded rectangle surrounds all the cells to be filled.
4. Release the mouse button to copy the contents (Figure 14.21).

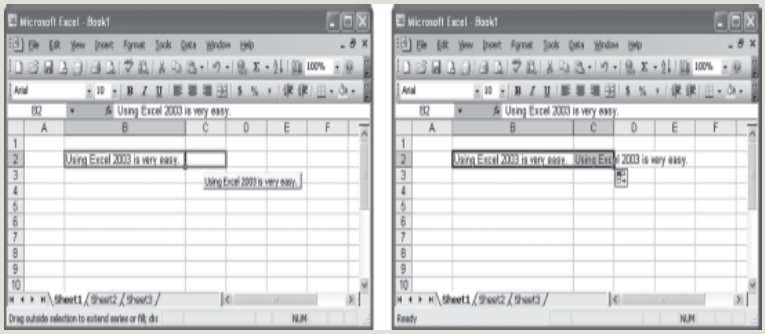


Figure 14.21 Using Fill Handle

THINGS TO REMEMBER

Linking Worksheets

In Excel, moving the information from a cell or a range of cells, rows, or columns is not just limited to a worksheet. Microsoft Excel 2003 gives the flexibility to link value from a cell in one worksheet to another. For example, the value of cell A1 in the worksheet1 and cell A2 in the worksheet2 can be added using the format "sheetname!celladdress." The formula for this would be "=A2+ [worksheet1.xls]sheet1!A1."

Dragging Cells: Cell contents can also be moved to another cell by simply dragging the cell to point to the desired cell. To move a range of cells using drag and drop, follow the steps given below:

1. Select the cell or the range of cells.
2. Point at the border around the selection to change mouse pointer to a pointing arrow.
3. Drag the border to the new location and release the mouse button (Figure 14.22).

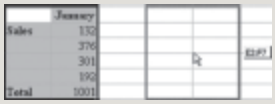


Figure 14.22 Dragging cells

To drag the selected cells to a different sheet, hold down the **Alt** key and drag the selected area onto the sheet label. Excel will then display that sheet and you can position your selected cells as normal. Using this technique, you can restructure your spreadsheet at any stage in its development.

Sometimes, you may wish to create a list of numbers in cells. When the list of numbers is large say 1,000, then the process of entering the number manually becomes very time-consuming and exhausting. To automate this process, Microsoft Excel 2003 also allows you to fill series of data in an “intelligent” manner. For example, type any number in a cell, then type the next immediate number in the adjacent cell. To follow the same series pattern to a range of cell, select both cells and drag to the desired range using fill handle (Figure 14.23).

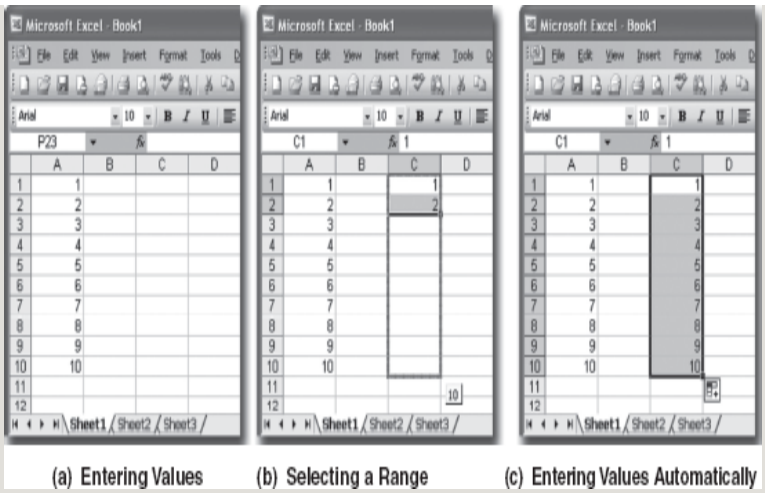


Figure 14.23 Inserting Series of Data

14.4.6 Finding and Replacing Cells

Find and replace option in Microsoft Excel is used to locate a particular word, phrase, or numeric value in a worksheet, or replace it with the new data. Microsoft Excel Find and Replace function swiftly and unerringly locates anything you are looking for, and once the desired text is located, it can automatically be replaced by the new data.

Data in Excel can be searched column- or row-wise in a worksheet. To find data within the worksheet, follow the steps given below:

1. Select **Find** from the **Edit** menu, which opens the Find and Replace dialog box with Find as active tab (see Figure 14.24).

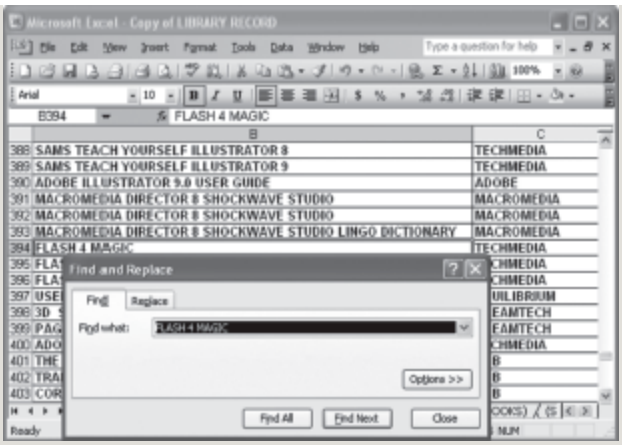


Figure 14.24 Find and Replace Dialog Box

2. Type the data that you want to search in the **Find what** box and click Find Next button to find the first occurrence of the data.
3. Keep on pressing the **Find Next** button until you are finished with finding data.
4. To close the Find and Replace dialog box, click the **Close** button or press the **Esc** key.

To replace data within the worksheet, follow the steps given below:

1. Select **Replace** from the **Edit** menu, which opens the Find and Replace dialog box with **Replace** as active tab (see Figure 14.25).

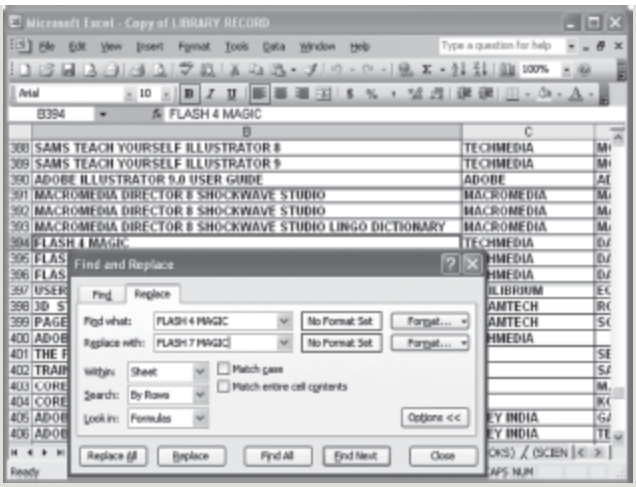


Figure 14.25 Find and Replace Dialog Box

1. Type the data that you want to search in the **Find what** box.
2. Click **Find Next** button to find the first occurrence of the data. If you want to replace the data then type the replacement data in the **Replace with** box.
3. Select **Replace** button to replace each occurrence of the word individually. Keep on pressing the Find Next button until you are finished with finding and replacing the data.
4. Click **Replace All** to replace all the occurrences of the data at once. Excel will display a message when it has replaced all the occurrences.
5. To close the Find and Replace dialog box, click the **Close** button or press the **Esc** key.

14.4.7 Undo and Redo

Just like Word 2003, Excel remembers the actions performed on a file. If a portion of a worksheet is deleted or changed, you can revert to the original state by using the Undo command. This feature instructs the application to ignore the last action (whether it was deleting, copying, or moving). Note that Excel can undo or redo up to previous 16 actions (since last Save state) only as compared to Word, which allows user to undo or redo all actions until the document is closed (Figure 14.26).

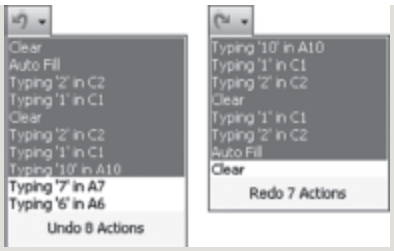


Figure 14.26 Undo and Redo Action

To undo the last action(s), click the **Undo** button ( ) on the Standard toolbar, or select **Undo** from the **Edit** menu. If you want to undo a number of actions at the same time then click the down arrow ( ) beside the **Undo** button to display a list of actions that can be undone.

If an undo action is set and you want to reverse it then *Redo* command can be used to reverse the undo action. To redo the last undo action, click on the **Redo** button ( ) on the Standard toolbar, or select **Redo** from the **Edit** menu. If you want to redo a number of actions at the same time then click the down arrow ( ) beside the **Redo** button to display a list of actions that can be redone.

14.5 FORMULAS AND FUNCTIONS

One of the distinguishing features of Excel is that it makes use of formulas and functions to dynamically calculate results from the data present in worksheets. Functions are routines built into the Excel spreadsheet while formulas are defined by the user and may include the built-in functions. Both Functions and formulas are widely used in simple as well as in advance computing. They provide the power to analyse data extensively.

Functions can be a more efficient way of performing mathematical operations than formulas. For example, if you want to add the values of cells C1 to C10, then using the formula you need to type "=C1+C2+C3+C4+C5+C6+C7+C8+C9+C10." This can be tedious, a better, and a shorter way would be to use the SUM function and simply type "=SUM(C1:C10)."

14.5.1 Functions

Microsoft Excel contains many pre-defined or built-in formulas, which are known as *functions*. These can be used to perform simple or complex calculations. They perform calculations b using specific values, called *arguments*, in a particular order.

Parentheses are used to separate different parts of a formula. For example, in the formula =SUM(A1:A5), the parentheses separate the worksheet function from the cell references that the function is referring to. This is particularly important in longer or more complicated formulas, for example =((A2/4)+(A5-B3))*5. If a mistake is made and the parentheses in a formula do not match, an error message is displayed. Note that the parts of a formula contained inside the parenthesis are calculated first. Some examples of functions are =SUM(B4,G43,T70),=COS(A2), =AVERAGE(B1:B10).You can type functions in the Formula Bar directly into the cell or use the Function Wizard.

THINGS TO REMEMBER

Functions

Functions have three parts:

- = sign, which tells Excel that a formula or function follows.

- Function name, such as SUM for addition or AVERAGE for determining the average of a series of numbers.
- Arguments on which the particular function operate. The argument contains cell references to let the function know which data to calculate. In Excel, a function can accept a maximum of 30 arguments. The argument must also be enclosed by parentheses.

To insert a function, follow the steps given below:

1. Click the cell in which the function is to be entered.
2. Select **Function** from **Insert** menu, or click **Insert Function** button ( ) on the Formula Bar to display the Insert Function dialog box (Figure 14.27).

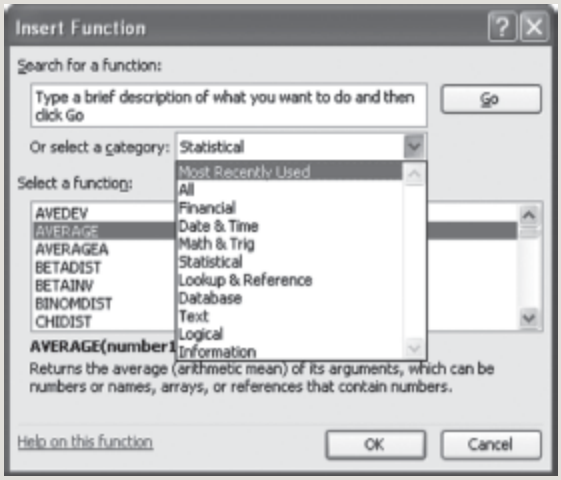


Figure 14.27 Insert Function Dialog Box

3. Select a desired function category from the **select a category** drop-down box and choose the function name from the **Select a function** list which contains a list of available functions in the selected category. Once the desired function has been selected, click **OK**.
4. After this, Excel displays a Function Arguments dialog box to help the user to create a function. In this dialog box, first click on the collapse button (labelled with a red arrow) to the right of the box labelled Number1 or Value1 (this depends on the function chosen) (Figure 14.28).

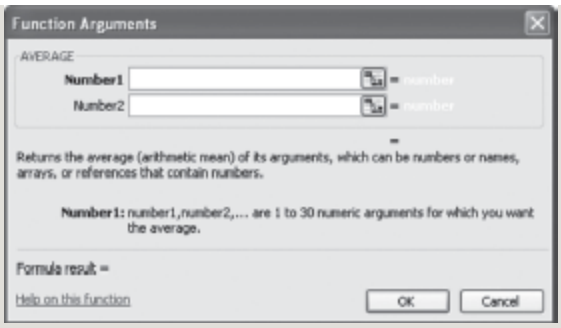


Figure 14.28 Function Arguments Dialog Box

5. Drag the mouse to select the range of cells to be included as the first argument of the function and press **Enter** key.
6. To specify additional arguments into the function, repeat Steps 4 and 5.
7. Click **OK** to insert the function (Figure 14.29).

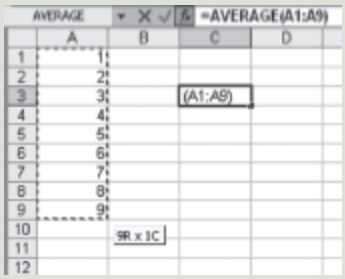


Figure 14.29 Inserting Function in Worksheet

Some of the categories of functions provided by Excel are:

- Math and Trig Functions
- Statistical Functions
- Date and Time Functions
- Logical Functions
- Text Functions

Math and Trig Functions: A user can perform simple calculations, such as rounding a number or calculating the total value for a range of cells with the help of mathematical and trigonometric functions. The various mathematical functions are listed in Table 14.5.

Table 14.5 Mathematical Functions

Formula	Description	Example
SUM(number1, number2,...)	Adds all the numbers in a range of cells	SUM(10,20,30) *Output:* 60
ROUND(number, num_digits)	Rounds off a number to specified places of digits	ROUND(3.628, 2) *Output:* 3.63
SQRT(number)	Returns a positive square root of a given number. If the number is negative, SQRT returns the #NUM! error value	SQRT(25) *Output:* 5
ABS(number	Returns the absolute value of a number	ABS(-100) *Output:* 100
TRUNC(number, num_digits)	Truncates a number to an integer by removing the fractional part of the number	TRUNC(8.99,1) *Output:* 8.9

Logical Functions: Excel has a number of functions that allow the user to evaluate values and make decisions based on the result of the evaluation. These functions are known as **logical functions.** The logical function returns either true or false value depending on the condition. The various logical functions used are listed in Table 14.6.

Table 14.6 Logical Fucntions

Formula	Description	Example
AND(logical1, logical2, ...)	Returns value TRUE if all its arguments are true and returns FALSE if one or more arguments are false	AND(1+2=3, 2-2=0, AND(1+2=3, 2-2=1) *Output:* False
NOT(logical)	Reverses the value of its argument	NOT(1+1=2) *Output:* False NOT(1+1=1) *Output:* True
OR(logical1, logical2,...)	Returns TRUE, if any argument is TRUE and returns FALSE only when all arguments are FALSE	OR(1+2=3, 2-2=1) *Output:* True OR(1+2=2, 2-2=1) *Output:* False

Statistical Functions: In addition to mathematical functions, Excel provides a great deal of assistance to ocompute statistical data. All the functions are simple and take only a set of obser vations as arguments. The various statistical functions used are listed in Table 14.7.

Table 14.7 Statistical Functions

Formula	Description	Example
MAX(number1, number2,...)	Returns the largest value in a given set of values	MAX(60,25,5) *Output:* 60
MIN(number1, number2,...)	Returns the smallest value in a given set of values	MIN(60,25,5) *Output:* 5
AVERAGE(number1, number2,...)	Calculates the arithmetic mean of all values in the list	AVERAGE(60,25,5) *Output:* 30

Text Functions: In Excel, text functions are used not only to convert a value to text but also to join several text strings into one text string. Many functions are available that enable you to manipulate text strings, convert numeric entries into text strings, and convert numeric text entries into numbers. The various text functions used are listed below in Table 14.8.

Table 14.8 Text Functions

Formula	Description	Example
CONCATENATE (text1, text2,...)	Joins several text strings into one text string	CONCATENATE (“GrandTotal”, “Total”) *Output:* Grand Total
LEN(text)	Returns the number of characters in a text string	LEN(“INDIA”) *Output:* 5
LOWER(text)	Converts all upper case letters in a text string to lower case	LOWER(“INDIA”) *Output:* India
PROPER(text)	Capitalizes the first letter in each word of a text	PROPER(“INDIA”) *Output:* India
UPPER(text)	Converts a text string to upper case	UPPER(“india”) *Output:* INDIA
TRIM(text)	Removes all spaces from a text string except for single spaces between words	TRIM(“INDIA”) *Output:* INDIA

Date and Time Functions: The date and time functions are used for working with date and time. Excel uses serial numbers to store dates, giving each day of each year a unique number. The var ous date and time functions used are listed below in Table 14.9.

Table 14.9 Date and Time Functions

Formula	Description	Example
DATE(year, month, day)	Returns the number that represents the date in Excel date-time code	DATE(1979,9,6) *Output:* 9/06/79
TIME(hour, minute, second)	Returns the number that represents a particular time	TIME(19,23,7) *Output:*7:23 PM
NOW()	Returns the current date and time	NOW() *Output:* 4/30/10 8:39

14.5.2 Using AutoSum

The SUM function is used more often than any other function. To make this function more accessible, Excel includes an **AutoSum** button on the Standard toolbar, which inserts the SUM function into a cell. It is a great tool to use when you want to quickly add the contents of a range of cells. To use AutoSum, follow the steps given below:

1. Select the cell where you want the sum to appear.
2. Click the **AutoSum** button (Σ) on the Standard toolbar. AutoSum inserts a formula that uses the SUM function. Clicking on the AutoSum, the cells get surrounded by a flashing dotted line. This dotted line is called a **marquee**. Excel puts this around the range of cells you want to add up and inserts the range reference in the formula.
3. If this is the correct range, then press the **Enter** key. If not, type or highlight the correct range and press the **Enter** key or **Enter** button (✓) to apply the formula (Figure 14.30).

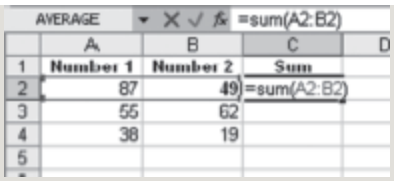


Figure 14.30 Using AutoSum

14.5.3 Formulas

Formulas are mathematical expressions built in Excel that instruct the computer to carry out calculations on specified sets of numbers in the rows and columns. A formula always begins with an equal sign (=) followed by some combination of numbers, text, cell references, and operators. If a formula is entered incorrectly, an *ERROR IN FORMULA* message will appear. If you forget to enter the initial (=) sign, Excel will treat the expression like a text string and the values will not be calculated. Note that Excel evaluates a formula in a specific order: from left to right following the order of operations.

Suppose, cell A1 contains 25 and cell A2 contains 90 and you want to add values in cell A1 and A2 in A3. For this, follow the steps given below:

1. Select the cell where formula is to be inserted. In our case, the cell is A3.
2. Type "=" followed by the operation (say, SUM) to be performed.
3. Type the first and second cell names, separated by a colon (A1:A2).
4. Press the **Enter** key or click the **Enter** button ( ) on the Formula Bar. Now the formula appears in the Formula Bar while the cell (A3) contains the result of the formula as shown in Figure 14.31.

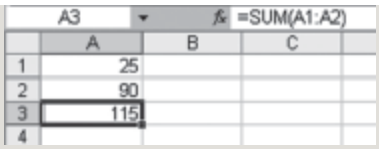


Figure 14.31 Using Formulas

Relative, Absolute, and Mixed Referencing:
Referring to cells by their column and row labels (such as "A1") is called *relative referencing*. When a formula contains relative referencing and it is copied from one cell to another, Excel does not create an exact copy of the formula. It will change cell addresses relative to the row and column they are moved to. For example, if a simple addition formula in cell C2 =SUM(A2:B2) (see Figure 14.32[a]) is copied to cell C3, the formula would change to =SUM(A3:B3) to reflect the new row. To prevent this change, cells must be called by *absolute referencing*, and this is accomplished by placing dollar signs "$" within the cell addresses in the formula.

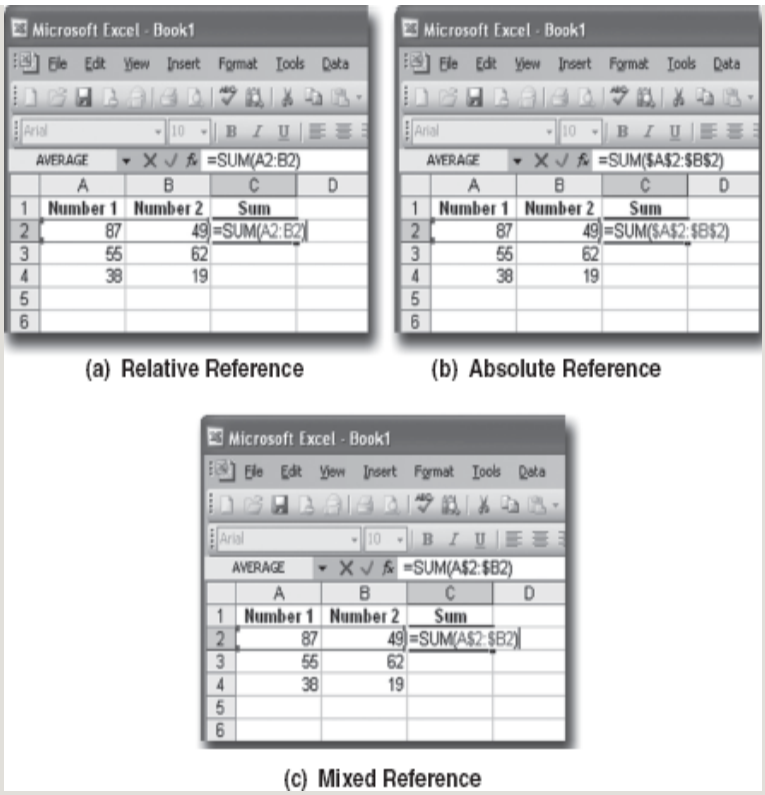


Figure 14.32 Relatice, Absolute and Mixed Referencing

For example, if the formula in cell C2 would read =SUM($A$2:$B$2) (see Figure 14.32[b]) in which both the column and row of both cells are absolute, the formula will not change when copied. *Mixed referencing* can also be used where only the row or column is fixed. For example, in the formula =SUM(A$2:$B2) (see Figure 14.32[c]), the cell reference A$2 indicates relative column and absolute row and the cell reference $B2 indicates absolute column and relative row.

Editing and Deleting Formulas: A formula can be edited or deleted if required. To delete a formula, simply click on the cell that contains the formula and press the

Delete key. If you want to alter the formula, follow the steps given below:

1. Click the cell that contains the formula.
2. Click in the Formula Bar, change the formula and press the **Enter key.**

Handling Operators in Formula: In Microsoft Excel 2003, operators specify the type of calculation that is to be performed on numbers or quantities. Excel includes four different types of calculation operators:

- Arithmetic Operator
- Text Operator
- Comparison Operator
- Reference Operator

All these operators have been listed in Table 14.10.

Table 14.10 Operators Used in Excel Formulas

Operator	Meaning
*	Multiplication
/	Division
+	Addition
–	Subtraction
%	Per cent
^	Exponentiation
=	Equal to
>	Greater than
<	Less than
>=	Greater than or equal to
<=	Less than or equal to
<>	Not equal to
&	Concatenates or combines two values to produce one continuous text value
: (colon)	Known as range operator. Produces one reference to all the cells between two references, including the two references, for example, D3:D7
, (comma)	Known as union operator, which combines multiple references into one reference, for example, SUM(D3:D7, F15, B4)
(space)	Known as intersection operator, which produces on reference to cells common to the two references (A7:C7 B6:B8)

Error Values: An error may occur in Excel while working with formulas when the intended task is not carried out in a proper way. For example, if you intend to add cells (A1: A4) and one of the cells, say A3, contains text, then the correct output will not be shown in the selected cell and an error state is said to have occurred. The errors may also occur due to some other reasons. Some of the common errors are listed in Table 14.11.

Table 14.11 Error Values

Type of Error	Description
######	This is not exactly a kind of error; it just specifies that the result is too long to fit in the selected cell. This error can be corrected by making the column wider
#DIV/0!	This type of error occurs when a number is divided by zero. The error can be corrected by making sure that the divisor is not zero
#NAME?	This type of error occurs when there is a name in the formula that Excel cannot recognize. This error can be avoided by selecting a name from the Name Box instead of typing it. If you type a function, check the spelling and verify that the function exists or if users perform operations on text, enclose it in double quotation marks.
#REF!	This type of error occurs when cell reference is invalid. For example, when the user deletes cells referred to in a formula or cells are pasted and moved elsewhere. This error can be corrected by entering the formula again
#VALUE!	This type of error occurs when wrong operands or arguments are used within a formula. This error can be corrected by checking the passed arguments again and if wrong then passing them again

14.6 INSERTING CHARTS

In Excel, numerical data can be easily converted into a chart for graphical presentation of the data. *Charts* provide more visual clarity than tables of data and, therefore, have more impact. Several types of chart are possible in Microsoft Excel, and each type has a number of variations. The chart types include pie charts, bar charts, line charts, etc. Whatever the type, all charts are linked to the worksheet data. This means that data values can be amended even after the chart has been created and the chart will automatically be updated.

Microsoft Excel allows a chart to be placed in two ways, as an *embedded chart* or a *chart sheet*. Embedded charts appear as objects on worksheets and can be resized and positioned to appear alongside tabular information or other charts and objects, whereas chart sheet exists as a separate sheet in a workbook. Chart sheets are, perhaps,

a little easier to manipulate but when printed, they appear alone on the page. It is, therefore, possible to adjust the page orientation, headers, footers and other page attributes for the chart without affecting other worksheets.

Types of Chart: Excel allows you to create charts of different types. You can choose from the list of chart types available. Some of the commonly used chart types are:

- **Area ( ):** This is used when a user wants to emphasize change over time.
- **Surface ( ):** In this chart, temperature and time are plotted together to show the tensile strength they produce.
- **Bar ( ):** This chart compares values with the different values given.
- **Radar ( ):** In this chart, each category of information has its own line radiating out from the centre.
- **Column ( ):** This chart is very similar to a bar chart, except that the bars are aligned vertically instead of horizontally.
- **Bubble ( ):** This chart shows three sets of variables, represented by the two axes and the size of the bubble.
- **Line ( ):** This chart is useful for comparing trends.
- **XY (Scatter) ( ):** This is useful for comparing a set of values with the average or predicted values.
- **Pie ( ):** This chart is used to compare a set of figures.
- **Doughnut ( ):** This chart is very similar to a pie chart, except that it can show more than one set of figures. Each ring of the doughnut represents a set of figures.

14.6.1 Creating Charts

One of the simplest and easiest methods to create a chart is by using the *Chart Wizard*. This wizard helps in creating a chart by displaying a series of dialog boxes. The dialog boxes included are selecting a chart type; selecting a format for the chart; specifying how the data are arranged; and adding a legend, axis titles, and a chart title. To insert a chart, follow the steps given below:

1. Enter the data into the worksheet, which is to be converted into a chart as shown in Figure 14.33.

	A	B	C
1	Subjects	Students	
2	English	34	
3	Maths	32	
4	Science	26	
5	History	28	
6	Computers	39	
7			

Figure 14.33 Worksheet Data

2. Select **Chart** from the **Insert** menu or click the **Chart Wizard** button ( ) on the standard toolbar to view the first **Chart Wizard** dialog box.
3. Choose the **Chart type** and **Chart sub-type** from the available list. Then click **Next** (Figure 14.34)

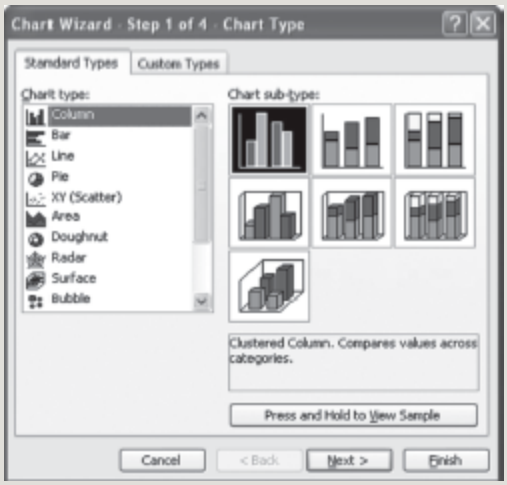


Figure 14.34 Chart Type Dialog Box

4. Select the data range (if different from the area highlighted in step 1) and click **Next** (Figure 14.35).

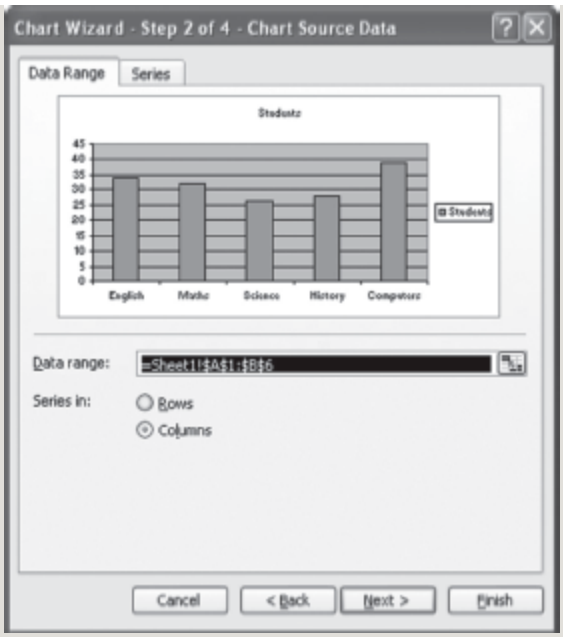


Figure 14.35 Chart Soource Data Dialog Box

5. Now the Chart Options dialog box appears (see Figure 14.36). This dialog box allows you to enter the name of the chart and titles for the X- and Y-axes. By clicking on the tabs, one can change the axes, gridlines, legend, data labels, and data table. Click Next.

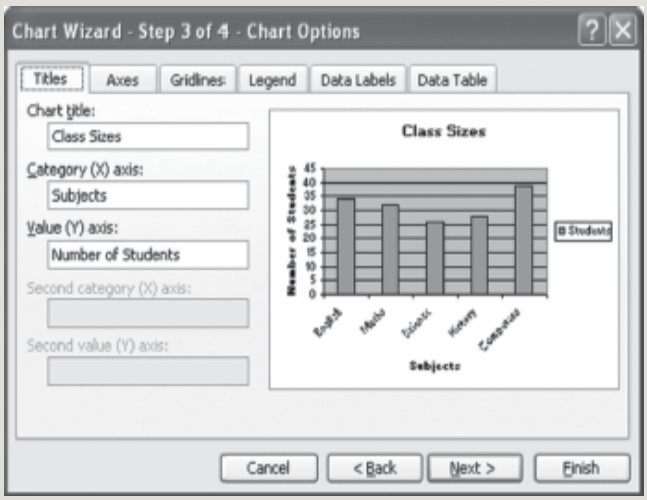


Figure 14.36 Chart Options Dialog Box

6. The Chart location dialog box appears. This is the last step in Chart Wizard. This dialog box prompts you to specify the location of the chart. Click **As new sheet** if the chart is to be placed on a new, blank worksheet or select **As object in** to embed the chart in an existing worksheet as shown in Figure 14.37.

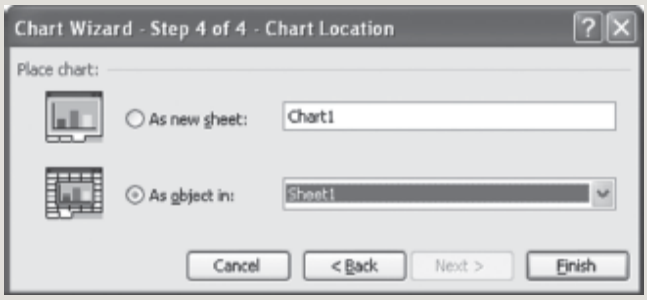


Figure 14.37 Chart Location Dialog Box

7. Click **Finish** to create the chart. The chart appears on the screen as shown in Figure 14.38.

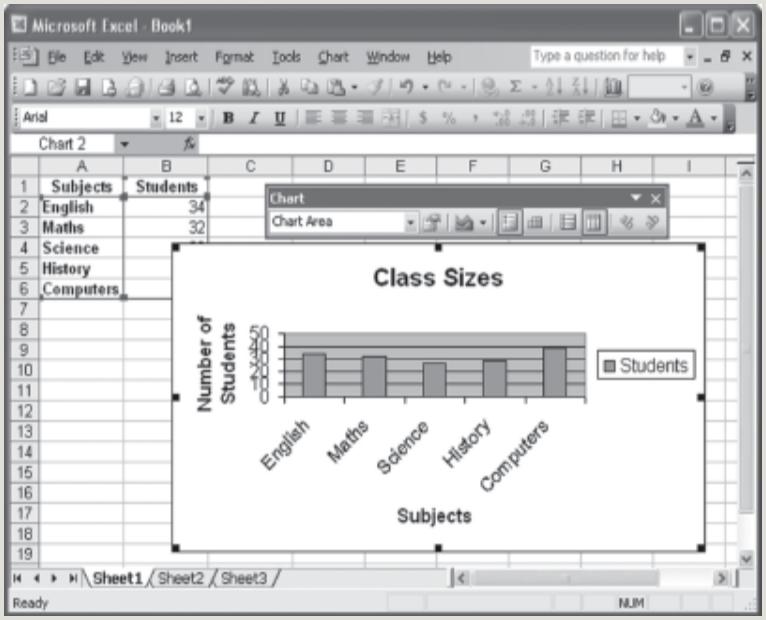


Figure 14.38 Chart on the Worksheet

Resizing and Moving Chart: Chart can be easily resized by clicking on the border of the chart. Doing this, the resizing handles appear on the border of the chart. Handles on the corners resize the chart proportionally while handles along the lines will stretch the chart.

A chart can be moved by selecting the chart, then holding down the left mouse button, and dragging the chart to a new location. Elements within the chart such as the title and labels can also be moved within the chart (Figure 14.39).

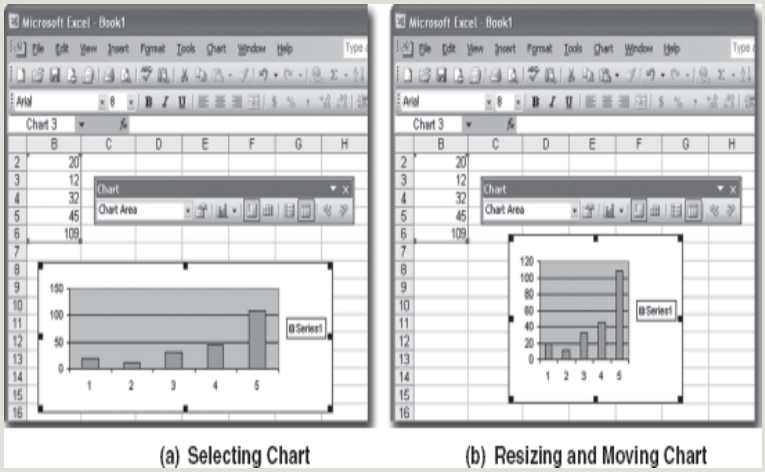


Figure 14.39 Resizing and Moving Chart

14.6.2 Using Chart Toolbar

Once a chart is inserted, it can be changed and enhanced by using the Chart toolbar. If the Chart toolbar does not appear, right-click anywhere on the menu bar and select **Chart** from the shortcut menu (Figure 14.40). The various buttons available on the Chart toolbar are listed in Table 14.12.

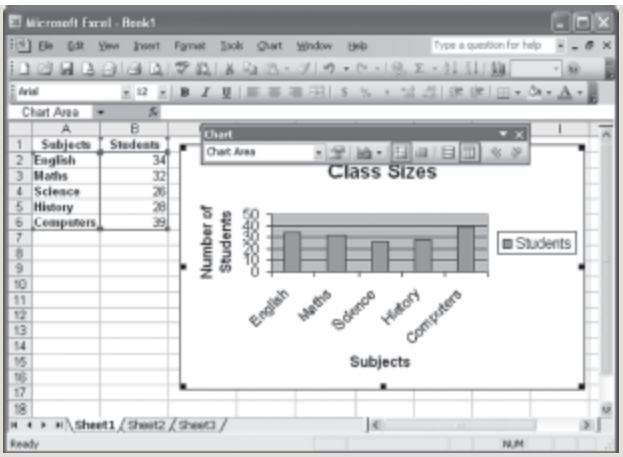


Figure 14.40 Selecting Chart for Editing

Table 14.12 List of All the Available Buttons on the Chart Toolbar

Command	Button	Description
Chart Objects		Used to select different objects in a chart
Format Chart Area		Used to edit the chart area
Chart Type		Used to select a different type of chart
Legend		Used to show or hide the chart legend
Data Table		Displays the data table instead of the chart
By Row		Displays the data by rows
By Column		Displays the data by columns
Angle Clockwise		Used to angle the text in the downward direction
Angle Counterclockwise		Used to angle the text in the upward direction

Changing the Chart Type: If a user after making a particular type of chart feels that the chosen type does not suit or meet the requirements, then a different type of chart can be chosen from the Chart toolbar to change the type of chart (Figure 14.41).

Figure 14.41 Different Chart Types

14.6.3 Saving a Chart

You can save charts for easy reference in future, thus reducing the time involved in making the same chart repeatedly. To save a chart, follow the steps given below:

1. Select the chart and choose **Chart Type** from the **Chart** menu to display the Chart Type dialog box.
2. In the **Custom Types** tab, choose **User-defined** and click **Add.** This displays the Add Custom Chart Type dialog box.
3. Type a name for the graph in the **Name** text box and the description for the graph in **Description** box (if required) and click **OK** to close the Add Custom Chart Type dialog box (Figure 14.42).

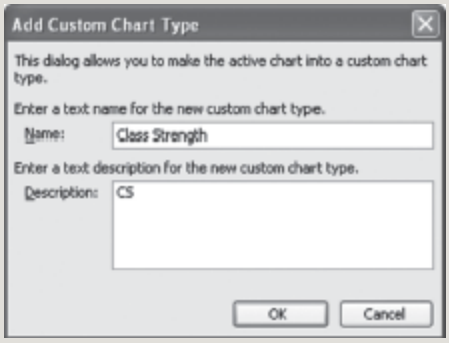


Figure 14.42 Setting a Default Chart

4. Click **Set as default chart** button, a message box for confirming the action appears. Click **Yes** to set the selected chart as the default chart.
5. Click **OK** to save the chart (Figure 14.43).

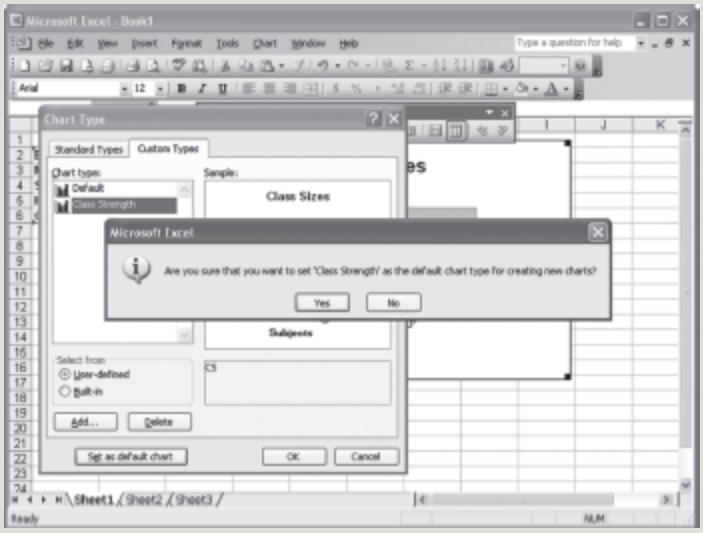


Figure 14.43 Saving a Default Chart

14.7 SORTING

Excel allows you to sort the data in any order. To perform this, follow the steps given below:

1. Click in the column by which you want to sort the data or select a specific range in a column that is to be sorted.
2. Select **Sort** from **Data** menu to display the Sort dialog box (Figure 14.44).

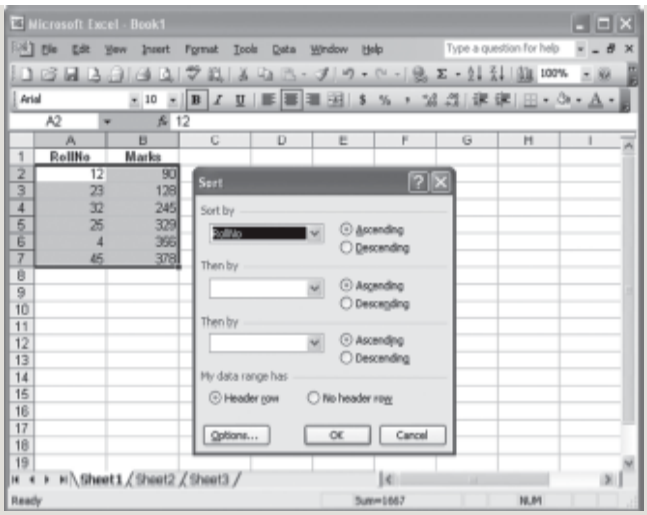


Figure 14.44 Sort Dialog Box

3. By default, Excel sorts all the data in ascending order. If you want to sort in descending order, select **Descending**. Note that you can sort the data up to three fields at a time.
4. Click **OK** to sort data and close the Sort dialog box (Figure 14.45).

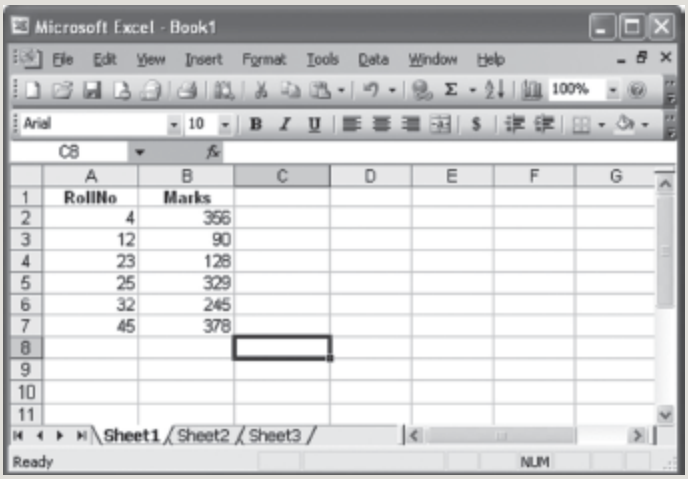


Figure 14.45 Sorted Data

Printing in Microsoft Excel is similar to printing in other Windows-based applications. However, several options, particularly those concerned with arranging the page, are specific to the application.

14.8.1 Setting Page Layout

Page Layout option is used to view the existing page layout or to set a new layout. To set a new layout, select **Page Setup** from the **File** menu to display the Page Setup dialog box. This dialog box has four tabs: *Page, Margins, Header/Footer,* and *Sheet*.

Page Tab: Click on the Page tab to make changes to the page layout. This tab allows a user to make changes to *any* of the following:

- **Orientation:** Select Portrait or Landscape in the Orientation section.
- **Paper size:** Display the **Paper size** drop-down box to select the required paper size.
- **Scaling:** This is used to reduce or enlarge the print.
- **Print quality:** Choose the quality needed. This depends upon the printer installed on the computer.
- **Page numbering:** To begin page numbering, select the **First page number** text box and enter the number that is needed (Figure 14.46).

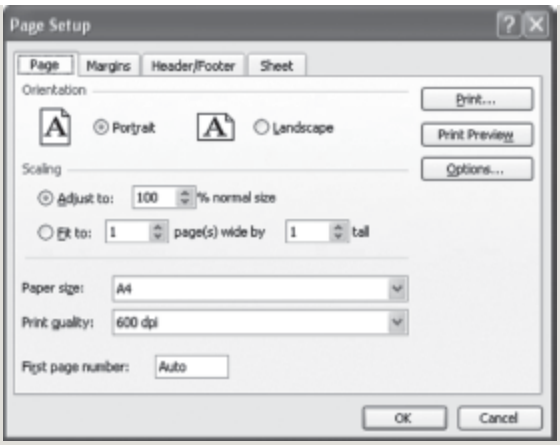


Figure 14.46 Page Setup Dialog Box

Margins Tab: Margins are the empty space between the text and the left, right, top, and bottom edges of a printed page. By default, margins are "1" inch at the top and bottom and ".75" inch from left and right. You can make changes to the top, bottom, left, and right margins of the document in order to create more space for the data present on the page or to add some extra space when binding a document (Figure 14.47).

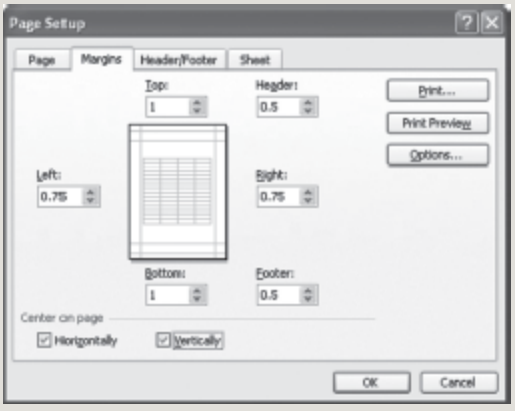


Figure 14.47 Setting Margins

Header/Footer Tab: Headers and footers can be easily created in Excel. There are some standard as well as customized options available for creating headers and footers. You can also create custom headers and footers by selecting Header/Footer tab in the Page Setup dialog box (see Figure 14.48).

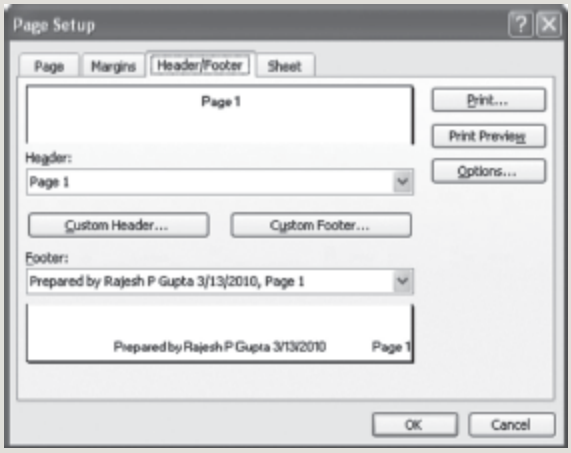


Figure 14.48 Adding Headers and Footers

To create header (or footer), click on the down arrow to the right of the **Header** (or **Footer**) list box to display a list of available headers (or footers). Click on the appropriate header (or footer) needed. If you want to create a different header (or footer), then click on the **Custom Header** (or **Custom Footer**) button to display the Header (or Footer) dialog box (Figure 14.49). In the **Left section** box, enter any data that are needed to be present at the left margin of the header (or footer). In the **Center section** box, enter the data that are required at the centre of the header (or footer). Similarly, in the Right section box, enter the data required at the right margin of the header (or footer). You can also add

date and time to header (or footer) using the **Date** and **Time** icons, respectively. When finished, click **OK** to close the dialog box. The new header (or footer) will be displayed in the Page Setup dialog box in the **Header** (or **Footer**) list box.

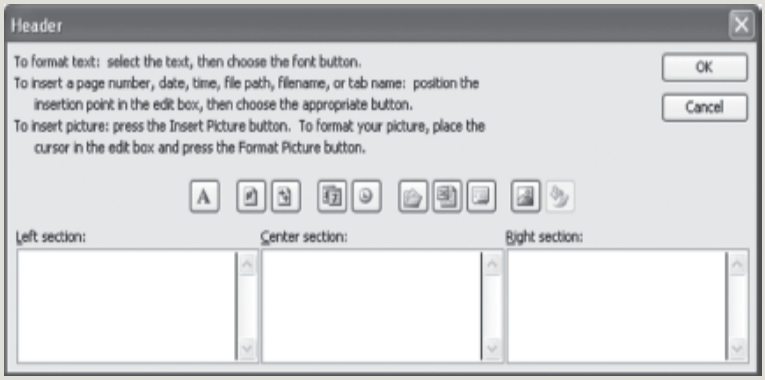


Figure 14.49 Customizing Header

Sheet Tab: You can change the print range and elements to be printed such as headings, gridlines, and comments, and also the order of pages to be printed in this tab. This tab allows a user to specify *any* of the following:

- **Columns or Rows to Repeat**: Click on the icon to the right of the **Rows to repeat at top** text box in the **Print titles** area and drag over the rows that the user wants to repeat at the top of the page. Click on the icon in the right of the **Columns to repeat at left** text box and drag over the columns that are to be repeated at the left of the page.
- **Elements That will Print**: Select the elements to print, that is, Gridlines, Comments, Draft Quality, Black and White, and Row and Column Headings.
- **Order of Pages to Print**: Select Down, then over, or Over, then down.
- **Print Range**: In the **Print area** text box, enter the worksheet range to be printed (Figure 14.50)

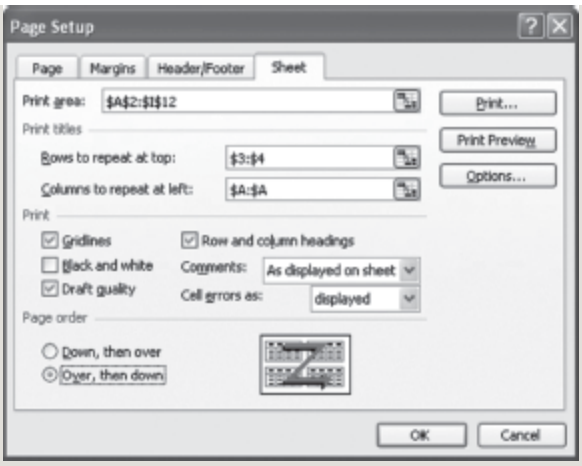


Figure 14.50 Sheet Tab

14.8.2 Print Preview

Print preview is a way to view the printed worksheet on-screen before printing the final output. Previewing the worksheet is a good way to identify formatting errors, such as incorrect margins, overlapped text, boldfaced text, and other text enhancements. This helps in saving costly printer paper, ink, and time. To view a worksheet in print preview mode, choose **Print Preview** from the **File** menu or click the **Print Preview** button ( ) on the Standard toolbar. Note that the options available in print preview mode in Excel are similar to Word. Click **Close** to return to the worksheet or **Print** to continue printing (Figure 14.51).

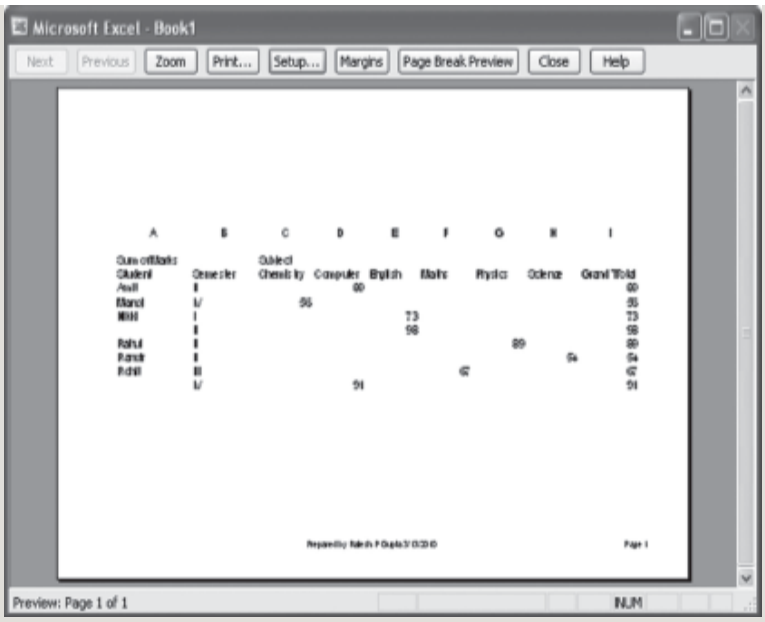


Figure 14.51 Print Preview

14.8.3 Printing Worksheets

Once you have completed formatting and editing worksheet(s), you can print the whole sheet (Active sheet), part of the sheet (selection), or several sheets (workbook). To print, click the **Print** button ( ) on the Standard toolbar, or select **Print** from the **File** menu to display the Print dialog box. This dialog box allows you to select the number of copies and how many pages of the document you want to print. It also contains **Print what** section where a user can choose the printing area. Some of the options provided under this section are:

1. **Selection**: Choose this option if only a selected part of the worksheet is to be printed.
2. **Entire workbook**: Choose this option if the entire workbook is to be printed.
3. **Active sheet:** Choose this option if the Active worksheet (active sheet is the sheet where user is working) is to be printed.

After selecting the appropriate option, click **OK** to print (Figure 14.52).

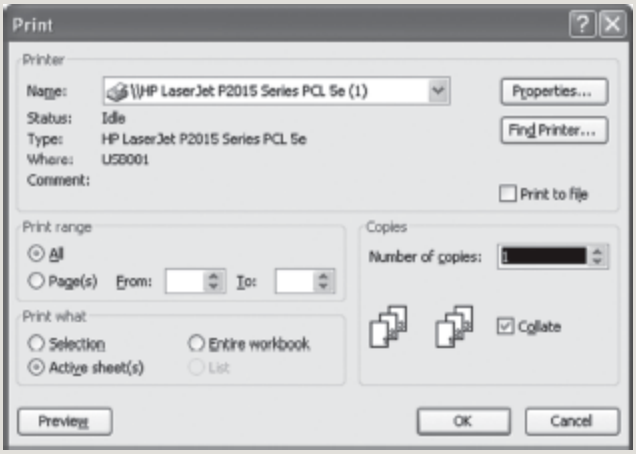


Figure 14.52 Print Dialog Box

Important Keyboard Shortcuts

Document Actions

Action	Shortcut Key
Open a Workbook	CTRL+O
Create a New Workbook	CTRL+N
Close a Workbook	CTRL+F4
Close a Workbook	CTRL+F4
Save As	F12
Save	CTRL+S
Help	F1
Print	CTRL+P
Close Excel Application	ALT+F4

Text-Formatting Actions

Action	Shortcut Key
Bold	CTRL+B
Italics	CTRL+I
Underline	CTRL+U
Edit Active Cell	F2
Format Cells Dialog Box	CTRL+1

Selecting Cells

Action	Shortcut Key
Entire Column	CTRL+Spacebar
Entire Row	SHIFT+Spacebar
Cells Left of Active Cell	SHIFT+Left Arrow
Cells Right of Active Cell	SHIFT+Right Arrow
Entire Worksheet	CTRL+A

Text-Editing Actions

Action	Shortcut key
Undo and Redo	CTRL+Z, CTRL+Y
Cut	CTRL+X
Copy	CTRL+C
Paste	CTRL+V
Find	CTRL+F
Replace	CTRL+H
Go To	CTRL+G

Miscellaneous

Action	Shortcut key
Move to Start of Row	Home
Move to Start of Column	CTRL+Left Arrow
Move to Next Worksheet	CTRL+Page Down
Apply AutoSum	ALT+=
Current Date	CTRL+;
Current Time	CTRL+:
Spelling	F7

LET US SUMMARIZE

1. Microsoft Excel is a spreadsheet program that allows you to perform various calculations, estimations, and formulations with data. It is the electronic counterpart of a paper ledger sheet, which consists of grid of columns and rows.
2. To open Microsoft Excel, click Start, point to All Programs, then point to Microsoft Office, and then select Microsoft Office Excel 2003.
3. To create a new workbook, click on New button ( ) on the Standard toolbar. Alternatively, select New from the File menu.
4. To open an existing workbook, click on the Open button ( ) on the Standard toolbar. Alternatively, select Open from the File menu.
5. To save the workbook, Microsoft Excel provides two menu options, namely Save and Save As.
6. To close the workbook, select Close from the File menu.
7. To find and replace text within the worksheet use the Find and Replace dialog box, which can be displayed by selecting Find and Replace from the Edit menu.
8. To undo the last action, select Undo from the Edit menu. To redo the last undo action, select Redo from the Edit menu. You can also use the Undo button ( ) and Redo button ( ) from the Standard toolbar.
9. In Excel, numerical data can be easily converted into a chart for graphical presentation of the data. Charts provide more visual clarity than tables of data and, therefore, have more impact.
10. One of the simplest and easiest methods to create a chart is by using the Chart Wizard. This wizard helps in creating a chart by displaying a series of dialog boxes.
11. *Print preview* is a way to view the printed workbook on-screen before printing the final output. To view a worksheet in the print preview mode, choose Print Preview from File menu or click the Print Preview button ( ) on the Standard toolbar.
12. To print a worksheet, click the Print button ( ) on the Standard toolbar. Alternatively, use the Print dialog box, which can be displayed by selecting Print from the File menu.

Fill in the Blanks

1. ..... are popular because they represent a better alternative to manually computing mathematical calculations and are more accurate and time saving.
2. By default, ... and ..... toolbars are displayed in the Excel environment.
3. Microsoft Excel worksheet contains columns ..... rows and ..... columns.
4. Press ..... or ..... the cell to modify the contents of the cell.
5. The options present in ..... tab allow you to change the font attributes such as font face, size, style, and effects if desired by the user.
6. The ..... provides an easy method of copying the contents of the selected cells to the adjacent cells in a column or in a row.
7. ..... are mathematical expressions built in Excel that instruct the computer to carry out calculations on specified sets of numbers in the rows and columns.
8. ..... reverses the value of its argument.
9. ..... chart shows three sets of variables, represented by the two axes and the size of the bubble.
10. A user can also create custom headers and footers by selecting ..... tab in the Page Setup dialog box.

Multiple-choice Questions

1. Which of these is not a font style?
    1. Bold
    2. Underline
    3. Italic
    4. Footer
2. A cell range is represented as follows:
    1. (A6.A8)
    2. (A6–A8)
    3. (A6:A8)
    4. (A6,A8)
3. ..... appear at the bottom of the Excel window.
    1. Sheet tabs
    2. Name Box
    3. Formula bar
    4. Title bar
4. Workbook is a collection of .....
    1. Page Setup
    2. Buttons
    3. Worksheets
    4. Charts
5. Borders tab in Format Cell dialog box is used to apply .....
    1. Borders
    2. Shading
    3. Background colours
    4. All of these

6. Excel provides a feature known as ..... that enables you to apply pre-defined layouts to selected tables in the worksheet.
    1. AutoFormat
    2. Header and Footers
    3. Undo and Redo
    4. None of these
7. ..... and ..... option in Microsoft Excel is used to locate a particular word, phrase, or numeric values in a worksheet, or replace it with something else.
    1. Undo, Redo
    2. Formulas, Functions
    3. Find, replace
    4. None of these
8. One of the distinguishing features of Excel is that it makes use of ..... to dynamically calculate results from the data present in worksheets.
    1. Formulas and functions
    2. Goto
    3. Tables
    4. All of these
9. Microsoft Excel contains many pre-defined or built-in formulas, which are known as .....
    1. Formatting Cells
    2. Functions
    3. Relative Referencing
    4. AutoSum
10. An error may occur in MS Excel while working with ..... when the intended task is not carried out in a proper way.
    1. Formulas
    2. Tables
    3. Headers and Footers
    4. None of these

State True or False

1. The Name Box displays the name of the active cell or selected range and can be used to name a range of cells.
2. The intersection of row 12 and column E gives cell 12E.
3. In Excel, the rows are labelled in a numbered order (1, 2, 3, ..., 65536).
4. The Save option is used to make the multiple copies of the same file.
5. To select a large area of adjacent cells, the Ctrl key is used to extend the selection.
6. The cell contents can also be moved to another cell by simply dragging the cell to point to the desired cell.
7. Excel allows the users to create mathematical formulas and execute functions.
8. In the Bar chart, each category of information has its own line radiating out from the centre.
9. Data entries are categorized into three basic types: labels, values, and formulas.
10. By default, Excel sorts all the data in descending order.

Descriptive Questions

1. Describe the steps involved in creating a chart.
2. Show the steps involved in naming a range of cells.
3. Explain Number, Font, and Alignment tab in Format Cell dialog box.
4. Explain Find and Replace in Excel 2003.
5. What is the difference between Formulas and Functions? Show the steps involved in inserting Functions.
6. Explain the AutoSum feature with the help of an example.
7. Describe the various types of error values that might occur while working with Functions and Formulas.
8. What are the various types of charts available in Excel 2003?
9. Explain the various options available on the Charts toolbar.
10. Explain the Margin tab, Sheet tab, and Header/Footer tab in the Page Setup dialog box.

ANSWERS

Fill in the Blanks

1. Spreadsheets
2. Standard, Formatting
3. 65,536, \256
4. F2, double-click
5. Font
6. Fill handle
7. Formulas
8. NOT(logical)
9. Bubble
10. Header/Footer

Multiple-choice Questions

1. (d)
2. (c)
3. (a)
4. (c)
5. (a)
6. (a)
7. (c)
8. (a)
9. (b)
10. (a)

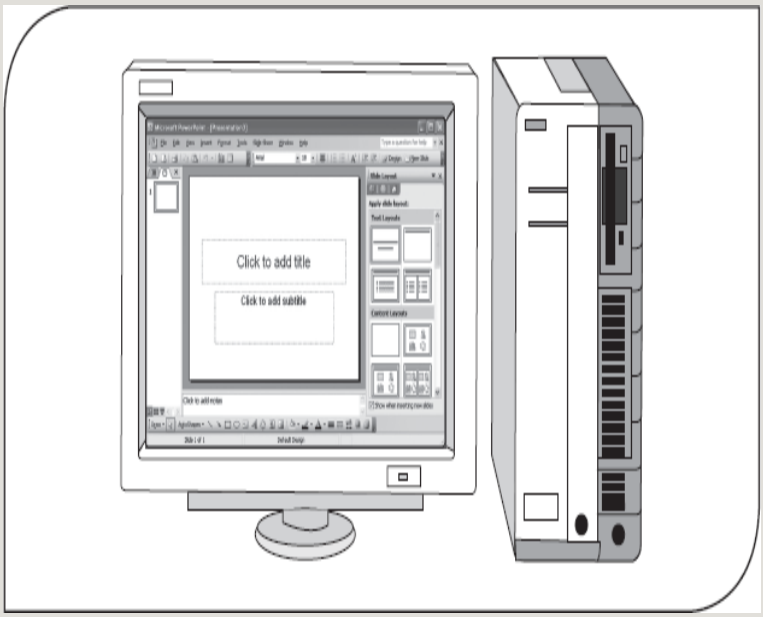
State True or False

1. True
2. False
3. True
4. False
5. False
6. True
7. True

8. False
9. True
10. False

Chapter 15

Microsoft Office PowerPoint 2003



This chapter is penned to provide a foundation to harness the power of creating presentations using Microsoft Office PowerPoint 2003. PowerPoint enables the user to effectively create presentations quickly and efficiently. The reader will learn to create and format slides using various in-built features of PowerPoint. The chapter also covers various tools and features such as animation effects, and transition effects, through which a presentation can be greatly enhanced. The chapter concludes by discussing how to pack and print a PowerPoint presentation.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- How a good presentation communicates effectively and efficiently to convince, motivate, inspire, and educate the audience
- The steps required for creating a PowerPoint presentation from scratch using AutoContent Wizard, Design Template, and so on
- The different types of screen layouts and views provided by PowerPoint
- How PowerPoint presentation is created with the help of Design Templates and Masters for consistent appearance of the slides
- How PowerPoint enables AutoLayouts to create slides using placeholders for text and WordArt and Clip Art to insert graphics for visual impact and emphasis on the presentation
- How multimedia effects can be added to make the presentation more interactive and interesting
- The different options of printing documents offered by PowerPoint

15.1 INTRODUCTION

In today's highly competitive world, representing ideas effectively has become the need of the hour. No matter how good the product or for that matter a project is, if the idea of the project or the product is not communicated effectively and efficiently, failure is guaranteed. One is also better heard if the ideas are presented stylistically. One of the best ways of expressing the ideas is visual images, which includes usage of graphics and pictures. Graphics, when used to convey ideas, expression, or thoughts are known as *presentation graphics or business graphics*. A good presentation can truly convince, motivate, inspire, and educate its audience.

THINGS TO REMEMBER

Features of PowerPoint 2003

1. PowerPoint 2003 combines slides, outlines, and notes into one view which makes it easy to add slides, text, notes, and navigate.
2. In PowerPoint, text is automatically resized to fit into a place holder. Slides automatically resize to fit the display resolution and window.
3. You can create tables directly in PowerPoint 2003 instead of importing from Word or Excel.
4. You can create mini-presentations within a single file that are tailored to different audiences.
5. The new Slide Show menu consolidates everything you need to deliver presentations electronically.
6. PowerPoint 2003 provides with every output options: slides, colour and black-and-white, speaker's notes, and on-screen electronic presentations.

Microsoft Office PowerPoint 2003 is a powerful, yet easy-to-use presentation package which is a part of Microsoft Office 2003 suite of products. Its primary purpose is to help you design professional-style presentations quickly and easily. PowerPoint offers the tools to create a basic presentation and to enhance and customize it to meet its goal. The tri-pane feature of PowerPoint 2003 allows you to view the slides, outline, and notes simultaneously,so that you can easily organize the information to present in the slides.

15.2 STARTING MICROSOFT OFFICE POWERPOINT 2003

To open Microsoft PowerPoint, perform *any* one of the following steps:

- Double-click the **Microsoft PowerPoint** icon located on the desktop.

- Click **start**, point to **All Programs**, then point to **Microsoft Office**, and then select **Microsoft Office PowerPoint 2003** (Figure 15.1).

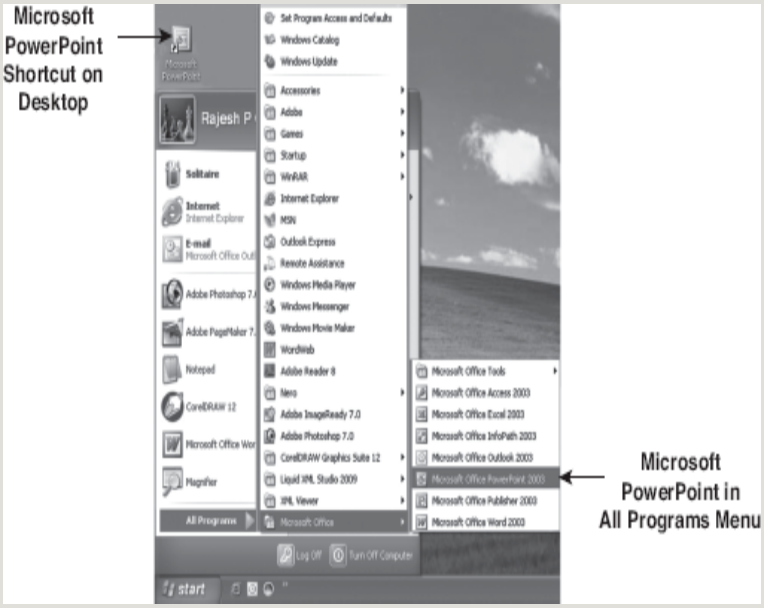


Figure 15.1 Starting Microsoft PowerPoint

15.2.1 Microsoft PowerPoint Environment

When Microsoft PowerPoint is opened, the main screen of the program appears. This main window contains many parts; these parts are described in detail below (Figure 15.2).

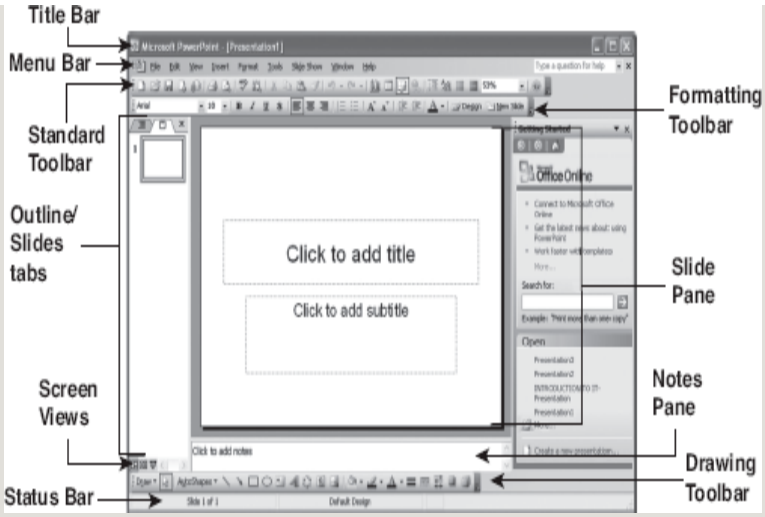


Figure 15.2 Main Microsoft PowerPoint Window

The **title bar** is located at the very top of the PowerPoint window and it displays the name of the application and active presentation. Below this bar is the **menu bar**, which contains different menus that control all options, functions, and commands for the entire Microsoft PowerPoint application. By default, Microsoft Office PowerPoint 2003 contains nine menus, each of which has an associated pull-down menu. For example, the **File** menu contains commands to open, create, and print a presentation. Generally, a **toolbar** is positioned just below the menu bar. Toolbar contains command buttons that provide single-click entry to carry out many tasks. By default, Standard and Formatting toolbars are displayed in the PowerPoint environment. Additional toolbars like Drawing and Picture toolbars can be added by right-clicking on the menu bar and selecting the desired toolbar(s) from the pop-up menu, as shown in Figure 15.3.

Figure 15.3 Various Toolbars

Outline/Slides Tabs: Below the toolbars, on the left side Outline/Slides tabs are located. They are used to organize and develop content for the presentation, move slides from one position to another, and edit titles and body text.

Screen Views: PowerPoint provides a range of screen views for creating a presentation. The different screen views available are *Normal, Slide Sorter,* and *Slide Show*. These views give you the opportunity to modify and develop your presentation as desired. You can select the desired view by clicking the button at the bottom of the page (Figure 15.4).

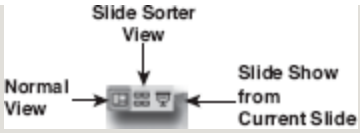


Figure 15.4 Different view Buttons

Slide Pane: In this pane, you can see the way the slide appears, with text, design elements, and graphics. You can also add or edit text, hyperlinks, graphics, movies, sounds, and animations to individual slides in this pane.

Notes Pane: Below the Slide pane is the Notes pane. This pane lets you add **Speaker Notes** to your presentation. These notes are kept by speaker for reference. These are typically a smaller version of the slides along with some notes.

15.3 WORKING WITH POWERPOINT 2003

In this section, you will learn the steps required to produce a PowerPoint presentation from scratch. These include:

- Creating a new presentation
- Opening an existing presentation
- Saving a presentation
- Closing a presentation

15.3.1 Creating a New Presentation

Whenever you start PowerPoint, it opens a new blank presentation and the Getting Started task pane, which enables to access online help and content, open existing presentations and create a new presentation. To create a new presentation, click **Create a new presentation** on the Getting Started task pane, to display New Presentation task pane (see Figure 15.5). A new

presentation can be created using the **AutoContent wizard, Design Template,** and **Blank presentation.**

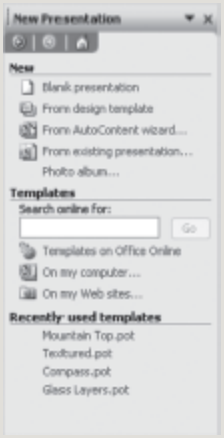


Figure 15.5 Creating a New Presentation

Using the AutoContent Wizard: The AutoContent Wizard helps you to create and structure the content of presentation in a clear and logical way. It helps in creating a presentation by leading you through some basic steps. The wizard then uses answers to automatically layout and format the presentation. PowerPoint 2003 then selects the best style and built-in outline to suit the presentation. This wizard helps to create a presentation quickly and easily and is very helpful for new users. To create a presentation using this wizard, follow the steps given below:

1. Select **From AutoContent wizard** from the New Presentation task pane to display the AutoContent Wizard dialog box. Click **Next** to continue to the next step (Figure 15.6).

Figure 15.6 Step 1: of AutoContent Wizar

2. The Wizard displays various categories of presentations, such as **All, General, Corporate, Projects,** and **Sales/Marketing.** The box adjacent to these categories displays the available presentations for the category selected. Select the desired presentation and click **Next** (Figure 15.7).

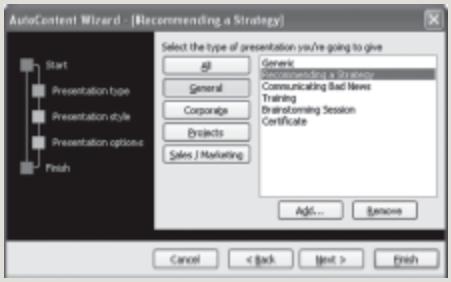


Figure 15.7 Step 2: of AutoContent Wizard

3. The next decision is to select the output type. You can present the slide show on an overhead projector, computer screen, and so on. Select the desired output type and click **Next** (Figure 15.8).

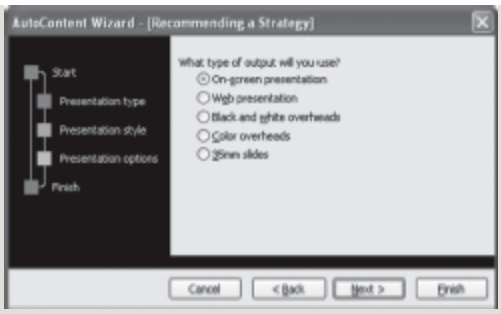


Figure 15.8 step 3: of AutoContent Wizard

4. Enter the presentation title, which you want to display in each slide inside the **Presentation title** box. You can also include information such as the company name or the name of your department in the **Footer** box. There are two checkboxes as well. One includes a date on which the presentation was last updated and the other adds the slide number. Select the desired option and click **Next** (Figure 15.9).

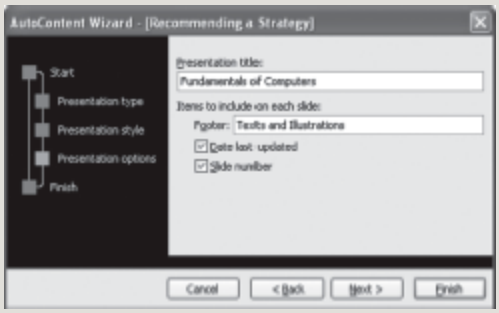


Figure 15.9 Step 4: of AutoContent Wizard

5. Click **Finish** to create the presentation (Figure 15.10).

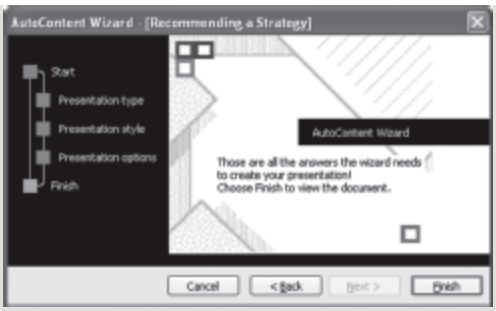


Figure 15.10 Step 5: of AutoContent Wizard

6. A presentation is created as shown in Figure 15.11.

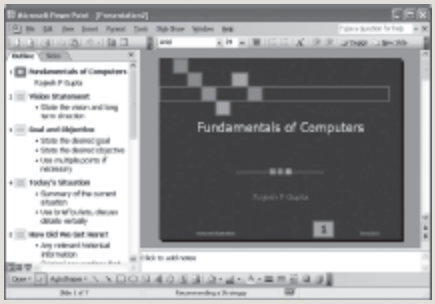


Figure 15.11 A Sample Presentation

Using Design Template: PowerPoint provides many built-in slide sets (templates) with different backgrounds and text formatting to create a presentation. A design template includes pre-formatted layouts, fonts, and colours, which blend together to give a consistent look to the presentation.

If you choose **Design Template**, a new presentation is displayed using the design selected by the user,consisting of a single slide. This option allows you to create a presentation, maintaining a consistent design scheme

like layouts, colours, fonts, and so on. To use the Design Template, follow the steps given below:

1. Select **From design template** from the New Presentation task pane to display Slide Design task pane, which shows various design templates (Figure 15.12).

Figure 15.12 Design Templates

2. Choose the template for the slide from **Apply a design template** section. Note that Microsoft PowerPoint 2003 contains readymade design templates each having its own layout, font styles, and font colours, and so on. Figure 15.13 illustrates the slide with the selected design template.

Figure 15.13 A Sample Slide

Using Blank Presentation: The most basic method of creating a slide show presentation is by using Blank Presentation. This method is used to start the whole process of creating presentation from scratch. In other words, it includes no preset design, colour, or content suggestion. To use the Blank Presentation, follow the steps given below:

1. Select **Blank presentation** from the New Presentation task pane to display a new blank slide and a Slide Layout task pane (see Figure 15.14).

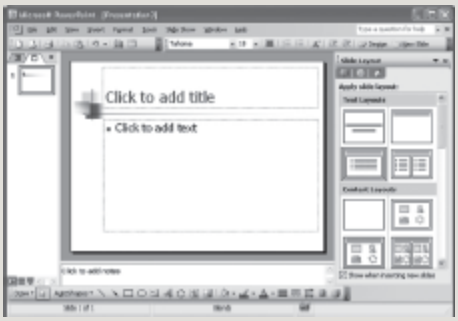


Figure 15.14 A Blank Presentation

2. Choose the layout of the slide. Note that PowerPoint 2003 includes 27 different types of layouts, known as **AutoLayouts.** Figure 15.15 illustrates the slide with selected layout.

Figure 15.15 A Blank Presentation

15.3.2 Opening a Presentation

To open an existing presentation, follow any of the steps given below:

- Select **Open** from the **File** menu to display the Open dialog box. You can also open a presentation by clicking on the **Open** button ( ) on the Standard toolbar. The Open dialog box allows you to specify the name of the desired file in the **File name** list box. If the desired file is not in the current location, you can locate it by navigating through **Look in** list box. Once the file is found, select the file and click the **Open** button.
- PowerPoint presentation can also be opened by double-clicking the presentation file icon, if it exists (Figure 15.16).

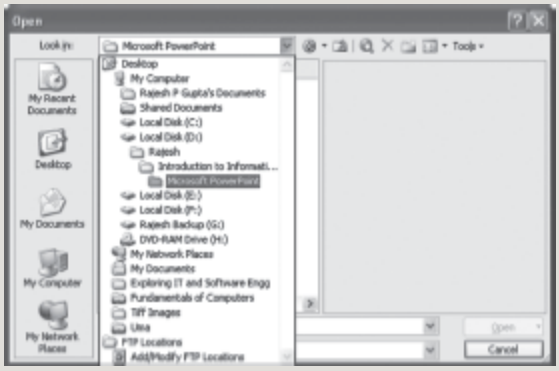


Figure 15.16 Open Dialog Box

15.3.3 Saving a Presentation

Once a presentation is created, you can start typing the text, inserting graphics, animations, charts, and so on. When finished, you must save the presentation for future references. To save the presentation, Microsoft PowerPoint provides two menu options, namely **Save** and **Save As**. If you are working on an untitled presentation which has never been saved, then using any of the options (**Save and Save As**) will have the same effect, that is, it displays a Save As dialog box using which you can save the PowerPoint file at the desired location (Figure 15.17).

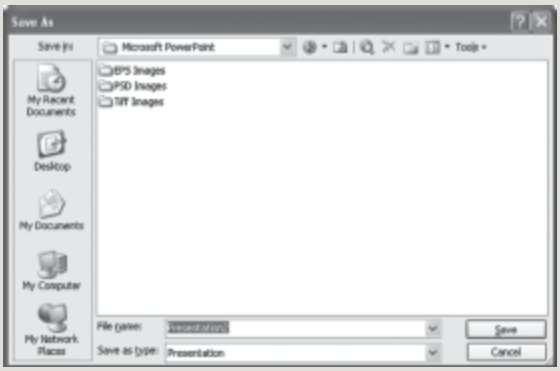


Figure 15.17 Save As Dialog Box

Once you have saved the new presentation, you can use the **Save** option to update changes made to a file, while the **Save As** option can be used to make the multiple copies of the same file.

Note: *By default, Microsoft PowerPoint presentation file is saved with .PPT extension. Although it can also be saved in other file formats like .PPS, .BMP, .GIF, and .HTML.*

Setting the Default Directory and AutoSave Property: PowerPoint 2003 allows to change the default location for saving the files as well as to save your files automatically. The files are saved automatically by enabling the *AutoRecover* feature. This feature helps in protecting your work even after power failure or when you have forgotten to save your file. To change the default location of your file and save your file using AutoRecover feature, follow the steps given below:

1. Select **Options** from the **Tools** menu to display the Options dialog box (Figure 15.18).

Figure 15.18 Options Dialog Box

2. Select **Save** tab and specify the location of default directory in the **Default file location** text box.
3. Click on the **Save AutoRecover info** checkbox and specify the time interval after which the file will be automatically saved. Click **OK.**

15.3.4 Closing a Presentation

After finishing all the work in PowerPoint, you may want to close the presentation. For this, follow *any* of the steps given below:

- Select **Close** from **File** menu.

- Click on the **Close Window** button at the upper-right corner of the PowerPoint window to close the current presentation. Note that there are two close buttons. One is located at the top of the application window, which closes the PowerPoint application, while the second (lower) one is present on the presentation which closes only the current file (Figure 15.19).

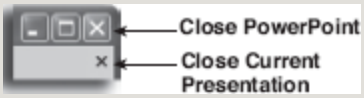


Figure 15.19 Closing a Presentation

Note: *If a user makes changes to a file and has not saved those changes, PowerPoint will ask whether to save the changes before closing the file.*

15.4 WORKING WITH DIFFERENT VIEWS

PowerPoint provides different types of screen layouts for constructing presentation in addition to the Slide Show. View is a way in which one can see and work with the presentation and slides. User can select the page view by clicking the screen view buttons on the left-hand side at the bottom of the window.

The three PowerPoint views are: *Normal View, Slide Sorter View, and Slide Show View.*

Normal View: The default view in PowerPoint 2003 is the Normal View, which makes the most commonly used components of a presentation easily accessible to the user. This view contains three working areas: slide and outline tabs on the left, the slide pane on the right, and notes pane at the bottom.

Normal view enables to see various aspects of the presentation simultaneously in two different tabs namely, *Outline tab* and *Slides tab*.

- **Outline Tab:** This tab enables a user to write and edit the content of presentation. When the text is selected in the outline tab, the buttons for working with outline become available on the Standard and

Formatting toolbars. One can use these buttons to organize the presentation like increasing and decreasing the text indents, expanding and reducing the content, etc (Figure 15.20).

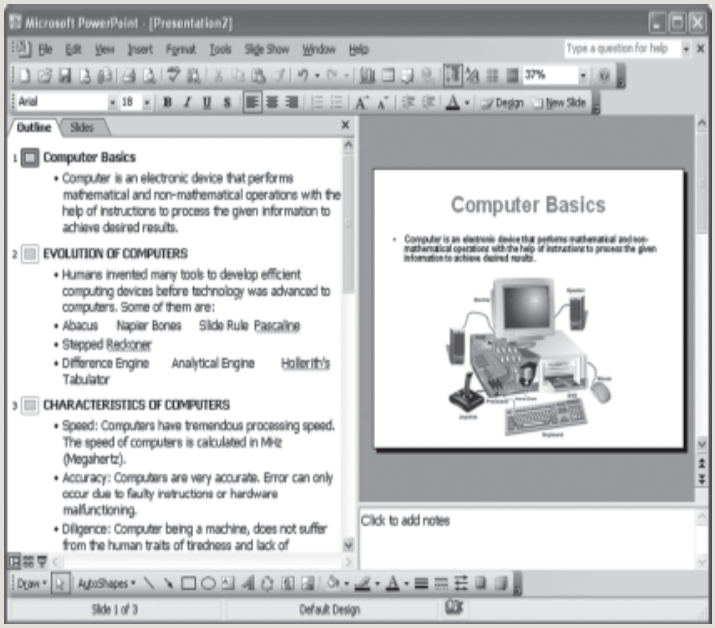


Figure 15.20 Outline Tab in Normal View

- **Slides Tab**: This tab displays the slides as thumbnail-sized images. This tab is helpful for navigating through presentation, seeing the effects of design changes, adding, deleting, or rearranging slides (Figure 15.21).

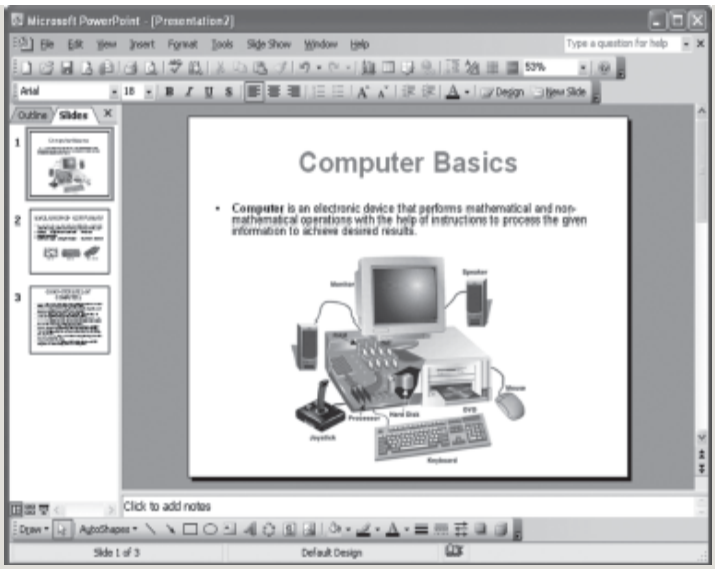


Figure 15.21 Slides Tab

Slide Sorter View: This view option displays all the slides in a presentation as thumbnails. In simple terms, all the slides in the presentation are displayed on the screen at the same time in a miniature form. This view helps user to add, delete, and rearrange the slides. Also, transition effects can be added for moving from one slide to another in a convenient manner. Thus, this view acts like a preview tool (Figure 15.22).

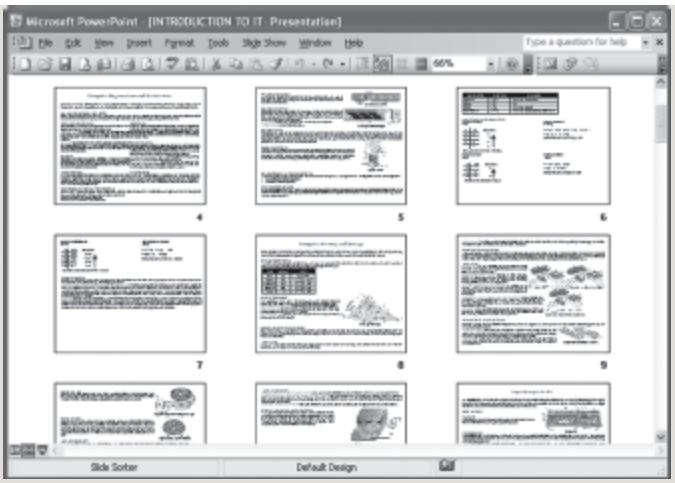


Figure 15.22 Slide Sorter View

Slide Show View: Generally, this view is used when you want to deliver a presentation. This view option allows you to view the slides (full screen), as they would appear in a slide show (Figure 15.23).

Figure 15.23 Slide Show View

15.5 DESIGNING PRESENTATION

A presentation is created using the slides, which contains text, graphics, animations,and so on. The slides are automatically created when a presentation is designed through the AutoContent Wizard or Blank Presentation. Slides can also be inserted in a presentation, if required.

THINGS TO REMEMBER

Template

A template is a presentation containing Power-Point masters with custom formatting, fonts, a colour scheme, placeholder for text, graphics, animations, and so on. PowerPoint offers various pre-designed templates to helpyou develop powerful yet simple presentations. With custom template, presentations will have a consistent look and feel, and you won't have to customize each presentation layout separately.

PowerPoint is designed in a way to give a consistent appearance to your slides. There are two ways in PowerPoint, which help you to choose the look of your slides: *design template* and *masters*.

15.5.1 Choosing Design Template

Design templates contain colour schemes, slide and notes masters with custom formatting, and fonts, all designed to create a desired look for your presentation. When you apply a design template to your presentation, the slide master and colour scheme of the new template replaces the slide master and colour scheme of the original presentation. After you apply a design template, each slide that you add has the same custom look. PowerPoint comes with a wide variety of professionally

designed templates. In addition, you can create your own templates. If you create a special look for one of your presentations, you can save it as a template. To apply a different design to a presentation, open the presentation and follow the steps given below:

1. Select **Slide Design** from the **Format** menu or click **Design** button on the Formatting toolbar, to display the Slide Design task pane (Figure 15.24)

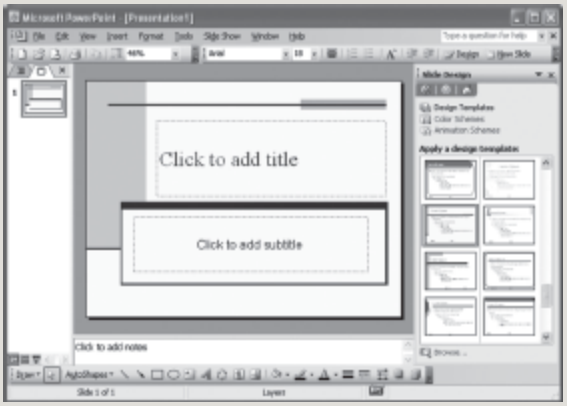


Figure 15.24 Apply Design Template Dialog Box

2. Click the design template that best suits your presentation. Note that the selected design template can be applied to all slides or the **selected** slides in a presentation.

15.5.2 Using Masters

Masters in PowerPoint are used to control many facets of the slides such as backgrounds, font typeface,font size, colours, bullets and locations for all main components, tabs, and indents. PowerPoint has three masters, namely, *Slide Master, Notes Master, and Handout Master*.

Slide Master: PowerPoint automatically assigns a slide master to every presentation. The slide master allows you to customize the look of each slide and also ensures consistency across your presentation. In the

slide master, you can change fonts, bullets, and header and footer information. Use the slide master to add a picture, change the background, adjust the size of the placeholders, and change the font style, size, and colour (Figure 15.25).

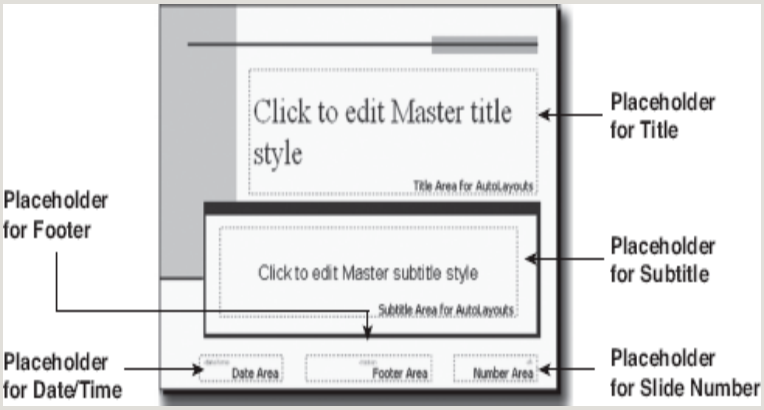


Figure 15.25 Slide Master

To change the formatting in the slide master, select **Master** from the **View** menu and choose **Slide Master.**

Click the appropriate placeholder to edit Master text styles and apply formatting from the Formatting toolbar. For example, if you want all first-level bullet styles in the presentation to appear in 28-point Tahoma, click the text **Click to edit Master text style** and set the font to Tahoma, 28-point. Using the Format AutoShape dialog box, you can draw a thick line around the box, change the line colour, fill colour in the box, set the line thickness and style, position the box in the slide, etc. To

display the Format AutoShape dialog box, select the placeholder and double-click it (Figure 15.26).

Figure 15.26 Using Format Text Box

Notes Master: PowerPoint helps to create notes that will stay with the presentation. This feature provides an area for the speaker notes, along with a reduced picture of the slide. To view the notes master, select **Master** from the **View** menu and choose **Notes Master.** In Notes Master, you can add bullets and page numbers, resize slides or the text placeholders on slides, or change fonts. You can also modify the style of the text that will appear in the printouts of your notes (Figure 15.27).

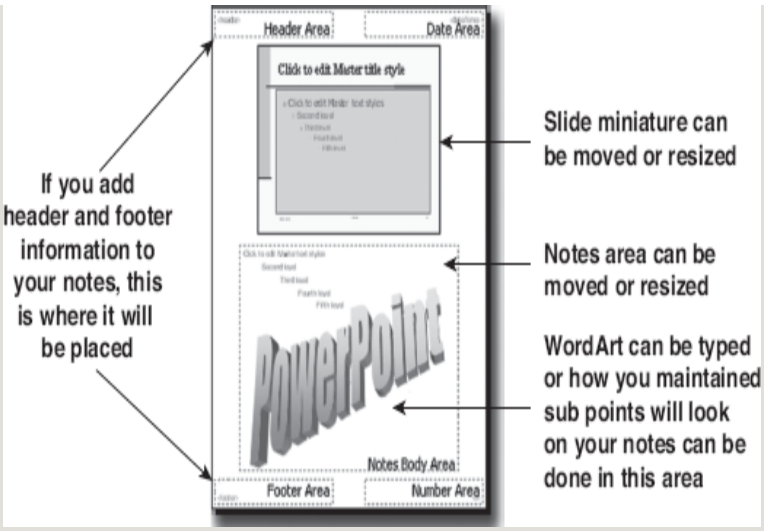


Figure 15.27 Notes Master

Handout Master: In handout master, you can add images, text, and colour to your handouts. In addition, you can add header and footer to the slides and also view these slides in different ways, such as one, two, three, four, six, nine, the outline, and the layout view (see Figure 15.28). To view the handout master, select **Master** from the **View** menu and choose **Handout Master.** The information added in the handout master is not visible while working with the presentation, but will be available when it is printed (Figure 15.28).

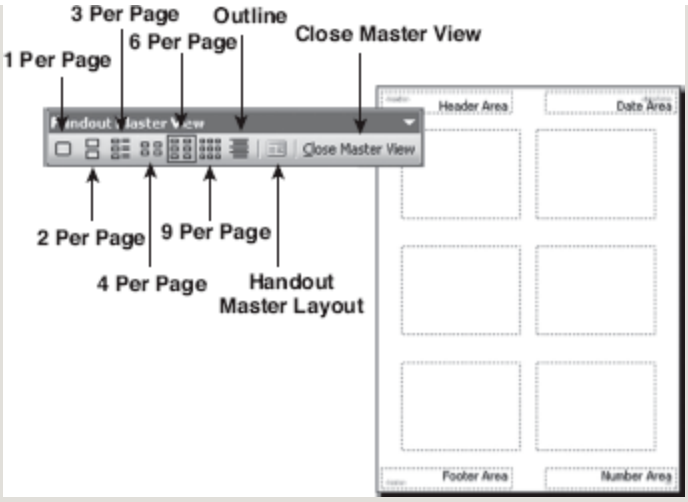


Figure 15.28 Handout Master

15.5.3 Working with Slides

In the previous section, we discussed how to create and save a PowerPoint presentation. Now let us learn how to insert, copy, delete, and navigate between slides. Apart from these activities, you will also learn how to work with text, format text, perform undo and redo operations, make lists, and insert headers and footers.

Inserting and Deleting Slides: Sometimes you may want to add new slides to your presentation. Slides can be added in all views except the Slide Show view. To insert slide, select **New Slide** from the **Insert** menu, or click the **New Slide** button on the Formatting toolbar. This displays the Slide Layout task pane from which you can choose the desired layout.

You can also remove a slide from the presentation. To delete slide, select the slide and follow *any* of the steps given below:

- Select **Delete Slide** from the **Edit** menu. To delete multiple slides, choose Slide Sorter view, hold down the **Ctrl** key to select multiple slides, and then click **Delete Slide**

- Right-click on the slide and perform the cut option. This option works only in case of Slide Sorter view.

Duplicating Slides: PowerPoint 2003 allows you to easily make copies of one slide. This may be required if you want to use the same data in more than one slide. Note that when duplication is done, the layout is also copied. To duplicate slide, follow the steps given below:

1. Select the slide or slides that you want to duplicate.
2. Select **Duplicate Slide** from the **Insert** menu. To duplicate multiple slides, click on one slide and press **Shift** (for selecting contiguous slides) or **Ctrl** (for selecting non-continuous slides) to select multiple slides.

Navigating between Slides: In PowerPoint, you can navigate to the next and previous slides by following any of the three view options: *Normal, Slide Sorter, and Slide Show*. To navigate to the next slide in the Normal view, follow any of the steps given below:

- Click the **Next Slide** icon ( ) on the vertical scrollbar.
- Click the next slide number.

To view the next slide in Slide Show view, right-click on the slide and select **Next** from the pop-up menu that appears (Figure 15.29).

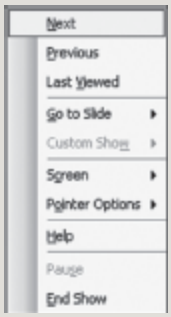


Figure 15.29 Pop-up Menu

To navigate to previous slide in Normal view, follow *any* of the steps given below:

- Click the **Previous Slide** ( ) icon on the vertical scrollbar.
- Click the previous slide number.

To view the previous slide in the Slide Show view, right-click on the slide and select **Previous** from the pop-up menu. Note that to move to a specific slide in Slide Sorter view, double-click the desired slide number, which displays the slide in the Normal view.

Numbering Slides: In PowerPoint, you can easily add numbers to the slides. Numbering the slides enables to locate or reference a particular slide quickly. You can insert slide numbers in a particular slide or in all slides of a presentation. To insert slide numbers, follow the steps given below:

1. Select **Slide Number** from the **Insert** menu to display the Header and Footer dialog box (Figure 15.30).

Figure 15.30 Header and Footer Dialog Box

2. Select the **Slide number** checkbox to display the slide numbers on the slides.
3. Select **Apply** to insert slide number in the current slide, or **Apply to All** to insert the numbering in all slides.

Changing Background: You can change the appearance of your slide background by changing its colour, shade, pattern, or texture. PowerPoint contains a number of pre-defined backgrounds or, if needed, you can create your own background. Note that backgrounds can be applied to individual slides or to the whole presentation. To change the background, select **Background** from the **Format** menu, .to display the Background dialog box (Figure 15.31).

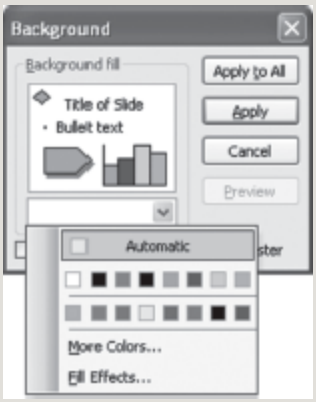


Figure 15.31 Bakcground Dialog Box

Click on the colour drop-down box, which contains following options:

- Colour scheme, which contains eight colours below **Automatic** to change the colour of the slide(s).
- **More Colors** option to select colour from a range of colours.
- **Fill Effects** to provide Gradient, Texture, Pattern, or Picture effects to the slide(s) (Figure 15.32).

Figure 15.32 Fill Effects

After you have made your selection, select **Apply** to apply the changes to the current slide, or click **Apply to All** to apply the changes to all slides.

15.5.4 Working with Text

There are a number of ways in which text can be added into the slides. Text can be added using text placeholder, text box, an AutoShape, and WordArt.

Adding Text using Placeholder: As stated earlier, PowerPoint includes 27 types of AutoLayouts for creating slides. Many of these layouts contain text placeholders for titles, body text, and bulleted lists. To add text in placeholder, click in the placeholder and type the text. You can change the size and position of a text placeholder or apply a different AutoLayout to an existing slide, at any time while creating a presentation (Figure 15.33).

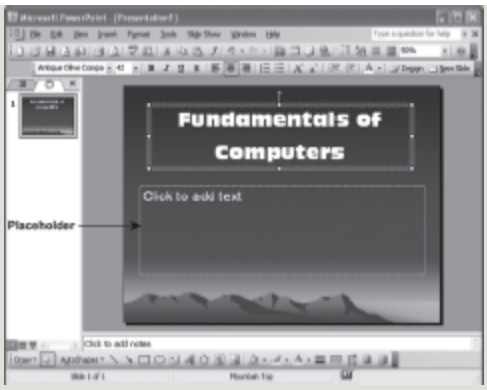


Figure 15.33 Inserting Text using Placeholder

Adding Text using Text Box: Text in slides can also be added through the use of text box. The text box is available on the Drawing toolbar. Note that if the Drawing toolbar is not visible, then right-click anywhere on the menu bar and select Drawing from the pop-up menu. To add text in the slide using text box, follow the steps given below:

1. Select the slide where you want to add a text box.
2. Click **Text Box** button ( ) on the Drawing toolbar.
3. Select the area in the slide where you want to place the text box and start writing the text in the text box (Figure 15.34).

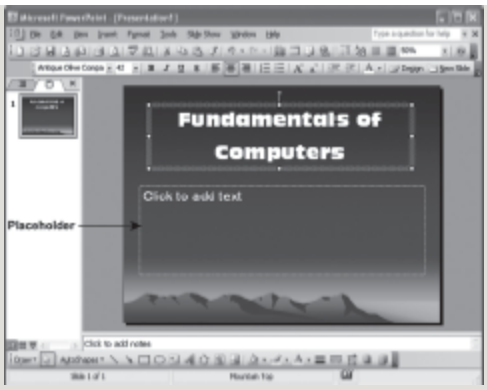


Figure 15.34 Inserting Text using Text Box

Adding Text to an AutoShape: Text can also be added in the AutoShape. AutoShapes contain several categories of shapes, including lines, basic shapes, flowchart elements, stars and banners, and callouts that you can use in your presentation. To add text in AutoShape, follow the steps given below:

1. Select the **AutoShapes** option from the Drawing toolbar.
2. Choose a shape from the list of options available. Note that the chosen shape has a background colour, which makes it invisible.
3. To add text, right-click the shape and click **Add Text**(Figure 15.35).

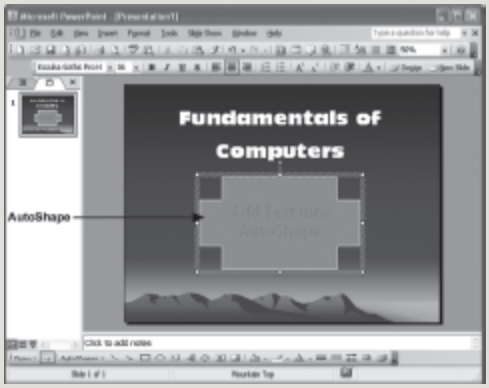


Figure 15.35 Inserting Text into AutoShape

15.5.5 Formatting the Text

As you enter and edit text, you can change its appearance to add emphasis and make the presentation easier to read. This is known as *formatting*.

It is the process of determining how the typed text will appear on a printed page. Formatting is applied by altering the appearance of text by setting the typeface (font), size, line spacing, and colour. Formatting can be done by using the Formatting toolbar or by using the **Format** menu. To format text, follow the steps given below:

1. Select the text to be formatted.
2. Use buttons on the Formatting toolbar to format the selected text. Some of the buttons used in formatting text are given in Table 15.1.

Table 15.1 Formatting Text

Command	Button	Description
Font		Changes the font of the selected text
Font Size		Changes the size of the selected text
Bold		Bolds the selected text
Italics		Italicizes the selected text
Underline		Underlines the selected text
Font Color		Changes font colour of the selected text

Text can also be formatted using the Font dialog box. For doing this, follow the steps given below:

1. Select the text to be formatted.
2. Select **Font** from the **Format** menu to display the Font dialog box (Figure 15.36).

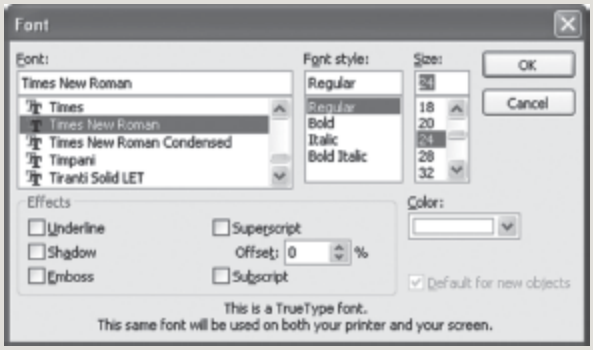


Figure 15.36 Font Dialog Box

3. In this dialog box, the font, font style, and text size can be specified. This dialog box also allows you to apply underline, colour, and other effects to the selected text.
4. Once formatting is complete, click **OK** to apply the formatting changes to the selected text.

Aligning Text and Paragraph Indenting:

Alignment and indentation operations are a part of paragraph formatting. These operations are performed to give a well-defined structure to the presentation. This gives the slide a professional look. To align or indent a paragraph, follow the steps given below:

1. Select the paragraph to be aligned or indented.
2. Use buttons on the Formatting toolbar to align or indent the paragraph. Some of the buttons used in aligning or indenting paragraph are given in Table 15.2.

Table 15.2 Paragraph Formatting

Command	Button	Description
Align Left		Left aligns the paragraph or selected text
Center		Centre aligns the paragraph or selected text
Align Right		Right aligns the paragraph or the selected text
Decrease Indent		Decreases the indent of the selected paragraph
Increase Indent		Increases the indent of the selected paragraph

Note: *Various other options for formatting the text and paragraph are available in the Format menu.*

15.5.6 Working with Graphics

In PowerPoint 2003, graphical elements can help you create eye-catching slides for a presentation. Graphics are frequently used with text to add emphasis and visual impact. PowerPoint 2003 provides you the option of adding AutoShapes, WordArt, and Clip Art.

AutoShapes: Microsoft PowerPoint comes with a set of readymade AutoShapes containing several categories of shapes, including lines, basic shapes, flowchart elements, stars and banners, and callouts that you can use in your presentations. They allow you to draw more complicated shapes in your slides quickly and more accurately. You can resize, rotate, flip, colour, and combine the shapes also. To use AutoShapes, follow *any* of the steps given below:

- Select **Picture** from the **Insert** menu and click **AutoShapes** to display AutoShapes toolbar.
- Select **Toolbars** from the **View** menu and click **Drawing** to display the Drawing toolbar that contains the **AutoShapes** menu (Figure 15.37).

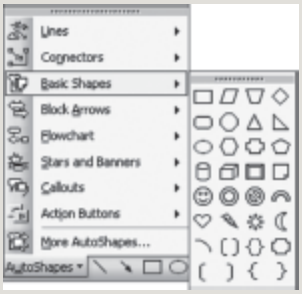


Figure 15.37 Autoshapes

Table 15.3 lists all the available buttons on the AutoShapes toolbar.

Table 15.3 AutoShapes Toolbar Options

Command	Button	Description
Lines		Draws a straight line, arrow, double arrow, curve, freeform, or scribble
Connectors		Draws connector lines to connect objects. There are three connector lines: straight, elbow, and curved
Basic Shapes		Draws two- and three-dimensional shapes, icons, braces, and brackets
Block Arrows		Draws block arrows from many types of two- and three-dimensional arrows
Flowchart		Draws flowchart symbols
Stars and Banners		Draws stars, bursts, banners, and scrolls
Callouts		Draws callouts, which are used to label some text on the document
Action Buttons		Readymade buttons that appear pressed in when you click them during a slide show
More AutoShapes		Used to choose more available AutoShapes

Now, to draw AutoShapes in the presentation, follow the steps given below:

1. Select the required shape on the toolbar. Use the drag-and-drop method to draw the shape in a slide (Figure 15.38).

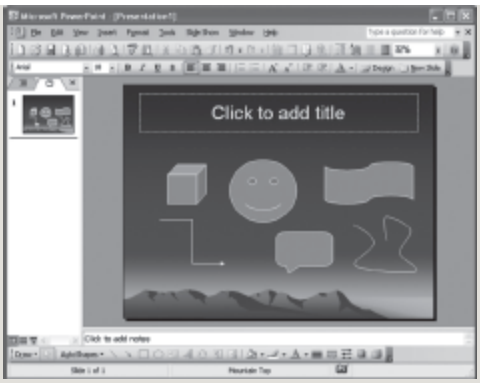


Figure 15.38 Inserting AutoShaps

2. When the shape has been created, you can resize it to fit your requirement. To resize AutoShape, use the sizing handle and other adjustments, which are specific to each shape. AutoShape can also be rotated. It is same as discussed in Microsoft Word.

Formatting Shapes: PowerPoint allows you to quickly and efficiently change the look of the AutoShape. There are several options available, which can be used to change the AutoShape appearance. To format the AutoShape, right-click on the AutoShape and select **Format AutoShape** from the shortcut menu. This displays the Format AutoShape dialog box, which contains six tabs to format the AutoShape. The options available under these tabs are listed in Table 15.4.

Table 15.4 Formatting Shapes

Tabs	Description
Colors and Lines	Allows you to choose the fill colour, line colour, and style for the shape
Size	Allows you to determine the size and scale of the image
Position	Allows you to control the position of the AutoShape on the slide
Picture	Allows you to crop and adjust the colour of the AutoShape
Text Box	Allows you to set the text alignment and internal margin of the text inside the AutoShape
Web	Allows you to enter text that will be displayed in place of graphics while the graphics is being loaded by the browser

Word Art: PowerPoint 2003's WordArt lets you take plain text and add some vigour to a presentation. Using WordArt you can create shadowed, skewed, rotated, and stretched text, as well as text that can be integrated into pre-defined shapes. It helps you to easily create banners, seals, and logos. To insert a WordArt, follow the steps given below:

1. Click **Insert WordArt** button ( ) on the Drawing toolbar. Alternatively, select **Picture** from the **Insert** menu, and then select **WordArt.** When you perform any of these actions, the WordArt Gallery dialog box is displayed (Figure 15.39).

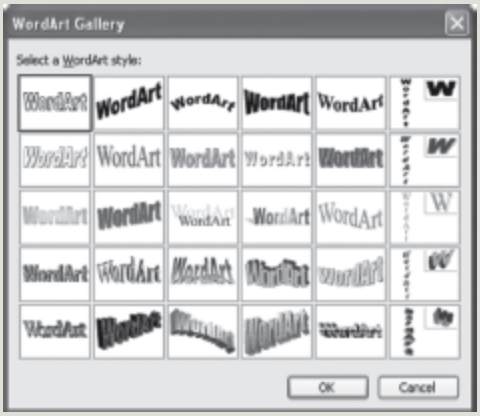


Figure 15.39 WordArt Gallery Dialog Box

2. Select a WordArt style to apply to the slide and click **OK** to close the dialog box. This opens Edit WordArt Text dialog box, where text is entered for the chosen style (Figure 15.40).

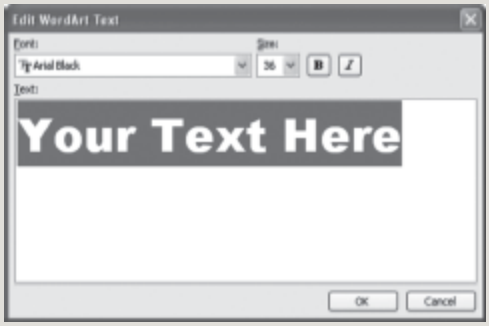


Figure 15.40 Edit WordArt Text Dialog Box

3. Select a font by clicking the down arrow next to the **Font** list. The font size can also be selected by clicking the down arrow next to the **Size** box. To make the text bold, click the (B) button, and to make the text italics, click the (I) button.
4. When text has been entered and font selection made, click **OK**. The WordArt now appears in the slide with the selection handles visible on it. You can resize, move, and rotate the WordArt by using these selection handles (Figure 15.41).

Figure 15.41 WordArt in a Slide

To add or change effects to the text, use the buttons on the WordArt and Drawing toolbars. The WordArt toolbar appears when you click the WordArt special text. The options available under the WordArt toolbar are the same as discussed in *Chapter 13*.

Drawing Toolbar: The Drawing toolbar is used to change and enhance the drawing objects with colours, patterns, borders, and other effects. The options included in the Drawing toolbar have also been discussed in *Chapter 13*.

Pictures: Pictures are used in slides to enhance its readability and understandability of the presentation. These are bitmaps, scanned pictures, photographs, and Clip Art. You can use any image-editing software to create photographs, line drawings, and other graphics, as long as you can save the file in a format that PowerPoint 2003 supports. You can insert a picture from a file or Clip Art in the same manner as discussed in *Chapter 13*.

15.5.7 Adding Multimedia

A presentation can be made interactive and interesting by adding multimedia effects, which helps in capturing audience's attention, as multimedia content can often communicate more information than a slide, which contains only few sentences. You can add multimedia effects such as animations, video, and sound effects to the slides.

Adding Animations: Animation effects are added to grab the audience's attention as well as to reinforce some important point that needs to be highlighted. Animation can be in terms of special sound or visual effects, which are added to the text or other objects such as a chart or

picture. To add animation effects to a slide or entire presentation, follow the steps given below:

1. Click on the placeholder that contains the text or graphic that you wish to animate.
2. Select **Animation Schemes** from the **Slide Show** menu to display Slide Design task pane.
3. Select the desired animation scheme from the **Apply to selected slides** section. The selected animation scheme is applied to the current slide.
4. To apply the selected animation scheme to all slides, select **Apply to All Slides.**

Custom animations offer more control over the animation. You can choose the order of the slides on which animation is applied as well as decide the direction the animation should travel in, apply a sound to the animation, and decide how the text appears in the slides. To add custom animation, follow the steps given below:

1. Select the text or graphic to animate.
2. Right-click text or graphic and select **Custom Animation** from the shortcut menu, or select Custom Animation from the Slide Show menu. This displays the Custom Animation task pane (Figure 15.42).

Figure 15.42 Custom Animation Dialog Box

3. Select various effects from **Add Effect** to assign animation to the objects.
4. Once the effects have been applied, they can also be modified by setting various options in the **Modify** section.

Adding Transition: Transition effects are used to insert effects between slides. They are also used to indicate a new section of a presentation or to emphasize a certain slide. You can choose from a variety of transitions, which are present in PowerPoint 2003. Transition effects can be added in the Slide or Slide Sorter view. To add transition effects, follow the steps given below:

1. Select **Slide Transition** from the **Slide Show** menu to display the Slide Transition task pane (Figure 15.43).

Figure 15.43 Slide Transition Task Pane

2. Select the desired transition from the **Apply to selected slides** section, to apply the transition to the displayed slide (Figure 15.44).

Figure 15.44 Selecting Trabsition

3. Select various options under **Modify transition** section to change the speed and sound of the transition.

Note: *Transition effects are the style of moving from one slide to the next, while animation effects are the style used to display the information on each slide.*

15.5.8 Running Slides

After creating slides, working with text and graphics, and adding transition and animation effects to the slides, the next thing is to view all the slides collectively as a presentation. PowerPoint 2003 provides a number of ways to view and deliver presentations so that they are well designed and look professional. You can use the Slide Show to review a presentation before printing it or to prepare to show it electronically. To view the slide show, perform any of the steps given below:

- Select **Slide Show** from the **View** menu.
- Select **View Show** from the **Slide Show** menu.
- Click on the Slide Show view button ( ) at the bottom-left of the screen (Figure 15.45).

Figure 15.45 Running Slide Show

While in the slide show, right-click anywhere on the screen to display a pop-up menu containing number of options. Some of the important options are listed in Table 15.5.

Table 15.5 Slide Show Buttons

Options	Description
Next	Used to move to the nest side
Previous	Used to move to the previous side
Last Viewed	Used to move to the last viewed slide in the presentation
Go to Slide	Used to move to spicific slide in the presentation
Pointer Options	Used to add information using pen
Help	Used to move to the Slide Show Help
End Show	Used to move to the slide show

Highlighting Parts of the Slide Show: You can use the highlighting option while you are giving the presentation to add emphasis to important information. To highlight a slide during a slide show, follow the steps given below:

1. Select View Show from the **Slide Show** menu.
2. Right-click in the slide and select **Highlighter** from the **Pointer Options**.
3. Click and hold the highlighter on the slide. Drag the highlighter around the area on the slide you want to highlight until a line encircles the text.
4. Right-click in the slide and select **Arrow** from the **Pointer Options** to change the highlighter back to the mouse pointer (Figure 15.46).

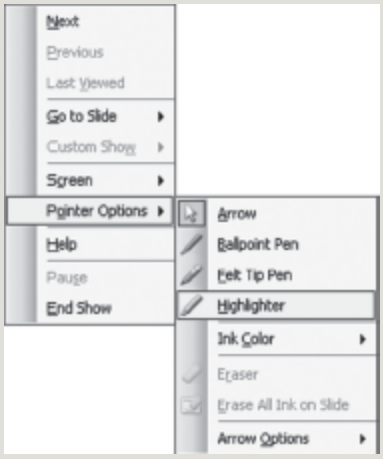


Figure 15.46 Pointer Options

Setting Time between Slides: You can insert time between the slides in the slide show. This is required when there is some narration over the slides, or there is some graphical representation, which needs proper explanation.

If you do not want to manually move through a slide show, there are two ways you can set the length of time for which the slide appears on the screen.

- You can set a time manually for each slide, and then run the slide show and view the timings you set.
- You can use the *rehearsal* feature. Here you can record timings automatically as you rehearse. You can also adjust the timings you have already set and then rehearse the new ones. To add time using this feature, follow the steps given below:

1. Select **Rehearse Timings** from the **Slide Show** menu. This starts the slide show with a Rehearsal toolbar that appears at the corner of the screen, as shown in Figure 15.47.

Figure 15.47 Rehearsal Toolbar

2. The timer appears in the upperleft-hand corner of the slide and begins to count down the seconds. The white box refers to the timing for each slide. The other timer refers to the total length of the slide show.
3. If you want to repeat the rehearsal of a slide, click the **Repeat** button ( ) on the Rehearsal toolbar. The current slide repeats and the timing for it starts again.
4. Rehearse your presentation until it is finished. After the presentation is complete, a message box appears. It displays the final running time and presents a message, as shown in Figure 15.48.

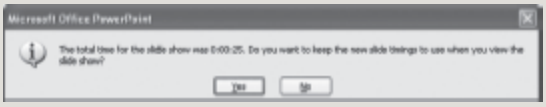


Figure 15.48 Message Box

5. Click **Yes** to record the new slide timings or **No** to return to the PowerPoint window.

15.5.9 Package for CD

Presentation can be created both for business and for professional use. For example, you may create a

presentation, which is to be delivered at some conference. To deliver a presentation only the presentation file is needed. In case of a small presentation, you can easily carry the file in one floppy. However, this approach may not work well for larger presentations especially one with lots of graphics. PowerPoint contains a useful feature called *Package for CD*, which solves the above problem and also allows you to view the presentation file on a computer that does not have the PowerPoint installed. To use this feature, follow the steps given below:

1. Select **Package for CD** from the **File** menu to display the Package for CD dialog box (Figure 15.49).

Figure 15.49 Package for CD Dialog Box

2. Type the name for your folder or CD in **Name the CD** text box.
3. Select **Add Files** to include more than one file from your computer.
4. Select Options to display **Options** dialog box. This dialog box has the following options (Figure 15.50):

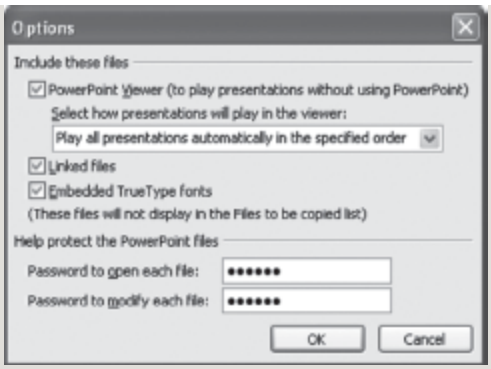


Figure 15.50 Options Dialog Box

1. **PowerPoint Viewer**: Select this checkbox to play the presentation, even if the destination computer does not have PowerPoint 2003 installed.
2. **Linked Files**: Select this checkbox to include all linked files.
3. **Embedded TrueType Fonts**: Select this checkbox to display the font correctly, even if the destination computer does not have particular font installed.
4. **Password to Open Each File**: Provide password to open the PowerPoint presentation.
5. **Password to Modify Each File**: Provide password to modify the PowerPoint presentation.

5. Click **OK** to close the Options dialog box and return to the Package for CD dialog box.
6. Finally, to copy the files to a folder or to a CD, click the corresponding button.

15.6 PRINTING IN POWERPOINT

Printing in Microsoft PowerPoint is similar to printing in other Windows-based applications. However, several options, particularly those concerned with arranging the page, are specific to the application. PowerPoint provides a number of choices for you to take printouts; colour or black and white, single slides, or several slides per page.

15.6.1 Setting Page Layout

The Page Layout option is used to view the existing page layout or to set a new layout. To set a new layout, select **Page Setup** from the **File** menu to display the Page Setup dialog box (Figure 15.51).

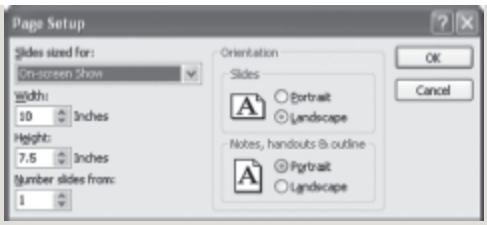


Figure 15.51 Page Setup Dialog Box

This dialog box contains options, which allows you to customize the page layout as desired. These options are:

Slide Size: PowerPoint allows you to determine the size of the slide by choosing from the number of options, which are listed below:

- On-screen Show (default)
- Letter Paper (8.5 × 11 in)
- A4 Paper (210 × 297 mm)
- 35 mm Slides
- Overhead
- Banner
- Custom

Slide Width and Height: You can adjust the width and height of the slide by adjusting the values in the Width and Height box, respectively.

Orientation: Orientation is the basic layout of the slides in the presentation. The default orientation is Landscape. To choose different orientation, click the radio button in front of the type of orientation and the change will be reflected in the slide.

15.6.2 Printing Presentation

Once you are ready with all the formatting and editing, you can print the presentation. To print a presentation,

click the **Print** button ( ) on the Standard toolbar or select **Print** from the **File** menu to display the Print dialog box. This dialog box allows you to select the number of copies and how many pages of the document you want to print. It contains **Print what** box where you can choose the printing area (Figure 15.52). The options available for printing in **Print what** box are listed in Table 15.6.

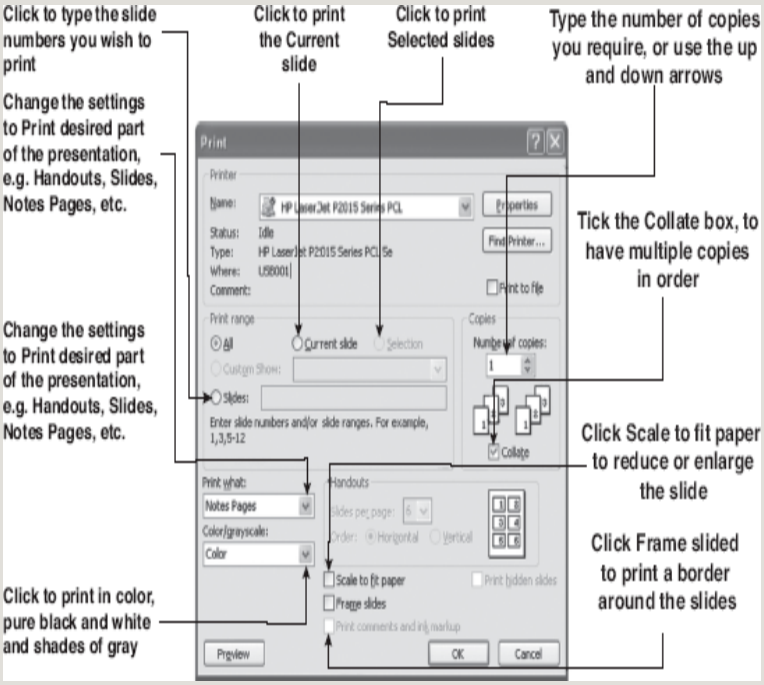


Figure 15.52 Print Dialog Box

Table 15.6 Printing Options

Options	Description
Slides	Prints all slides in the presentation
Handouts	Prints handouts
Notes Pages	Prints notes pages. In the printout, slide appears at the top of the page and the presentation notes at the bottom of the page
Outline View	Prints the outline text of your presentation (as it appears in the outline pane)

Printing Handouts: Power-Point 2003 allows you to print the handouts prepared in presentation. Using handouts as an option lets you print two, three, four, six, or nine slides per page. It also provides space for the audience to add notes as you give the presentation. To print handouts, select the **Handouts** option in the **Print what** box (Figure 15.53).

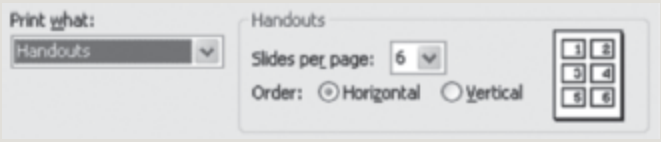


Figure 15.53 Printing Handouts

- Select the number of slides you want on the handouts in the **Slides per page** box.
- Select **Horizontal** or **Vertical** to specify the order in which you want the slides to appear on the page. After selecting appropriate options, click **OK** to print the handouts.

LET US SUMMARIZE

1. Microsoft PowerPoint 2003 is powerful, yet easy-to-use presentation software. Its primary purpose is to help you design professional-style presentations quickly and easily.
2. Click Start, point to All Programs, then point to Microsoft Office, and then select Microsoft Office PowerPoint 2003.
3. A new presentation can be created using the AutoContent wizard, Design Template, and Blank presentation.

4. To open an existing presentation, click on the Open button ( ) on the Standard toolbar. Alternatively, select Open from the File menu.
5. To save a presentation, Microsoft PowerPoint provides two menu options, namely Save and Save As. By default, Microsoft PowerPoint files are saved with .PPT extension.
6. To close the presentation, select Close from the File menu.
7. PowerPoint provides different types of screen layouts for constructing presentation in addition to the Slide Show. View is a way in which one can see and work with the presentation and slides.
8. A presentation is created using the slides, which contains text, graphics, animations, and so on. PowerPoint is designed in a way to give a consistent appearance to your slides. There are two ways in PowerPoint, which help you to choose the look of your slides: design template and masters.
9. Design templates contain colour schemes, slide and notes masters with custom formatting, and fonts, all designed to create a desired look.
10. Masters in PowerPoint are used to control many facets of the slides such as backgrounds, font typeface, font size, colours, bullets and locations for all the main components, tab, and indent. PowerPoint has three masters, namely Slide Master, Notes Master, and Handouts Master.
11. Slides can be added in the outline view or in the slide view. To insert slide, select New Slide from the Insert menu, or click the New Slide ( ) button on the Standard toolbar. This displays New Slide dialog box from which you can choose the desired layout.
12. In PowerPoint, you can navigate to the next and previous slides by following any of the five view options, that is, Normal, Outline, Slide View, Slide Sorter, and Slide Show.
13. You can change the appearance of your slide background by changing its colour, shade, pattern, or texture. PowerPoint contains a number of pre-defined backgrounds or, if needed, you can create your own background.
14. There are a number of ways in which the text can be added into the slides. Text can be added using text placeholder, text box, an AutoShape, and WordArt.
15. Formatting is the process of determining how the typed text will appear on a printed page. It is applied by altering the appearance of text by setting the font type and font size, line spacing, and colour. To format the text, first select the text to be formatted, and then use various buttons on the Formatting toolbar. Text can also be formatted by using the Format menu.
16. In PowerPoint 2003, graphical elements can help you create eye-catching slides for a presentation. Graphics are frequently used with text to add emphasis and visual impact. PowerPoint 2003 gives you the option of adding AutoShapes, WordArt, and ClipArt.
17. A presentation can be made more interactive and interesting by adding multimedia effects such as animations, video, and sound effects to the slides. This helps in capturing audience's attention as multimedia content can often communicate more information than a slide, which contains only few sentences.
18. Animation effects are added to grab the audience's attention as well as to reinforce the point that needs to be highlighted.

Animation can be special sound or visual effects which can be added to the text or other objects such as a chart or picture.

19. Transition effects are used to insert effects between slides. They are also used to indicate a new section of a presentation or to emphasize a certain slide. You can choose from a variety of transitions, which are present in PowerPoint 2003.
20. After you have finished with creating slides, working with text and graphics, adding transition and animation effects to the slides, the next thing is to view all the slides collectively as a presentation. PowerPoint 2003 provides a number of ways to review and deliver presentations so that they are well designed and look professional. You can view the presentation as a slide show in the Slide view.
21. You can add time between the slides in the slide show. This may be required when there is some narration over the slides, or there is some graphical representation, which requires a proper explanation.
22. Sometimes, presentation may be created for business or professional use. For example, you may create a presentation, which is to be delivered at some conference. To give a presentation all you have to do is bring your presentation file with you. This approach is suited for small presentations, but it may not work that well for larger presentations especially one with lots of graphics. PowerPoint contains a useful feature called Package for CD, which solves the above problem.
23. To print a presentation, click the Print button ( ) on the Standard toolbar. Alternatively, use the Print dialog box, which can be displayed by selecting Print from the File menu.
24. PowerPoint provides four options while printing—Slides, Handouts, Note Pages, and Outline View.

EXERCISES

Fill in the Blanks

1. By default, ... and ..... toolbars are displayed in the PowerPoint environment.
2. ..... helps to create a presentation using a series of some basic steps.
3. ..... option helps to update changes made to a file while option is used to make multiple copies of the same file.
4. ..... is a way in which one can see and work with the presentation and slides.
5. ..... view option allows you to view the slides (full screen), as they would appear in a slide show.
6. ..... is a presentation containing PowerPoint masters with custom formatting, fonts, colour schemes, etc.
7. PowerPoint provides ..... types of AutoLayouts for creating slides.
8. ..... is the process of changing the appearance of text and making presentation easier to read.
9. To insert time between the slides in the slide show, feature is used.
10. ..... feature is used to run a presentation even if PowerPoint is not installed on a computer.

Multiple-choice Questions

1. AutoShapes contain several categories of shapes, including .....
    1. Lines
    2. Basic shapes
    3. Flowchart elements
    4. All of these
2. Using WordArt you can easily create .....
    1. Banners
    2. Seals
    3. Logos
    4. All of these
3. ..... are added to grab the audience's attention as well as to reinforce the point that needs to be highlighted.
    1. Animation effects
    2. Transition effects
    3. Screen views
    4. Package for CD
4. ..... are used to insert effects between slides. They are also used to indicate a new section of a presentation or to emphasize a certain slide.
    1. Drawing toolbar
    2. Transition effects
    3. Logos
    4. None of these
5. ..... In master, you can add images, text, and colour to your handouts.
    1. Slide
    2. Notes
    3. Handout
    4. Title
6. Orientation is the basic layout of the slides in the presentation. The default orientation is the .....
    1. Landscape
    2. Animation effects
    3. Portrait
    4. Page Size
7. Blank Presentation does not include .....
    1. Preset design
    2. Colour
    3. Content suggestion
    4. All of these
8. ..... wizard helps to create a presentation in a quick way and is very helpful for new users.
    1. AutoContent
    2. Design Template
    3. Blank Presentation
    4. None of these
9. ..... view option displays all the slides in a presentation as thumbnails.
    1. Normal view
    2. Slide Show view
    3. Slide Sorter view
    4. None of these

10. A presentation is created using the ....., which contains text, graphics, and animations.
    1. Page Setup
    2. Slides
    3. Animation
    4. Transition

State True or False

1. A new presentation is created using AutoContent wizard, Design template, and Blank presentation.
2. Design template includes no pre-formatted layouts, fonts, and colours.
3. By default, Microsoft PowerPoint presentation files are saved with .PPT extension.
4. Slide Show displays all the slides in a presentation as thumbnails.
5. In Notes master, you can add header and footer to the slides.
6. While duplicating the slides, their layout is also copied.
7. To insert AutoShapes, select Picture from the Insert menu and click WordArt.
8. To highlight some parts of the slide show, highlighter is used.
9. Transition effects are the style of moving from one slide to the next.
10. The default orientation in PowerPoint is Landscape.

Descriptive Questions

1. Describe the steps involved in creating a new presentation using AutoContent Wizard.
2. What views are available in PowerPoint 2003? How is Slide Sorter view different from Slide Show view?
3. What is the role played by Masters in PowerPoint? Differentiate between Slide Master and Handout Master.
4. How are slides inserted and deleted in PowerPoint?
5. Explain Formatting in PowerPoint.
6. What is the role played by multimedia in presentation? Give the steps involved in adding animations.
7. Why is timing required between slides in PowerPoint?
8. Explain the benefit of using Package for CD. Give the steps involved in it.
9. Explain the options available in Page Setup dialog box.
10. Is it possible to take printouts of the handouts in PowerPoint? If yes, explain how.

ANSWERS

Fill in the Blanks

1. Standard, Formatting
2. AutoContent wizard
3. Save, Save As
4. View
5. Slide Show
6. Template

7. 27
8. Formatting
9. Rehearsal
10. Package for CD

Multiple-choice Questions

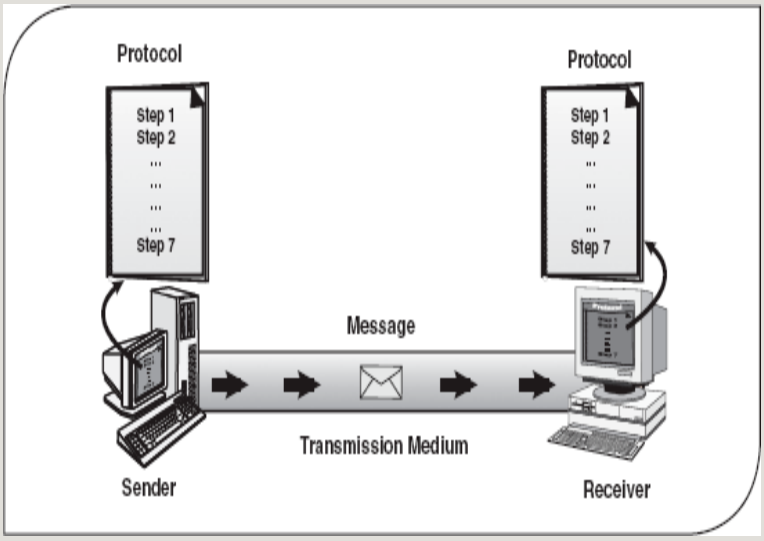
1. (d)
2. (d)
3. (a)
4. (b)
5. (c)
6. (a)
7. (d)
8. (a)
9. (c)
10. (b)

State True or False

1. True
2. False
3. True
4. False
5. False
6. True
7. False
8. True
9. True
10. True

Chapter 16

Data Communication and Computer Networks



This chapter introduces the reader to data communication and computer network. Data communication is the exchange of data between two devices via some form of transmission medium. It is carried within a collection of computers, which are linked together to share information and resources. The discussion begins with an outlook on data communication and its components. It is continued with various transmission modes, transmission media, modulation, and switching techniques. In the later stages of the chapter, the reader is acquainted with different types of network and topologies. The chapter concludes with an overview of various networking devices.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The basic concepts of communication and networking
- Data communication, which is the exchange of data between two devices using some form of transmission medium
- Transmission media, which refers to the physical media through which communication signals are transmitted
- Modulation, which is the addition of information to a signal carrier wave
- Multiplexing, which is a technique used for transmitting several signals simultaneously over a single communication channel
- Switching, which means routing traffic by setting up temporary connections between two or more network points
- Computer network—a collection of two or more computers connected together to share information and resources
- Network topology, which is the way a network is laid out, either physically or logically
- Communication protocol—a set of rules that coordinate the exchange of information
- Network devices, which interconnect individual computers for efficient communication

16.1 INTRODUCTION

The term *communication* in simple words means sending or receiving information. In ancient times, people used the beating of drums, smoke signals, mirrors reflecting sunlight, homing pigeons, and so on to send messages. Another way of long-distance communication was the use of homing pigeons to carry messages. With the advancement in science and technology, various devices were invented for communication. Telegraph was one of the most prevalent ways of data communication in

the nineteenth century. In 1870, Alexander Graham Bell invented the telephone, which revolutionized the way long-distance communication used to take place. Voice communication became common after the invention of the telephone. In the 1940s and 1950s, a major breakthrough was made in the development of computers and setting up a network of computers for electronic transfer of information or data communication. In fact, the first network was established in 1940, when George Stibitz and Samuel Williams, researchers at Bell Laboratories, developed complex number generator—one of the earliest digital computers. They made it possible for the users to operate the computer remotely from teletype machines. They also developed a system to allow users to communicate with the machine even more remotely, through telex machines. The advancement of computer networks and later the Internet acted as a boost for data communication.

Whenever more than one computer is used at the same location, networking them together facilitates exchange of data and information (data communication) between connected computers in a fast and efficient manner. Networking also helps sharing expensive resources like printers, backup tape drives, scanners, Internet connections, and so on. In this chapter, we discuss how communication system works, various transmission medias and devices, different techniques of data communication, types of computer network, and various network topologies.

16.2 DATA COMMUNICATION

Data *communication* is the exchange of data between two devices using some form of wired or wireless transmission medium. It includes the transfer of data, the method of transfer, and the preservation of the data during the transfer process. To initiate data

communication, the communicating devices should be a part of an existing communication system. For data communication to be effective, the following three fundamental characteristics should be considered:

THINGS TO REMEMBER

Communication System

A *communication system* can be defined as the collection of hardware and software that facilitates intersystem exchange of information between different devices. When we communicate, we are sharing information. This sharing can be local (face-to-face communication) or it may be remote (communication over distance).

- **Delivery**: The system must deliver data to the correct or the intended destination.
- **Accuracy**: The system must deliver data accurately (error free).
- **Timeliness**: The system must deliver data in a timely manner without enough time lags.

16.2.1 Components of Data Communication

There are five basic components in data communication system.

- **Message**: The message is the information that is to be communicated.
- **Sender**: The sender is the device that sends the message.
- **Receiver**: The receiver is the device that receives the message.
- **Medium**: The transmission medium is the physical path that communicates the message from the sender to receiver.
- **Protocol**: Protocol refers to a set of rules that coordinate the exchange of information. Both sender and receiver should follow the same protocol to communicate data. Without the protocol, the sender and receiver cannot communicate with each other; just as a person speaking English cannot be understood by a person who knows only Hindi (Figure 16.1).

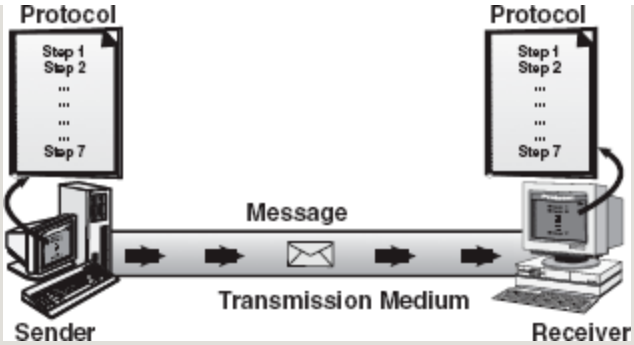


Figure 16.1 Data Communication Components

Suppose you want to convey your final marks to your mother. You call her via telephone and inform her about your marks. Here your marks is the message, you are the sender, your mother is the receiver, the telephone line is the medium, and the language in which you are conversing is the protocol.

16.2.2 Data Transmission Mode

Data transmission mode refers to the direction of signal flow between two linked devices. There are three types of transmission modes: simplex, *half-duplex, and full-duplex* (Figure 16.2).

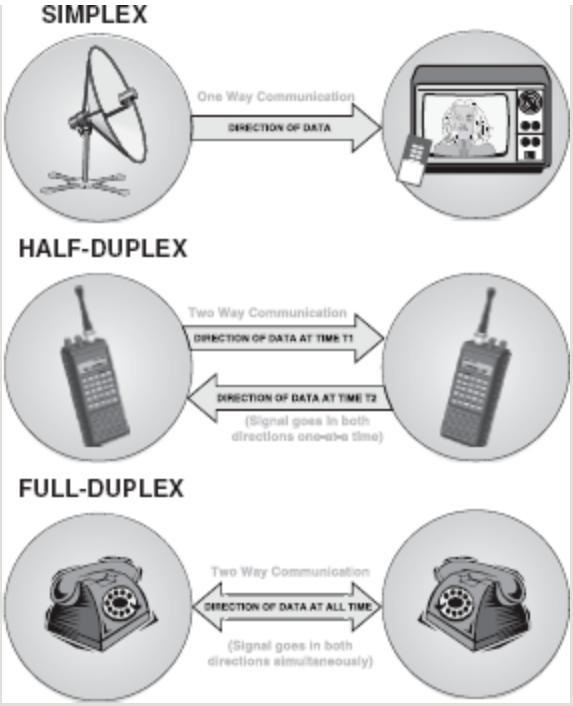


Figure 16.2 Data Transmission Modes

Simplex: Simplex transmission is unidirectional. The information flows in one direction across the circuit, with no capability to support response in the other direction. Only one of the communicating devices transmits information, the other can only receive it. Television transmission can be considered as an example of simplex mode of transmission where the satellite only transmits the data to the television, vice versa is not possible.

Half Duplex: In half-duplex mode, each communicating device can receive and transmit information, but not at the same time. When one device is sending, the other can only receive at that point of time. In half-duplex transmission mode, the entire capacity of the transmission medium is taken over by the device, which is transmitting at that moment. Radio

wireless set is an example of half-duplex transmission mode where one party speaks and the other party listens.

Full Duplex: Full-duplex transmission mode, also known as the *duplex mode*, allows communicating devices to both transmit and receive data simultaneously. A full-duplex mode can be compared to a two-way road with traffic flowing in both directions. A standard voice telephone call is a full-duplex call because both parties can talk at the same time and be heard.

16.2.3 Analog and Digital Data Transmission

The major role of the physical medium is to move the data from one communicating device to another. The data can be transmitted by means of electromagnetic signals. These signals may be in one of the two forms: *analog* or *digital*. Depending on the form of signal used, data transmission can be analog data transmission or digital data transmission.

Analog Signals: An analog signal is a continuous waveform that changes smoothly over time. The sine wave is the most fundamental form of an analog signal (see Figure 16.3). Sine waves can be described by three characteristics.

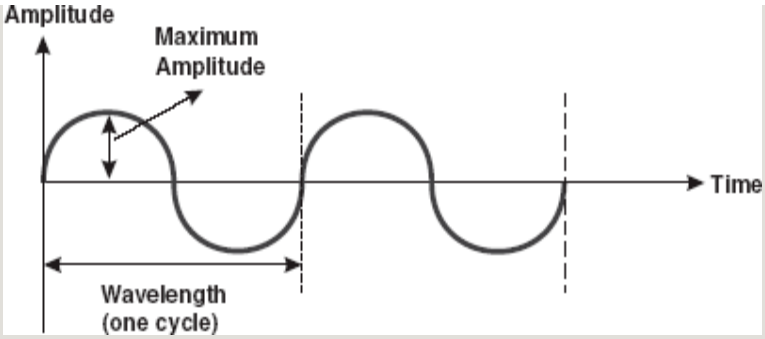


Figure 16.3 Analog Signal

- **Amplitude**: It is the value of the signal at any point on the wave. The maximum amplitude of a sine wave is the highest value it reaches on the vertical axis. The unit for amplitude depends on the type of the signal. For electrical signals, the unit is normally volts and amperes.
- **Frequency**: It refers to the number of cycles a signal completes in 1s. In other words, frequency means the number of times a signal wave goes up and down in a second and it is measured in Hertz (Hz). For example, if a signal wave completes one cycle in 1s, its frequency is 1 Hz.
- **Wavelength**: It refers to the distance between successive similar points of a given wave, that is, one cycle of the waveform.

Analog signals are perfect for carrying data such as voice or sound. However, these signals are prone to errors or noise that can be caused by the outside sources. Attenuation is another problem with analog signals because the amplitude of the wave naturally changes over distance.

Digital Signals: Digital signal is the data stored in the form of 0s and 1s. When the signal is at a high point, its value is 1 and when it is low, its value is 0. A signal in digital format has precise voltages that are not affected by noise or attenuation as compared with analog signals, which are very prone to noise. Digital signals can be represented by a graph similar to a bar graph. In Figure 16.4, 1 can be encoded as a positive voltage and 0 as a zero voltage.

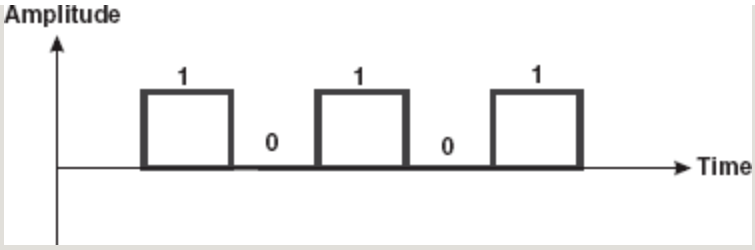


Figure 16.4 Digital Signal

To transmit data over analog telephone lines, a modem is required to convert the digital data signals to analog signals. When transmitted over long distances, analog signals require to be amplified, which can possibly distort the value of the data transmitted. When analog data is converted to digital data, it can be transmitted over digital signals faster and without distortion.

16.2.4 Data Communication Measurement

The measurement of the quantity of data that can be passed over a communication link in a given time is done in terms of bandwidth. Fundamentally, bandwidth refers to the maximum volume of data that can be transferred over any communication medium at a given point of time. The more the data needed to be transmitted in a given period, the more is the bandwidth required. On digital circuits, bandwidth is measured in *bits per second* (*bps*), which refers to the number of bits transmitted per second. A 1,000 bps is 1 kilobit per second or 1 kbps.

In the digital context, the level of bandwidth falls into three categories.

- **Narrowband**: In narrowband, there is a single transmission channel of 64 kbps or less. With advances in network technology, narrowband has come to be associated with any channel operating at less than 1.544 Mbps.

- **Wideband**: It is a medium-capacity communication channel, which carries data at speeds between 1.544 Mbps and 45 Mbps.
- **Broadband**: It is a transmission medium, capable of supporting a wide range of frequencies. It can carry multiple signals by dividing the total capacity of the medium into multiple, independent bandwidth channels, where each channel operates only on a specific range of frequencies. The bandwidth capacity in broadband is equal to 45 Mbps. Broadcast television, cable television, microwave, and satellite are examples of broadband technologies. These technologies are capable of carrying a great deal of information in a short amount of time, but are more expensive to use than technologies like the telephone, which require less bandwidth.

16.3 TRANSMISSION MEDIA

Transmission media refers to the physical media through which data can be transmitted from one point to another. As already stated, the data is transmitted from one device to another through electromagnetic signals. An *electromagnetic signal* is the combination of electric and magnetic fields, vibrating in conjugation with each other. Electromagnetic signals include power, voice, radio waves, infrared light, visible light, ultraviolet light, X-rays, and gamma rays. All these together constitute an electromagnetic spectrum. These signals can travel through vacuum, air or any other transmission medium. Voice signals are generally transmitted as current over metal cables. Radio frequencies are generally transmitted through air or space. Third type of electromagnetic energy is the visible light, which is currently being used for communication through fibre optic cable. Transmission media can be divided into two broad categories: *guided/wired media* and *unguided/wireless media*.

16.3.1 Guided/Wired Media

Guided transmission media use a cabling system that guide the data signals along a specific path. Cable is the medium through which information usually moves from one network device to another. It consists of a cable composed of metals like copper, tin, or silver. The data signal in the guided medium is bound by the cabling system; hence, guided medium is also known as *bound*

medium. There are three basic types of guided media—*twisted pair, coaxial cable*, and *optical fibre*.

Twisted Pair: The biggest network in the world, the telephone network, originally used only twisted-pair cabling and still does for most local connections. The name comes from the fact that each individually insulated conductor is part of a pair, making this a balanced medium, and that each pair is twisted together along its length, which helps to further protect it from interference from external sources. Each pair consists of a wire, used for receiving data signal, and a wire used for transmitting data signal. Twisted pairs are used in short-distance communication (less than 100 m). It comes in two forms, namely, unshielded and shielded with a metal sheath or braid around it. The two are commonly known as *unshielded twisted pair* (*UTP*) and *shielded twisted pair* (*STP*).

- **UTP Cable**: It is the most common type of telecommunication medium in use today. It is most suited for both data and voice transmission, hence is commonly used in telephone systems. The cable has four pairs of wires inside the jacket. Each pair is twisted with a different number of twists per inch to help eliminate interference from adjacent pairs and other electrical devices. The tighter the twisting, the higher the supported transmission rate and the greater the cost per foot. Each twisted pair consists of two metal conductors (usually copper) that are insulated separately with their own coloured plastic insulation (Figure 16.5).

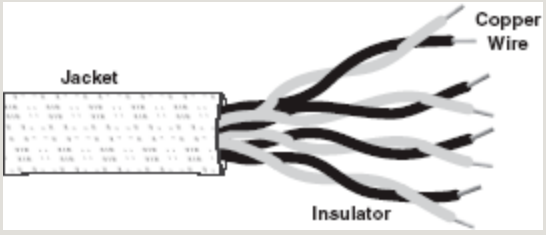


Figure 16.5 Unshielded Twisted-Pair Cable

- **STP Cable**: This cable has a metal foil or braided mesh covering that covers each pair of insulated conductors. The metal foil is used to

prevent infiltration of electromagnetic noise. This shield also helps to eliminate *crosstalk*, a phenomenon that can be experienced during telephone conversation when one can hear another conversation in the background. A disadvantage of UTP is that it may be susceptible to radio and electrical frequency interference. Shielded twisted pair is suitable for environments with electrical interference; however, the extra shielding can make the cables quite bulky (Figure 16.6).

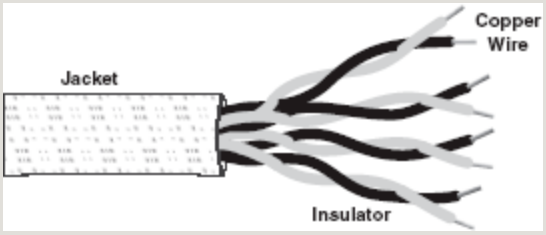


Figure 16.6 Shielded Twisted-Pair Cable

Note: *All signalling over twisted pair is of the baseband type. Baseband transmission use the entire media bandwidth for a single channel, that is, only one signal can be carried at a time, and all the systems take turns using it.*

Coaxial Cable: Coaxial cables have a single central conductor, which is made up of solid wire (usually copper). This conductor is surrounded by an insulator over which a sleeve of metal mesh is woven to block any outside interference. This metal mesh is again shielded by an outer covering of a thick material (usually PVC) known as *jacket* (Figure 16.7).

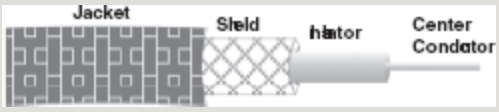


Figure 16.7 Coaxial Cable

Although coaxial cabling is difficult to install, it is highly resistant to signal interference. It can support greater cable lengths between network devices than twisted-pair cable. In addition, as compared to twisted pairs, it also offers higher bandwidth. A coaxial cable is capable of transmitting data at a rate of 10 Mbps. It is more expensive per foot, but cheaper per bytes of data transferred in a second. Coaxial cable is very robust and is commonly used in cable TV network.

Optical Fibre: Twisted-pair and coaxial cable transmit data in the form of current. Optical fibre, on the other hand, consists of thin glass fibres that can carry information in the form of visible light. The typical optical fibre consists of a very narrow strand of glass called the *core*. Around the core is a concentric layer of glass called the *cladding*. A typical core diameter is 62.5 microns (1 micron=10-6m). Cladding generally has a diameter of 125 microns. The cladding is covered by a protective coating of plastic, known as jacket (Figure 16.8).

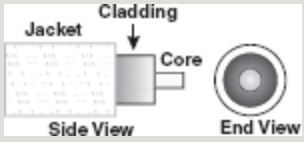


Figure 16.8 Optical Fibre

Optical fibre has the ability to transmit signals over much longer distances than coaxial and twisted pair. Light signal, unlike electrical signal, does not heat up the wire and is not susceptible to outside interference. It is extremely hard to tap into, making it desirable from the

security point of view. Optical fibre also has the capability to carry information at vastly greater speeds. A single optical fibre can pack hundreds of fibres, where each fibre has the capacity equivalent to that of thousands of twisted-pair wires. This capacity broadens communication possibilities to include services such as video conferencing and interactive services. However, fibre optics cable is by far the most expensive per foot. Moreover, it is more difficult to install and modify (Figure 16.9).

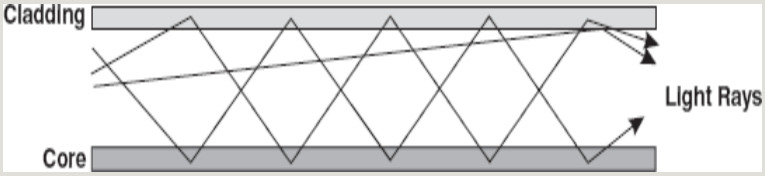


Figure 16.9 Signals Carried Over an Optical Fibre

THINGS TO REMEMBER

Optical Fibres

Optical fibres work on the principle that the core refracts the light and the cladding reflects the light. The core refracts the light and guides the light along its path. The cladding reflects the light back into the core and stops it from escaping through the medium. These light pulses, which can be carried over long distances via optical fibre cable, carry information.

16.3.2 Unguided/Wireless Media

Unguided media facilitates data transmission through earth's atmosphere (air, water, or vacuum) at a much faster rate and provides wide area coverage. The data signals are not guided or bound to a fixed channel to follow. Some commonly used unguided media of transmission include *radio frequency (RF) propagation, microwave transmission, satellite transmission, infrared wave transmission, Bluetooth*, and *Wi-Fi*.

RF Propagation: In RF propagation, the signal is carried over carrier waves (waves which carry signals over them), which have frequencies in the range of RF spectrum. There are three types of RF propagation, namely, *ground wave, ionospheric, and line of sight*.

Ground wave propagation follows the curvature of the earth. They have carrier frequencies of up to 2 MHz. AM radio is an example of ground wave propagation (Figure 16.10).

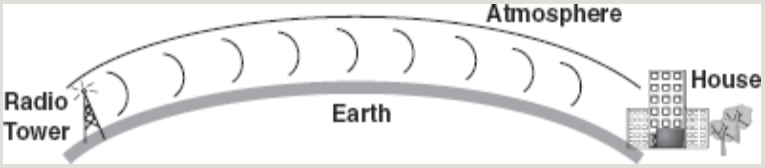


Figure 16.10 Ground Wave Propagation

In ionospheric propagation, the signal wave bounces off the earth's ionosphere layer in the upper atmosphere. It operates in the frequency range of 30–85 MHz. As this type of propagation depends on the earth's ionosphere, it changes with the day timings and weather (Figure 16.11).

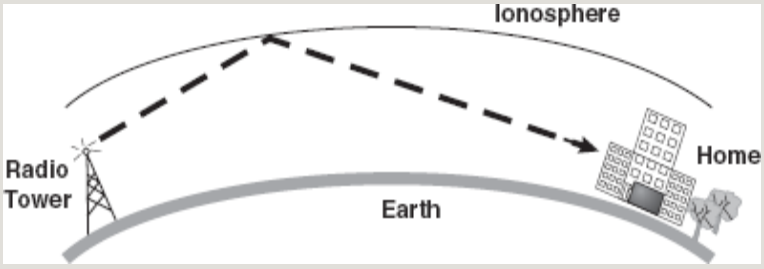


Figure 16.11 Ionospheric Propagation

Line of sight propagation transmits exactly in the line of sight. The receiving station must be in view of the transmitting station. It is sometimes called *space waves* or *tropospheric propagation*. It is limited by the curvature of the earth for ground-based stations (50 km) (Figure 16.12).

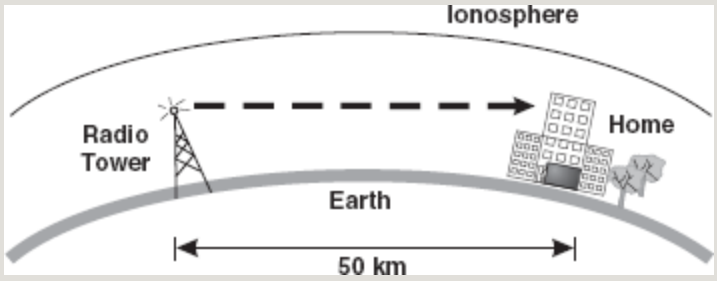


Figure 16.12 Line of Sight Propagation

Microwave Transmission: Microwave transmission use microwave signals that travel at a higher frequency and are popularly used for transmitting data over long

distances. Microwave communication uses parabolic antennas (satellite TV dish antennas) that are mounted on the tower, which could be tens of kilometers away, but in line of sight. Due to long distances, repeaters are used at regular intervals in between the transmitter and the receiver to boost the signals. Normally, the distance between the repeaters goes up roughly with the square root of the tower height. For example, for a 100 m high tower, repeaters can be spaced at 80 km apart (Figure 16.13).

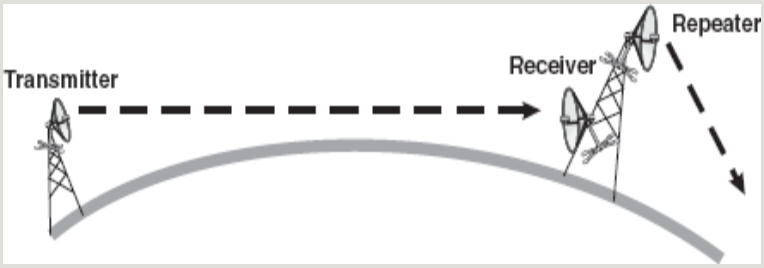


Figure 16.13 Microwave Transmission

Satellite Transmission: Satellite transmission is also a kind of line of sight transmission. Satellites are set in geostationary orbits directly over the equator, which rotates in synchronization to earth, hence they look stationary from any point on earth. These geostationary orbits are placed 36,000 km above the earth's surface. The communication is carried through uplinks and downlinks. The uplink transmits the data to the satellite and downlink receives the data from the satellite. Uplinks and downlinks are also called *earth stations* because they are located on the earth. The area shadowed by the satellite in which the information or

data can be transmitted and received is called the *footprint* (Figure 16.14).

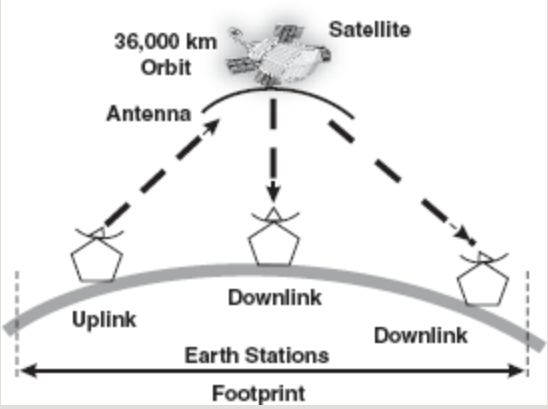


Figure 16.14 Satellite Transmission

Infrared Wave Transmission: Infrared frequencies are just below the visible light spectrum. These are the high-frequency waves used for short-range communication. The waves are cheap, directional, and can be easily built; however, the waves do not pass through solid objects. Infrared waves are used in TV remotes, garage doors, and wireless speakers. Computer machines in an office can easily be connected in a wireless local area network (LAN) using infrared waves (Figure 16.15).

Figure 16.15 Infrared Wave Transmission

Bluetooth: Bluetooth system is named after a Danish king Harold Blatand (Bluetooth in English) who is known for his contribution to unification of tribes from Denmark, Norway, and Sweden. The Bluetooth technology is used to create a personal area network (PAN) or piconet that lets electronic devices, such as computers, cell phones, PDAs, wireless keyboards, etc., to communicate with each other within the range of 10 m. It uses radio signals that are omni-directional, that is, signals travel in all directions from the source. Hence, the communicating devices need not be aligned with each other.

When Bluetooth-capable devices come within range of one another, an electronic conversation takes place to determine whether they have data to share or whether one needs to control the other. The user does not have to manually operate on the devices, as connections are set automatically. Once the conversation has occurred, devices form a piconet. Once the piconet is established, the member devices randomly hop frequencies in unison

so they stay in touch with one another and avoid other piconets that may be operating in the same room.

The key features of Bluetooth are its roubtness, low complexity, and low power requirements that make it suitable for handheld devices like cell phones, PDAs, etc. Today, almost all the small handheld devices are Bluetooth enabled. The Bluetooth radio-link uses a fast acknowledgement and frequency-hopping scheme to make. However, the maximum data transmission rate that it supports is only 1 Mbps, which makes it unsuitable for exchanging large files or folders. Moreover, the maximum range for this technology is 30 feet, which limits the space of PAN and the connection accessibility for electronic devices to perform an action (Figure 16.16).

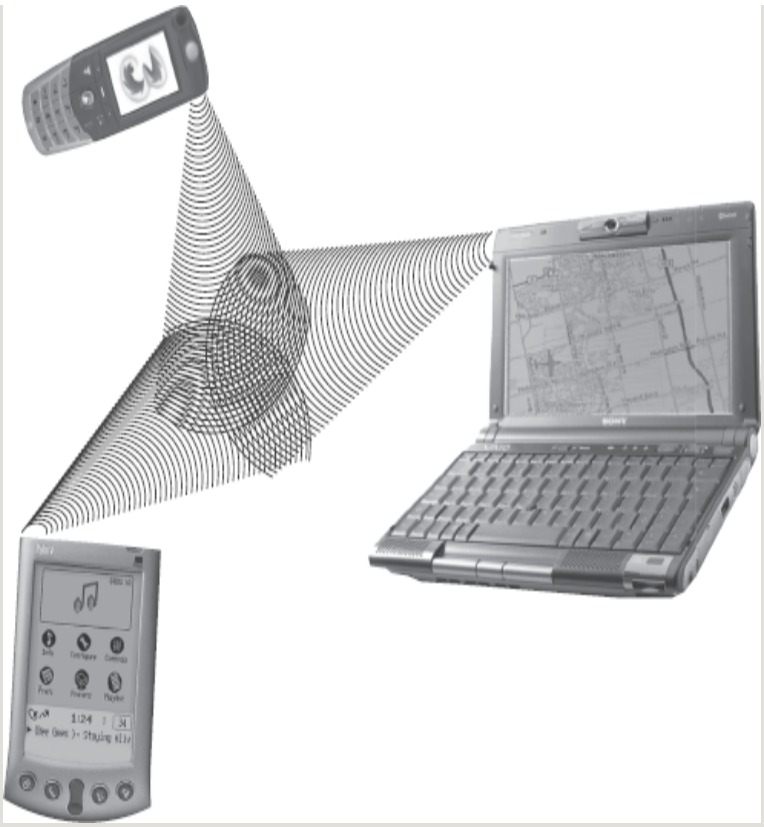


Figure 16.16 Bluetooth Communication

Wi-Fi: The term Wi-Fi which stands for *Wireless Fidelity* is widely used for creating wireless LANs and also for providing wireless Internet access. To create wireless LAN, the devices in the network, such as computer, laptop, PDA, etc., should be equipped with wireless network interface card (NIC) to send and receive Wi-Fi signals. A Wi-Fi-enabled device should be within a specific range (that is, up to 50 m) from the network to connect to it.

The major benefit of Wi-Fi is that it allows developing low-cost wireless networks by reducing the cost of wires for connecting devices. Elimination of wires enables the wireless devices to move from one location to another

(within the range) without losing the connectivity. Though Wi-Fi enables availing network services at low cost, it has some disadvantages also.

The weak security is one of the disadvantages of Wi-Fi that makes the wireless network vulnerable to attack from the intruders. Wi-Fi signals are not confined to the boundaries of building (in which the wireless LAN is operating); thus any intruder can easily gain access to the network by standing outside the building Another disadvantage of Wi-Fi network is its susceptibility to interference from other devices operating in the same area such as wireless phones, other wireless networks, etc (Figure 16.17).

Figure 16.17 Wi-Fi

16.4 MODULATION

Consider a scenario where a boy standing over the roof of his house has to deliver a paper to the boy standing over the roof of the house in front. Due to the distance between the two roofs, the first boy cannot just hand over the paper to the other. In addition, paper alone cannot travel the distance (if thrown) because it is lightweight. Therefore, the first boy wraps the paper onto a stone and throws it towards the other boy, who catches the stone wrapped in the paper. The boy takes off the paper from the stone, reads the message, and discards

the stone. In this manner, the first boy is able to deliver a message to the other boy. In the same way, before a signal is transmitted in a wide communication system, the signals (paper) are superimposed on a carrier signal (stone), which propagates by means of an electromagnetic wave. This process is called *modulation*. Modulation is the addition of information (or the signal) to a signal carrier wave. These carrier waves carry the signals to travel over long distances. Generally, there are three forms of modulation: *amplitude, frequency*, and *phase*.

16.4.1 Amplitude Modulation (AM)

In this modulation, the amplitude of a carrier wave is varied in accordance with the characteristic of the modulating signal. The frequency of the carrier remains the same, only the amplitude changes to follow variations in the signal. In simpler words, the two discrete binary digits (0 and 1) are represented by two different amplitudes of the carrier signal. Figure 16.18 depicts how the modulating signal is superimposed over the carrier signal that results in an amplitude-modulated signal.

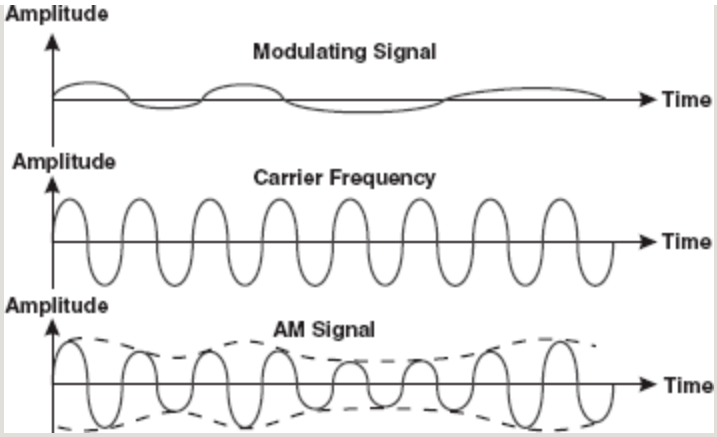


Figure 16.18 Amplitude Modulation

16.4.2 Frequency Modulation (FM)

In this modulation, the instantaneous frequency of carrier wave is caused to depart from the centre frequency by an amount proportional to the instantaneous value of the modulating signal. In simple words, FM is the method of impressing modulating signal onto a carrier signal wave by varying its instantaneous frequency rather than its amplitude (see Figure 16.19).

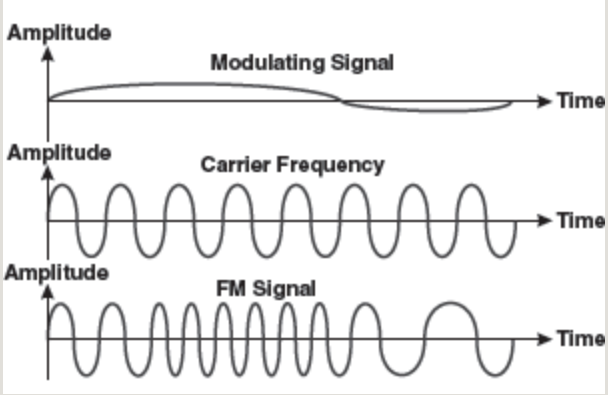


Figure 16.19 Frequency Modulation

16.4.3 Phase Modulation (PM)

PM is the encoding of information into a carrier wave by variation of its phase in accordance with an input signal. In this modulation technique, the phase of sine wave carrier is modifed according to the amplitude of the message to be transmitted. This technique is also called *phase shift keying* (*PSK*), which refers to the simple case of PM by a simple signal with only two states (Figure 16.20).

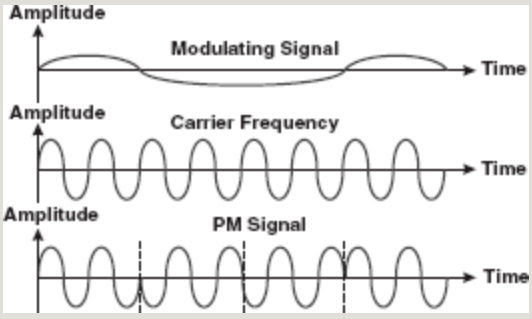


Figure 16.20 Phase Modulation

In a network environment, it is common that the transmission capacity of a medium linking two devices is greater than the transmission needs of the connected devices. Hence, the medium can be shared so that it can be used fully. This can be done by multiplexing. *Multiplexing* is a technique used for transmitting several signals simultaneously over a single communication channel. An analogy of multiplexing can be made with a multilane highway. Just as a multilane highway can carry increased volumes of traffic in multiple lanes at higher speeds and at relatively low incremental cost per lane, higher-capacity circuit can carry multiple conversations in multiple channels at relatively low incremental cost per channel.

16.5.1 Multiplexers

In a multiplexed system, several devices share the capacity of outputs single communication link called *common medium*. The communication device (on the sender's side) that multiplexes (combines) several signals into one data communication signal to be transmitted over the common medium is called a *multiplexer* (*MUX*). At the receiving end, a *demultiplexer* (*DEMUX*) completes the communication process by separating multiplexed signals from a transmission line and distributing it to the intended receiver (Figure 16.21).

Figure 16.21 A Multiplexed System

Signals are multiplexed using two basic techniques: *frequency-division multiplexing* (*FDM*) and *time division multiplexing* (*TDM*).

FDM: FDM is used when the bandwidth of the transmission medium between the MUX and DEMUX is much greater than the requirements from any one stream being multiplexed. FDM is usually used for broadband analog transmissions. In this technique, signals from each sending device are modulated using carrier waves with different frequencies. These modulated signals are then combined into a single composite signal that is transported over the common medium. Each modulated signal acts as a channel. One of the most common examples of FDM is the transmission of radio stations. Figure 16.22 gives a conceptual view of FDM.

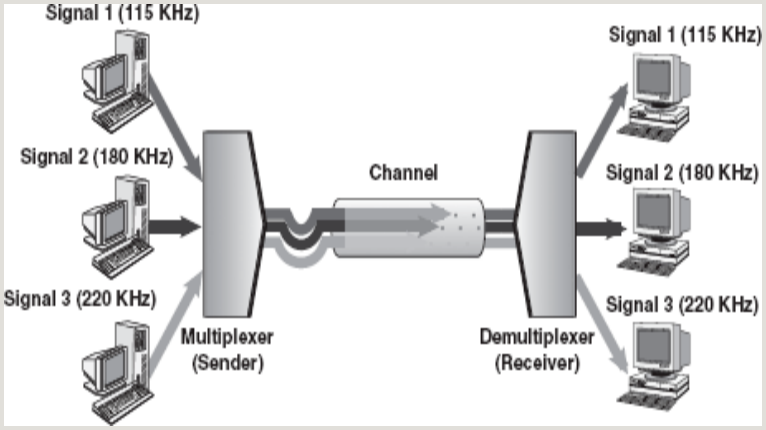


Figure 16.22 Frequency Division Multiplexing

TDM: TDM divides the main signal into time slots, with each time slot carrying a separate signal. It is used for digital communication and can be applied when the data rate capacity of the transmission medium is greater than the data rate required by the sending and receiving devices. Time division MUX allocates each input channel a period of time or time slot. Figure 16.23 gives a conceptual view of TDM. Each sending device is assigned the transmission path for a predefined time slot. Three sending signals, Signal 1, Signal 2, and Signal 3, occupy the transmission sequentially. As shown in the Figure 16.23, time slots A, B, P, Q, X, and Y follow one after the other to carry signals from the three sources, which upon reaching the DEMUX are sent to the intended receiver.

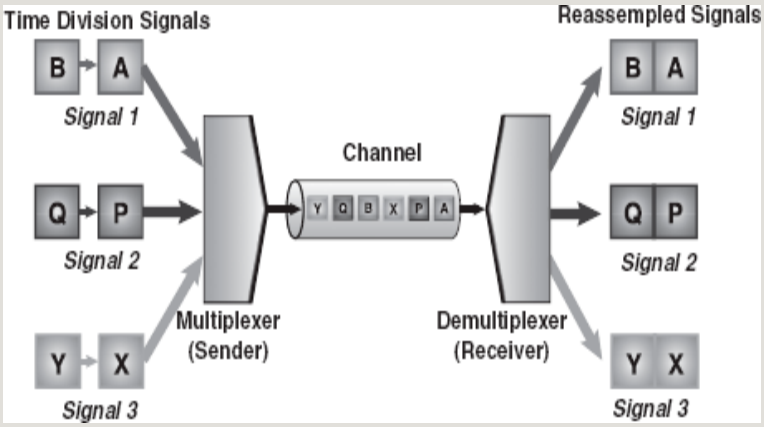


Figure 16.23 Time Division Multiplexing

16.5.2 Asynchronous and Synchronous Transmission

The primary concern while considering the transmission of data from one device to another is to decide whether to send the data one bit at a time (serial mode) or to send a group of bits into a large group (parallel mode). In serial mode, one bit is sent at a given instant; hence, only one communicating channel is needed to transmit data.

Since only one communication channel is required, the cost is largely reduced. On the other hand, in parallel mode, multiple bits are sent at any given instance; therefore, more than one channel is needed to transmit data. There is only one way of sending data in parallel mode, but there are two subclasses of serial transmission, namely, *asynchronous* and *synchronous* transmission.

Asynchronous Transmission: Asynchronous transmission is so called because the timing of the signal is not important. The information that is received or transmitted follows a predefined pattern. As long as the patterns are followed, the receiving device can retrieve the information without any regard to the timing of the signal sent. However, a synchronizing pulse is necessary for the receiver to know when the data is coming and when it is ending. Hence, each byte of information is preceded by a start bit (denoted by 0) and ended by a stop bit (denoted by 1). Therefore, the information in one byte, that is, eight bits becomes ten bits, increasing the overheads. In addition, the transmission of each byte may be followed by a gap of varying duration, which can further help in synchronizing the information with the data stream or channel. As soon as receiver detects the stop bit, it ignores any received pulses until it detects the next start bit. The asynchronous transmission is slower than the other forms of transmission but at the same time is cheaper and an attractive choice for low-speed communication (Figure 16.24).

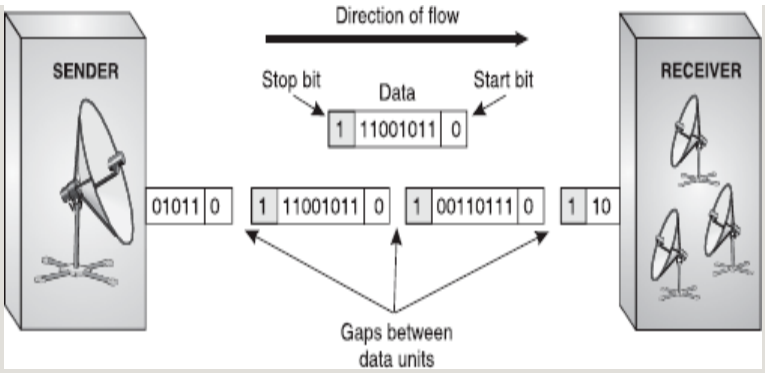


Figure 16.24 Asynchronous Transmission

Synchronous Transmission: Synchronous mode of transmission works on the same media as the asynchronous transmission, but the transmitter does not send start and stop bits to the receiver. The receiver's clock is synchronized with the transmitter's clock. In other words, data is transmitted as an unbroken string of 1s and 0s, and the receiver, on the basis of clock timings, separate the string into bytes. Timing becomes very important in synchronous transmission because without start and stop bits, there is no in-built mechanism to help the receiving device access the incoming information. The advantage of synchronous transmission is speed. With no extra start and stop bits, overhead is lessened, increasing the speed of transmission. Therefore, synchronous transmission is useful for high-speed application like transfer of large data from one computer to another (Figure 16.25).

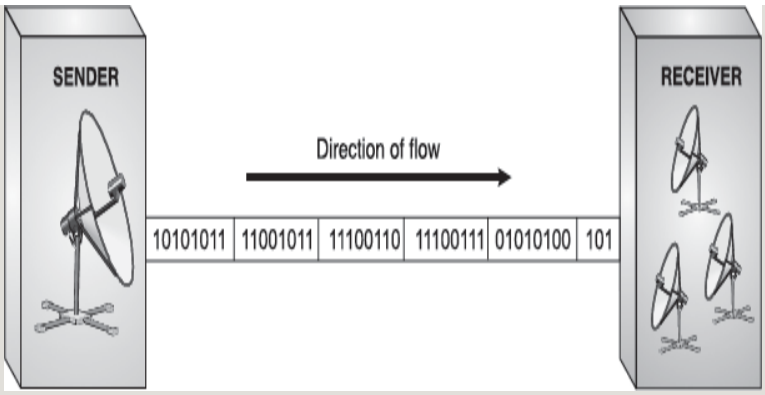


Figure 16.25 Synchronous Transmission

16.6 SWITCHING

Consider a scenario of a small office having four telephone sets used by the four employees for communication. If direct lines were to be used for all the people, six duplex lines are required $n(n - 1)/2$ lines, where n is the number of telephone sets. This is called *point-to-point connection* (Figure 16.26).

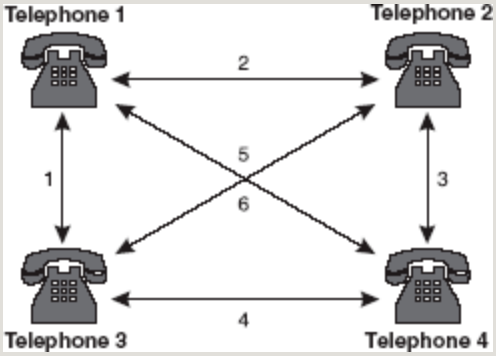


Figure 16.26 Point-to-Point Connection

This method, however, is impractical and wasteful when applied to very large networks. The number and length of the links require too Telephone 3 much of infrastructure; in addition, majority of these links would remain idle and wasted most of the time. A better solution is switching. On a network, *switching* means routing traffic by setting up temporary connections between two or more network points. This is done by devices located at different locations on the network, called switches (or *exchanges*). In a switched network, some switches are directly connected to the communicating devices while others are used for routing or forwarding information.

Figure 16.27 depicts a switched network in which communicating devices are labelled A, B, C, and so on and switches are labelled I, II, III, I V, and so on. Each switch is connected either to a communicating device or to any other switch for forwarding information. Notice that multiple switches are used to complete the connection between any two communicating devices at a time, hence saving the extra links required in case of a point-to-point connection.

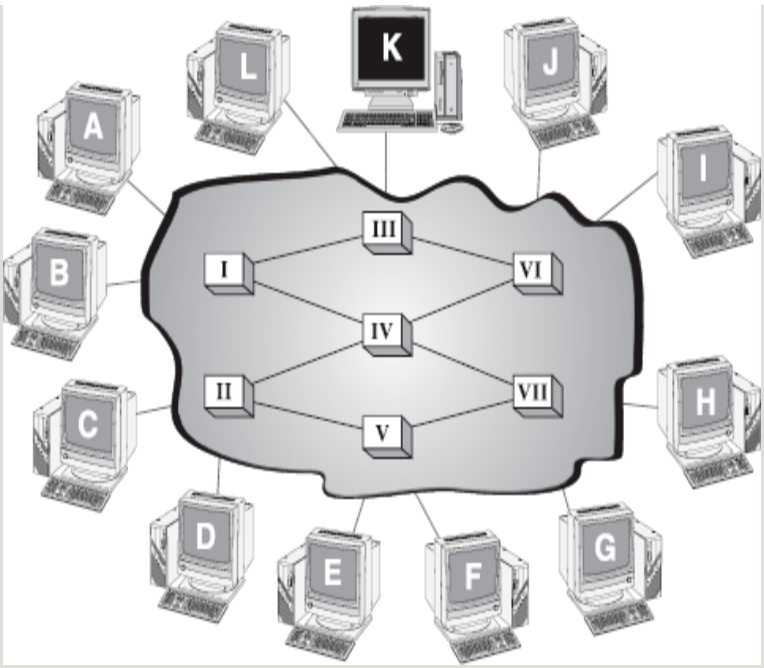


Figure 16.27 Switched Network

Switching traditionally employs three methods: *circuit switching*, *packet switching*, and *message switching*. Out of these, only circuit and packet switching are in use nowadays, message switching has been phased out in general communications.

16.6.1 Circuit Switching

When a device wants to communicate with another device, circuit switching technique creates a fixed-bandwidth channel, called a *circuit*, between the source and the destination. This circuit is reserved exclusively for a particular information flow, and no other flow can use it. Other circuits are isolated from each other, and thus their environment is well controlled.

THINGS TO REMEMBER

Circuit Switching

Circuit switching is plagued by many disadvantages. When there is no flow of data, the capacity of incurring link is wasted. Circuit switching is less suited to data communication, where the data come in surges with idle gaps between them. It is inflexible as well, that is, once the circuit is established, that circuit is the only path taken by the flow of information whether or not it remains the most efficient and available.

As illustrated in Figure 16.28, if device A wants to communicate with device D, sets of resources (switches I, II, and III) are allocated which act as a circuit for the communication to take place. The path taken by data between its source and destination is determined by the circuit on which it is flowing and does not change during the lifetime of the connection. The circuit is terminated when the connection is closed. Therefore, this method is called *circuit switching*. A common example of a circuit switched network is *public switched telephone network* (*PSTN*).

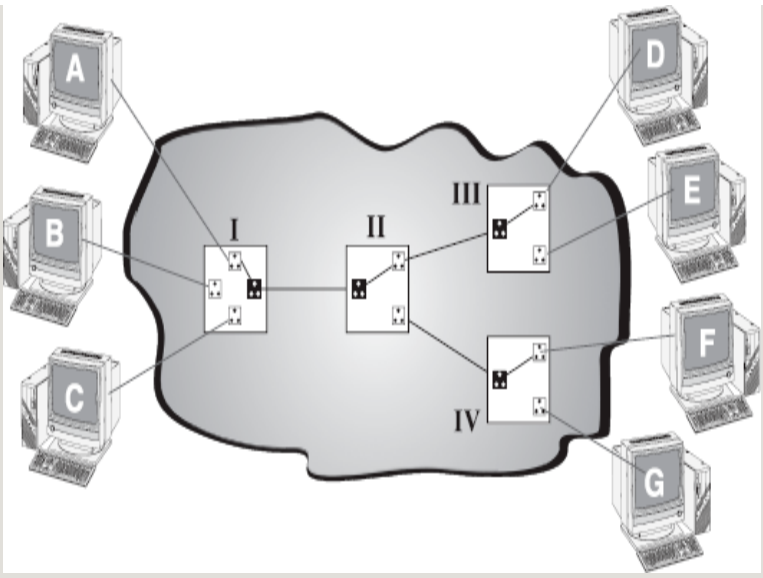


Figure 16.28 Circuit Switching

In circuit switching, data is transmitted with no delay (except for negligible propagation delay). This method is simple and requires no special facilities. Therefore, it is well suited for low-speed data transmission.

16.6.2 Packet Switching

Circuit switching was designed for voice communication. For example, in voice communication such as a telephonic conversation, once a circuit is established it remains busy for the duration of the conversation session. However, circuit switching is marred with many limitations such as in case of idle time (that is, if no data is communicated), the capacity of the link is wasted. This situation is addressed by packet switching.

Packet switching introduces the idea of breaking data into packets, which are discrete units of potentially variable length blocks of data. Apart from data, these packets also contain a header with control information

like the destination address, priority of the message, and so on. These packets are passed by the source point to its local *packet switching exchange* (*PSE*). When the PSE receives a packet, it inspects the destination address contained in the packet. Each PSE contains a navigation directory specifying the outgoing links to be used for each network address. On receipt of each packet, the PSE examines the packet header information and then either removes the header or forwards the packet to another system. If the channel is not free, then the packet is placed in a queue until the channel becomes free. As each packet is received at each transitional PSE along the route, it is forwarded on the appropriate link mixed with other packets. At the destination PSE, the packet is finally passed to its destination. Note that not all packets of the same message, travelling between the same two points, will necessarily follow the same route. Therefore, after reaching their destination, each packet is put into order by a *packet assembler and disassembler* (*PAD*). Figure 16.29 shows that four packets (1, 2, 3, and 4) once divided on machine A are transmitted via various routes, which arrive on the destination machine D in an unordered manner. The destination machine then assembles the arrived packets in order and retrieves the information.

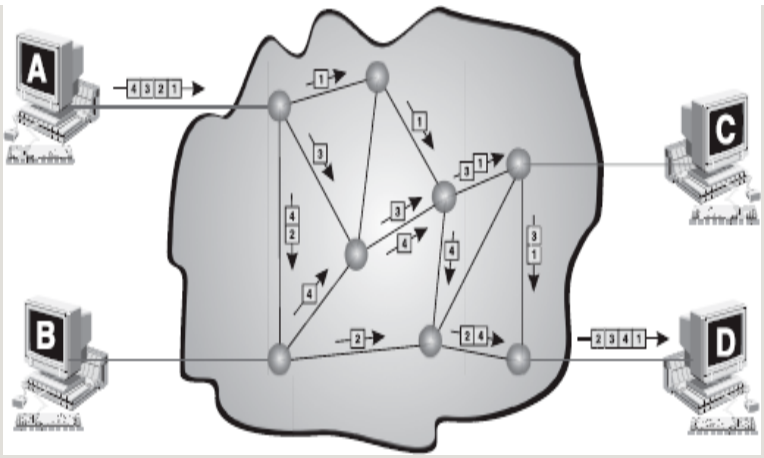


Figure 16.29 Packet Switching

The benefit of packet switching is that since packets are short, the communication links between the nodes are only allocated to transferring a single message for a short period while transmitting each packet. Longer messages require a series of packets to be sent, but do not require the link to be dedicated between the transmission of each packet. This also allows packets belonging to other messages to be sent between the packets of the original message. Hence, packet switching provides a much fairer and efficient sharing of the resources. In addition, if no data is available to the sender at some point during a communication, then no packet is transmitted over the network and no resources are wasted. Due to these characteristics, packet switching is widely used in data networks like the Internet.

16.6.3 Message Switching

Message switching technique employs the “store and forward” mechanism. In this mechanism, a special device (usually, a computer system with large storage capacity) in the network receives the message from a

communicating device and stores it into its memory. Then, it finds a free route and sends the stored information to the intended receiver. In this kind of switching, a message is always delivered to one device where it is stored and then rerouted to its destination.

Message switching is one of the earliest types of switching techniques, which was common in 1960s and 1970s. As the delays in such switching are inherent (time delay in storing and forwarding the message) and a large capacity of data storage is required, this technique has virtually become obsolete (Figure 16.30).

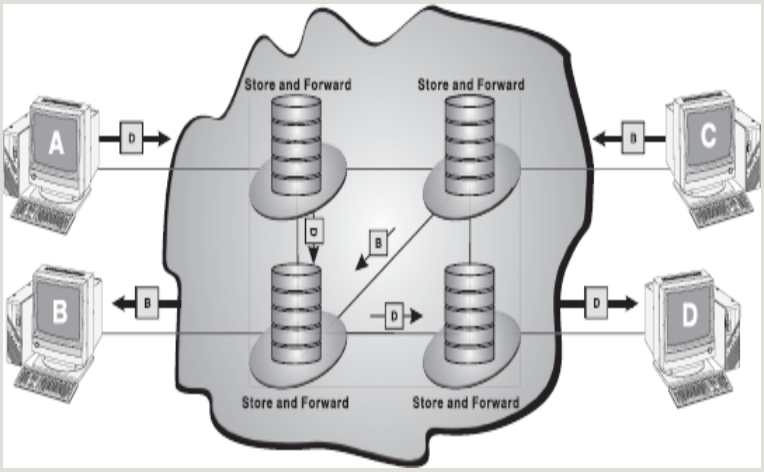


Figure 16.30 Message Switching

16.7 COMPUTER NETWORK

A *computer network* is a collection of two or more computers, which are connected together to share information and resources. The keyword in the definition is “sharing”—the main purpose of computer networking. The ability to share information efficiently is what gives computer networking its power and appeal. Envision a network as a project team. Through the efforts of all

involved—the sharing of time, talent, and resources—a goal is accomplished or a project is completed.

Personal computers are powerful tools that can process and manipulate large amounts of data quickly, but they do not allow users to share that data efficiently. Before networks, users needed either to print documents or to copy files to a disk for others to edit or use them Figure 16.31(a). If others made changes o the document, there was no easy way to merge the changes. This is known as *stand-alone environment*. However, when two computers are linked together using a cable that allows them to share data, it is known as *networked environment*. Observe in Figure 16.31(b), the computers and other devices (like printers) are connected together through a cable. This is called a *network* and the concept of connected computers sharing resources is called *networking*. Note that computers in a network are interconnected by telephone lines, coaxial cables, satellite links, radio, and so on.

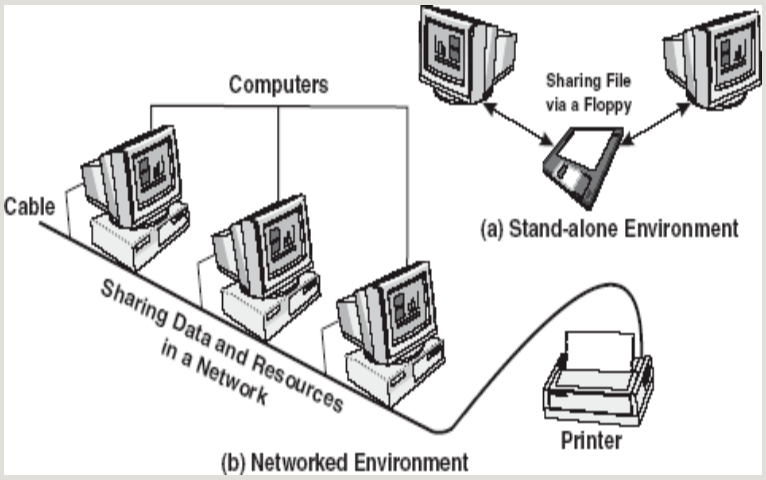


Figure 16.31 Stand-alone and Networked Environment

Technological advances in networking hardware and software have led to greater throughput on all scales and to increasingly tighter integration of networking with all aspects of computing. In tandem with these advances, the idea of networking has entered the common consciousness to an extent that would have been unimaginable a few years ago. This shift in perception has led to an expansion of networking beyond the workplace, which is already beginning to shape developments in networking technology.

Unlike operating system such as DOS that is designed for a single user to control one computer, *network operating system* (*NOS*) is an interconnected system of computing devices that provide shared and economical access to computer services. Thus, NOSs coordinate the activities of multiple computers across a network. NOSs are broadly divided into two types namely, *client/server networks* and *peer-to-peer (P2P) networks*.

Client/Server Network: In this network, each computer is either a client or a server. To complete a particular task, there exists a centralized host computer known as *server* and a user's individual workstation known as *client*. In simple words, client requests a service from the server and the server responds by providing that service. Note that a client/server network centralizes functions and applications in one or more dedicated file servers.

The servers provide access to resources, while the clients have access to the resources available on the servers. A typical example of client–server is sending an e-mail, where the e-mail is first transmitted to the source server. After this, the e-mail is transmitted to the destination server. Lastly, the destination server sends the e-mail to the client at the other end (Figure 16.32).

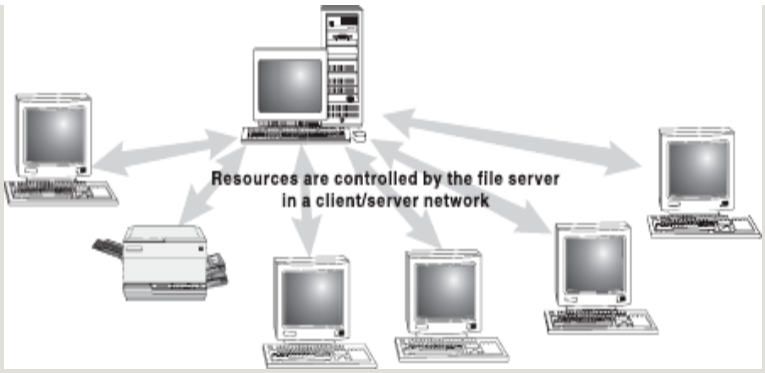


Figure 16.32 Client/Server Network

Generally, the server software runs on a computer, dedicated solely for hosting that software and supporting its services. On the other hand, client software runs on common personal computers or workstations. In a client/server relationship, both the client and server carry out some of the data processing on their ends. In the example, discussed earlier (e-mail) servers are responsible for transmitting the message, while client is responsible for sending and receiving the message.

P2P Network: A *P2P network* does not rely on dedicated servers for communication; instead, it uses direct connections between clients (peers). A pure P2P network does not have the notion of clients or servers, but only equal peer nodes that simultaneously function as both “clients” and “servers” to the other nodes on the network. That is, every node is able to initiate or complete any supported transaction (file transfer) with the other connected node (Figure 16.33).

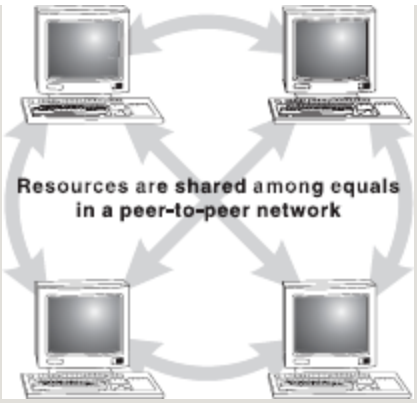


Figure 16.33 P2P Networks

The upper limit of the number of nodes that can function as both clients and servers on a P2P network is between 10 and 25. If there are more nodes, then a P2P machine can be used as a dedicated server with additional high-performance hardware. Note that the peer nodes involved in the network may differ in local configuration, processing speed, network bandwidth, and storage quantity.

P2P networks have been designed primarily for small to medium LANs. Many file-sharing services such as Morpheus and Kaaza are used extensively to find files (sound, images, and videos) on the network. However, most of these file-sharing services actually integrate both P2P and client/server networking design. Therefore, these file-sharing applications are technically known as *hybrid networks*.

Generations in P2P Networks: P2P file-sharing networks can be classified according to their "generations," which are based on the popular Internet-based file-sharing networks. These generations are listed below:

- **First Generation**: This generation had a centralized file list like Napster. In the United States of America, judiciary ruled that whoever controlled this centralized file list was responsible for any infringement of copyright, or any other illegal activities, which might occur while transferring files.
- **Second Generation**: This generation had decentralized file lists such as Gnutella and FastTrack. This was a necessary step for network creators, as a central authority (creators) will be held liable for copyright infringement.
- **Third Generation**: This generation is an improvement upon previous generations. These networks have in-built features such as efficiency, reliability, and anonymity. Examples of anonymous networks are Freenet, I2P, GNUnet, and Entropy. These networks allow only the known users (friends) to connect to your computer. After connecting, each node can forward requests and files anonymously between its own "*friends*" nodes.

Difference between P2P and Client/Server Network: Both P2P and client/server networks have associated advantages and disadvantages. These advantages and disadvantages form a part of distinction between the two. These differences are listed in Table 16.1.

Table 16.1 Difference between Peer-to-Peer Network and Client/Server Network

Basis	P2P Network	Client/Server Network
Client/Server Network	No central repository for files and applications	Resources and data security controlled through the server
Maintenance	Low maintenance cost file	A large network requires extra staff to ensure efficient operation
Installation	Can be easily installed	Requires experts for proper installation of the network
Expensive	No need for a dedicated server, thus not much expensive	Expensive, as it requires a dedicated server
Security	Lack of proper drawback	Provides high level of security
Dependence	All nodes are independent of each other. Failure occurring in one node does not affect the functioning of other nodes in the network	When server goes down, it affects the functioning of the entire network

16.7.1 Types of Computer Networks

A network can be as few as several personal computers on a small network or as large as the Internet, a worldwide network of computers. Today, when talking about networks, we are generally referring to three primary categories: *LAN, metropolitan area network* (*MAN*), and *wide area network* (*WAN*). These categories are defined depending upon various factors like the size of the network, the distance it covers, and the type of link used in interconnection.

LAN: A LAN is a computer network that covers only a small geographical area (usually within a square mile or less) such as an office, home, or building. In a LAN, connected computers have a NOS installed onto them. One computer is designated as the file server, which stores all the software that controls the network. It also stores the software that can be shared by the computers attached to the network. Other computers connected to

the *file server* are called *workstations*. The workstations can be less powerful than the file server, and they may have additional software on their hard drives. On most LANs, cables are used to connect the computers. Generally, LAN offers a bandwidth of 10–100 Mbps (Figure 16.34).

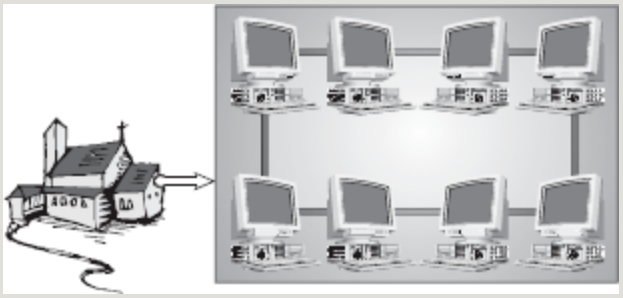


Figure 16.34 Local Area Network

MAN: A MAN is a network of computers spread over a "metropolitan" area such as a city and its suburbs. As the name suggests, this sort of network is usually reserved for metropolitan areas where the city bridges its LANs with a series of backbones, making one large network for the entire city. It may be a single network such as a cable television network or it may be a means of connecting a number of LANs. Note that MAN may be operated by one organization (a corporate with several offices in one city), or be shared and used by several organizations in the same city (Figure 16.35).

Figure 16.35 Metropolitan Area Network

WAN: A WAN is a system of interconnecting many computers over a large geographical area such as cities, states, countries, or even the whole world. These kinds of networks use telephone lines, satellite links, and other long-range communications technologies to connect. Such networks are designed to serve an area of hundreds or thousands of miles such as public and private packet switching networks and national telephone networks. For example, a company with offices in New Delhi, Chennai, and Mumbai may connect the LANs for each of those locations to each other through a WAN. Although a WAN may be owned or rented by private business, it is usually a public network designed to connect small-and intermediate-sized networks together. The largest WAN in existence is the Internet (Figure 16.36).

Figure 16.36 Wide Area Network

WAN offers many advantages to business organizations. Some of them are as follows:

- It offers flexibility of location because not all the people using the same data have to work at the same site.
- Communication between branch offices can be improved using e-mail and file sharing. It facilitates a centralized company-wide data backup system.
- Companies located in a number of small, interrelated offices can store files centrally and access each other's information.

16.8 NETWORK TOPOLOGIES

The term *topology* refers to the way a network is laid out, either physically or logically. A topology can be considered as the network's shape. It is the geometric representation of the relationship of all the links. There are five basic topologies: *Bus, Ring, Star, Tree*, and *Mesh*.

16.8.1 Bus Topology

Bus topology uses a common bus or backbone (a single cable) to connect all devices with terminators at both other. As the message passes each station, the station checks the message's destination address. If the address in the message matches the station's address, the station receives the message. If the addresses do not match, the

bus carries the message to the next station, and so on. Figure 16.37 illustrates how devices such as file servers, workstations, and printers are connected to the linear cable or the backbone.

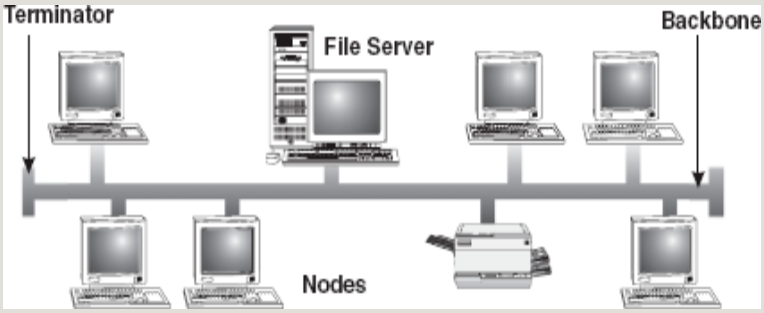


Figure 16.37 Bus Topology

Advantages of Bus Topology: The advantages of bus topology are:

- Connecting a computer or peripheral to a linear bus is easy.
- This topology requires least amount of cabling to connect the computers and, therefore, less expensive than other cabling arrangement.
- It is easy to extend a bus since two cables can be joined into one longer cable with a connector.

Disadvantages of Bus Topology: The disadvantages of bus topology are:

- Entire network shuts down if there is a failure in the backbone.
- Heavy traffic can slow down a bus because computers on such networks do not coordinate with each other to reserve time to transmit.

16.8.2 Ring Topology

In ring topology, computers are placed on a circle of cable without any terminated ends since there are no unconnected ends. Every node has exactly two

neighbours for communication purposes. All messages travel through a ring in the same direction (clockwise or counter-clockwise) until it reaches its destination. Each node in the ring incorporates a repeater. When a node receives a signal intended for another device, its repeater regenerates the bits and passes them along the wire (Figure 16.38).

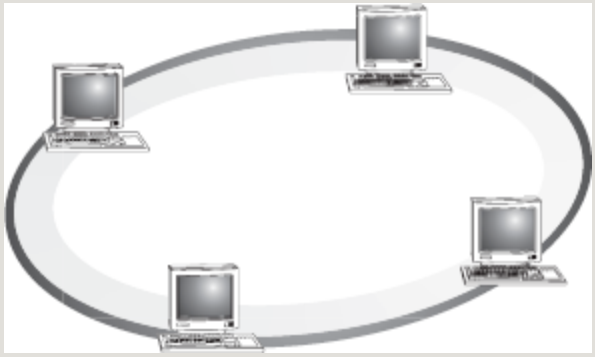


Figure 16.38 Ring Topology

Advantages of Ring Topology: The advantages of ring topology are:

- Ring topology is easy to install and reconfigure.
- Every computer is given equal access to the ring. Hence, no single computer can monopolize the network.

Disadvantages of Ring Topology: The disadvantages of ring topology are:

- Failure in any cable or node breaks the loop and can take down the entire network.
- Maximum ring length and number of nodes are limited.

16.8.3 Star Topology

In star topology, devices are not directly linked to each other, but they are connected via a centralized network component known as *hub* or *concentrator*. The hub acts as a central controller and if a node wants to send data to another node, it boosts up the message and sends the message to the intended node. This topology commonly uses twisted-pair cable; however, coaxial cable or fibre optic cable can also be used Figure 16.39

Figure 16.39 Star Topology

Advantages of Star Topology: The advantages of star topology are:

- Star topology is easy to install and wire.
- The network is not disrupted even if a node fails or is removed from the network.
- Fault detection and removal of faulty parts is easier in star topology.

Disadvantages of Star Topology: The disadvantages of star topology are:

- It requires a longer length of cable.
- If the hub fails, nodes attached to it are disabled.

- The cost of the hub makes the network expensive as compared to bus and ring topologies.

16.8.4 Tree Topology

A tree topology combines the characteristics of linear bus and star topologies. It consists of groups of star-configured workstations connected to a bus backbone cable. Not every node plugs directly to the central hub. The majority of nodes connect to a secondary hub that, in turn, is connected to the central hub. Each secondary hub in this topology functions as the originating point of a branch to which other nodes connect (Figure 16.40).

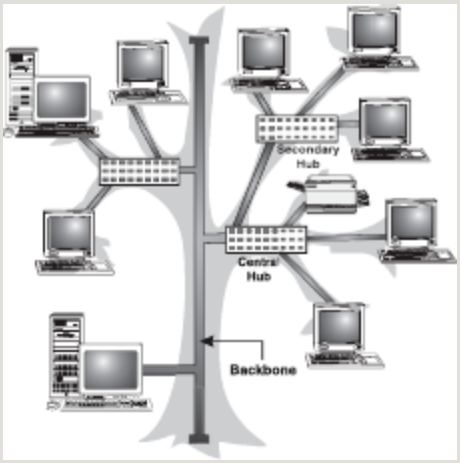


Figure 16.40 tree Topology

Advantages of Tree Topology: The advantages of Tree Topology are:

- The distance to which a signal can travel increases as the signal passes through a chain of hubs.
- Tree topology allows isolating and prioritizing communications from different nodes.
- Tree topology allows for easy expansion of an existing network, which enables organizations to configure a network to meet their needs.

Disadvantages of Tree Topology: The disadvantages of tree topology are:

- If the backbone line breaks, the entire segment goes down.
- It is more difficult to configure and wire than other topologies.

16.8.5 Mesh Topology

In a mesh topology, every node has a dedicated point-to-point link to every other node. Messages sent on a esh network can take any of several possible paths from source to destination. A fully connected mesh network has n(n - 1)/2 physical links to link n devices. For example, if an organization has five nodes and wants to implement a mesh topology, then 5(5 - 1)/2, that is, 10 links are required. In addition, to accommodate those links, every device on the network must have n - 1 communication (input/output) ports (Figure 16.41).

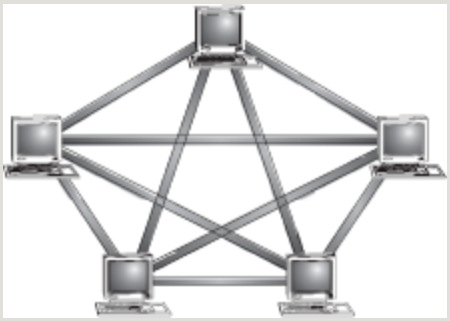


Figure 16.41 Mesh Topology

Advantages of Mesh Topology: The advantages of mesh topology are:

- The use of large number of links eliminates network congestion.
- If one link becomes unusable, it does not disable the entire system.

Disadvantages of Mesh Topology: The disadvantages of mesh topology are:

- The amount of cabling required is very large.
- As every node is connected to the other, installation and reconfiguration is very difficult.
- The amount of hardware required in this type of topology makes it expensive to implement.

16.9 COMMUNICATION PROTOCOL

Imagine yourself standing near a traffic crossing. You can notice that for smooth movement of the traffic, a functioning traffic light is essential. The red signal sends a message to stop, a yellow signal to wait, and a green signal to cross. This set of rules, which tells a driver when to move and when to stop, are traffic protocols. Similarly, computers adhere to certain protocols that define the manner in which communication takes place. A *communication protocol* is a set of rules that coordinate the exchange of information. If one computer is sending information to another and they both follow the same protocol, the message gets through, regardless of what types of machines they are and on what operating systems they are running. As long as machines have software that can manage the protocol, communication is possible.

16.9.1 The OSI Model

Open Systems Interconnectione (*OSI*) is the standard reference model for communication between two end-users in a network. In 1983, International Organization for Standardisation (ISO) published a document called "The Basic Reference Model for Open Systems Interconnection," which visualizes network protocols as a seven-layered model. The model lays a framework for the design of network systems that allow for communication across all types of computer systems. It consists of seven separate but related layers, namely,

physical, data link, network, transport, session, presentation, and *application* layer.

A layer in the OSI model communicates with two other OSI layers, the layer directly above it and the layer directly below it. For example, the data link layer communicates with the network layer and the physical layer. When a message is sent from one machine to another, it travels down the layers on one machine and then up the layers on the other machine. This route is illustrated in Figure 16.42. As the message travels down the first stack, each layer (except the physical layer) adds header information to it. These headers contain control information that are read and processed by the corresponding layer on the receiving stack. At the receiving stack, the process happens in reverse. As the message travels up the other machine, each layer strips off the header added by its peer layer.

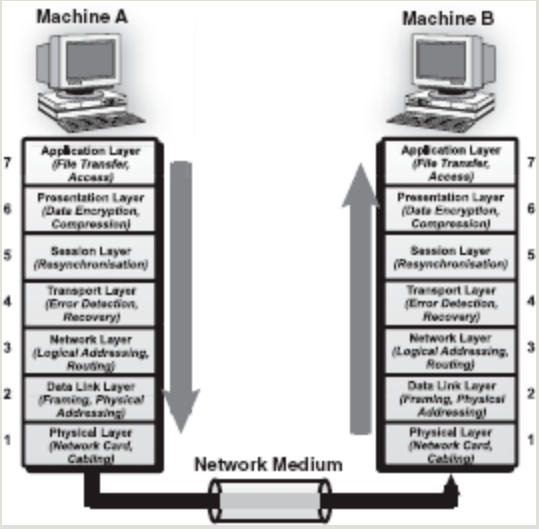


Figure 16.42 Message Transfer in Layered Architecture

Physical Layer: The physical layer defines the physical and electrical characteristics of the network. This layer acts as a conduit between computer's networking hardware and its networking software. It handles the transfer of bits (0s and 1s) from one computer to another. This is where the bits are actually converted into electrical signals that travel across the physical circuit. Physical layer communication media include various types of copper or fibre optic cable, as well as many different wireless medias.

Data Link Layer: The function of the data link layer is to provide service interface to the network layer and is responsible for error-free transmission and preventing the slow receivers from being overflow by the fast senders. To achieve this, the data link layer divides the packets received from the network layer into manageable form known as *frames*. These data frames are then transmitted sequentially to the receiver. On the receiver end, data link layer detects and corrects any errors in the transmitted data, which it gets from the physical layer.

Network Layer: The network layer is responsible for transporting traffic between devices that are not locally attached. For example, a router (a network layer device) provides the routing services in a network. When a packet is received on a router interface, the destination IP address is checked. If the packet is not destined for the router, then the router will look up the destination network address in the routing table. Once an exit interface is chosen, the packet will be sent to the interface to be framed and sent out on the local network. If the entry for the destination network is not found in the routing table, the router drops the packet.

Transport Layer: The basic function of the transport layer is to handle error recognition and recovery of the data packets. This layer establishes, maintains, and

terminates communications between the sender and the receiver. At the receiving end, transport layer rebuilds packets into the original message, and to ensure that the packets arrived correctly, the receiving transport layer sends acknowledgments.

Session Layer: The session layer comes into play primarily at the beginning and end of a transmission. At the beginning of the transmission, it lets the receiver know its intent to start transmission. At the end of the transmission, the session layer determines if the transmission was successful. This layer also manages errors that occur in the upper layers such as a shortage of memory or disk space necessary to complete an operation, or printer errors.

Presentation Layer: The function of presentation layer is to ensure that information sent from the application layer of one system would be readable by the application layer of another system. This is the place where application data is packed or unpacked and is made ready to use by the running application. This layer also manages security issues by providing services such as data encryption and compresses data so that fewer bits need to be transferred on the network.

Application Layer: The application layer is the entrance point that programs use to access the OSI model and utilize network resources. This layer represents the services that directly support applications. This OSI layer is closest to the end-user. Application layer includes network software that directly serves the end-users of the netwrork by providing them user interface and application features like electronic mail, USENET, etc.

Networks are becoming more complicated and more pervasive everyday. Terms like switch, router, gateway, and hub have become part of our everyday language, yet many people do not fully understand the differences between these devices or how to choose the right one for their network. These devices interconnect individual computers and ensure that they communicate efficiently. This section provides some background information on how each device works and gives some guidance in picking the right solution.

NIC: *NIC* is a hardware device that connects clients, servers, and peripherals to the network through a port. Most network interfaces come as small circuit board that can be inserted onto one of the computer motherboard's slots. Alternatively, modern computers sometimes include the network interface as part of their main circuit boards (motherboards). Each network interface is associated with a unique address called its *media access control* (*MAC*) address. The MAC address helps in sending information to its intended destination(Figure 16.43)

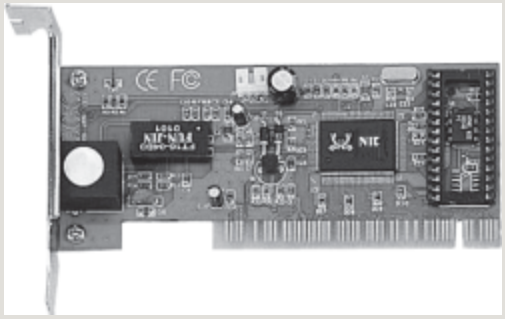


Figure 16.43 Network Interface Card

NICs are a major factor in determining the speed and performance of a network. It is a good idea to use the fastest network card available for the type of workstation one is using.

Repeater: A *repeater* is the most basic device on a network. Signals that carry information within a network can travel a fixed distance before attenuation endangers the integrity of the data. A repeater installed on the link receives signal, regenerates it, and sends the refreshed copy back to the link. Doing this means that the new signal is clean, free from any background noise introduced while travelling down the wire. In Figure 16.44, two sections in a network are connected by the repeater.

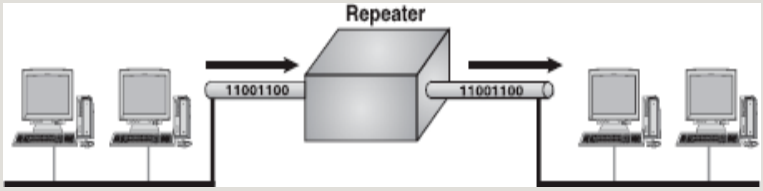


Figure 16.44 Repeater

Repeaters are most commonly used to extend a network cable. All network cable standards have maximum cable length specification. If the distance between two network devices is longer than this specification, a repeater is needed to regenerate the signal. Without the repeater, the signal will be too weak for the computers on each end to reliably understand. A good example of the use of repeaters would be in a LAN using a star topology with unshielded twisted-pair cabling. The length limit for unshielded twisted-pair cable is 100 m. The repeater

amplifes all the signals that pass through it allowing for the total length of cable on the network to exceed the 100 m limit. Nonetheless, repeaters have no in-built intelligence and do not look at the contents of the packet as they regenerate the signal. Thus, there is no processing overhead in sending a packet through a repeater. This also means that a repeater will repeat any errors in the original signal.

Hub: A *hub* is a small box that connects individual devices on a network so that they can communicate with one another. The hub operates by gathering the signals from individual network devices, optionally amplifying the signals, and then sending them onto all other connected devices. Amplification of the signal ensures that devices on the network receive reliable information. A hub can be thought of as the centre of a bicycle wheel, where the spokes (individual computers) meet.

Nowadays, the terms repeater and hub are used synonymously, but they are actually not the same. Although at its very basic level, a hub can be thought of as a multi-port repeater. Typically, hubs have anywhere from 4 to over 400 ports. When a signal is received on one port of the hub, it is regenerated out to all the other ports. It is most commonly used to connect multiple machines to the same LAN. Administrators connect a computer to each port on the hub, leaving one port free to connect to another hub or to a higher-level device like a bridge or router.

Bridge: A *bridge* is a device that allows division of a large network into two or more smaller and efficient networks. It monitors the information traffic on both sides of the network so that it can pass packets of information to the correct location. Most bridges can “listen” to the network and automatically figure out the address of each computer on both sides of the bridge. It

examines each packet as it enters though one of the ports. A bridge first looks at the MAC address of the sender and creates a mapping between the port and the sender's MAC address. It then looks at the address of the recipient, comparing the MAC address to the list of all learned MAC addresses. If the address is in the list, the bridge looks up the port number and forwards the packet to the port where it thinks the recipient is connected. If the recipient's MAC address is not in the list, the bridge then does a flood; it sends the signal to all the ports except the one from where it was received. As a result, a bridge reduces the amount of traffic on a LAN by dividing it into two segments. It inspects incoming traffic and decides whether to forward or discard it (Figure 16.45).

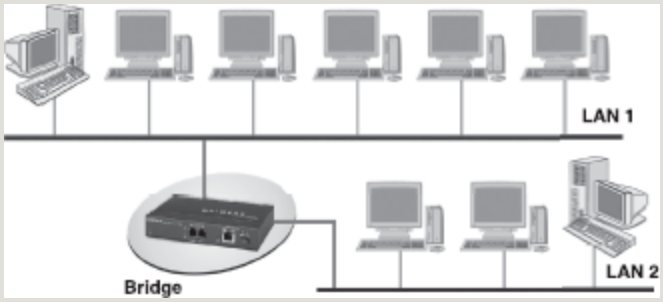


Figure 16.45 Bridge

Bridges can be used to connect networks with different types of cabling or physical topologies. They must, however, be used between networks with the same protocol. Since a bridge examines the packet to record the sender and lookups the recipient, there is overhead in sending a packet through a bridge. On a modern bridge, this overhead is miniscule and does not affect network performance.

Switch: A switch is a multi-port bridge. It connects individual devices on a network so that they can communicate with one another. The behaviour of a switch is the same as that of a bridge. It is capable of inspecting the data packets as they are received, determining the source and destination device of that packet, and forwarding that packet appropriately. The difference is that most switches implement these functions in hardware using a dedicated processor. This makes them much faster than traditional software-based bridges.

Router: A *router* is an essential network device for interconnecting two or more networks. Router's sole aim is to trace the best route for information to travel. As network traffic changes during the day, routers can redirect information to take less congested routes. A router creates and/or maintains a table, called a *routing table* that stores the best routes to certain network destinations. While bridges know the addresses of all computers on each side of the network, routers know the addresses of computers, bridges, and other routers on the network. Routers can even “listen” to the entire network to determine which sections are the busiest—they can then redirect data around those sections until they clear up (Figure 16.46).

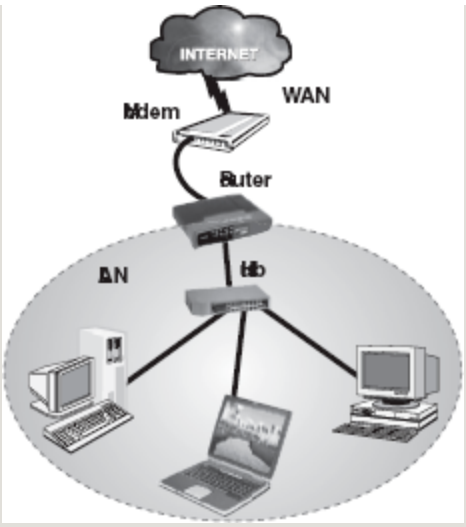


Figure 16.46 Router

Routers are generally expensive and difficult to configure and maintain. They are critical components of a network and if they fail, the network services will be significantly impaired. Most routers operate by examining incoming or outgoing signals for information at the network layer. In addition, they can permit or deny network communications with a particular network.

Gateway: A *gateway* is an internetworking device, which joins networks operating on different protocols together. It is also known as *protocol converter*. A gateway accepts the packet formatted for one protocol and converts the formatted packet into another protocol. It can be implemented completely in software, hardware, or as a combination of both. For example, a gateway can receive e-mail message in one format and convert it into another format. One can connect systems with different protocols, languages, and architecture using a gateway (Figure 16.47).

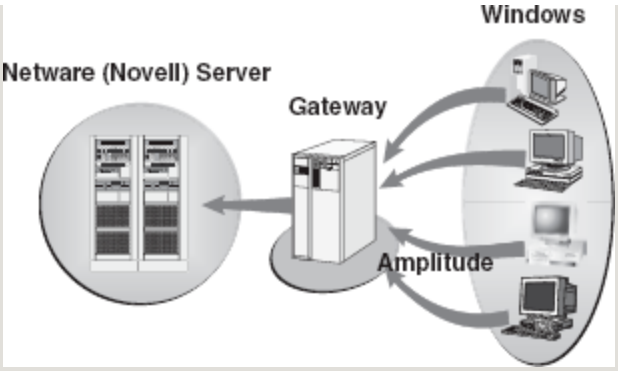


Figure 16.47 Gateway

LET US SUMMARIZE

1. *Data communication* is the exchange of data between two devices using some form of wired or wireless transmission medium. There are five basic components in data communication system: *message, sender, receiver, medium, and protocol*.
2. *Data transmission* mode refers to the direction of signal flow between two linked devices. There are three types of transmission modes: *simplex* (unidirectional data flow), *half-duplex* (bi-directional data flow, but one at a time), and *full-duplex* (simultaneous bi-directional data flow).
3. The data can be transmitted from one point to another by means of electromagnetic signals. These signals may be in one of the two forms: *analog* or *digital*. An analog signal is a continuous waveform that changes smoothly over time. Digital data refer to the data stored in the form of 0s and 1s.
4. *Bandwidth* refers to the maximum volume of data that can be transferred over any communication medium at a given point of time. The more the data needed to be transmitted in a given period, the more is the bandwidth required.
5. *Transmission media* refers to the physical media through which data can be transmitted from one point to another. It can be divided into two broad categories: *guided/wired media* and *unguided/wireless media*.
6. There are three basic *types of guided media: twisted-pair, coaxial cable,* and *optical fibre*. Some commonly used unguided media of transmission include *RF propagation, microwave transmission, satellite transmission, infrared wave transmission, Bluetooth,* and *Wi-Fi*.
7. *Modulation* refers to the process of superimposing the signals on a carrier wave by changing some of the wave's characteristics (such as amplitude, frequency, or phase) so that it is more suitable for transmission over the medium between transmitter and receiver. Generally, there are three forms of modulation: *amplitude, frequency,* and *phase*.
8. *Multiplexing* is a technique used for transmitting several signals simultaneously over a single communication channel. Signals are multiplexed using two basic techniques: *FDM* and *TDM*.

9. *Asynchronous transmission* refers to the data transmission of one character at a time, with intervals of varying lengths between transmittals, and with start bits at the beginning and stop bits at the end of each character, to control the transmission.
10. *Synchronous transmission* is a method of communication in which data are sent in blocks, without the need for start and stop bits between each byte. Synchronization is achieved by sending a clock signal along with the data.
11. *Switching* refers to routing traffic by setting up temporary connections between two or more network points. A temporary connection is achieved by devices located at different locations on the network, called switches. Switching traditionally employs three methods: *circuit switching, packet switching,* and *message switching*.
12. A *computer network* is a collection of two or more computers, which are connected together to share information and resources. A network can be classified into three categories: *LAN, MAN,* and *WAN*.
13. *Topology* refers to the way a network is laid out, either physically or logically. It is the geometric representation of the relationship of all the links. There are five basic topologies: *bus, ring, star, tree,* and *mesh*.
14. *OSI* is a standard reference model for communication between two end-users in a network. It consists of seven separate but related layers, namely, *physical, data link, network, transport, session, presentation,* and *application* layer.
15. To communicate over a network, a particular set of network devices such as *NIC, repeater, hub, bridge, switch, router,* and *gateway* are used. These devices interconnect individual computers and ensure that they communicate efficiently.

EXERCISES

Fill in the Blanks

1. ..... is the exchange of data between two devices using some form of wired or wireless transmission medium.
2. The three types of data transmission modes are....., and .....
3. The maximum volume of data that can be transferred over any communication medium at a given point of time is known as.....
4. In ..... modulation, the amplitude of a carrier wave is varied in accordance with the characteristic of the modulating signal.
5. The communication technique, which combines several signals from different devices and transmits over a common medium, is called .....
6. The switching technique where messages are stored and then forwarded to the nearby communicating device is .....
7. A network spread around Los Angeles and its suburbs can be termed as ..... network.
8. A tree topology combines characteristics of ..... and ..... topologies.
9. The device, which joins networks operating on different protocols together, is .....

10. The device, ftted onto the computer's motherboard, which enables computer to communicate over the network, is .....

Multiple-choice Questions

1. The transmission mode which allows both communicating devices to transmit and receive data simultaneously is .....
    1. Simplex
    2. Full duplex
    3. Half duplex
    4. None of these
2. Which of the following medias does not come under the guided media?
    1. Optical fibres
    2. Coaxial cable
    3. Microwave
    4. Twisted Pair
3. LAN stands for .....
    1. Long area network
    2. Local area network
    3. Local audible network
    4. Limited area network
4. A public switched telephone network (PSTN) uses ..... switching technique.
    1. Packet
    2. Message
    3. Circuit
    4. None of these
5. The network topology in which devices are not linked to each other and where hub acts as a central controller is .....
    1. Mesh topology
    2. Star topology
    3. Ring topology
    4. Tree topology
6. In ring topology, each .....
    1. Node is connected to exactly two others for communication purposes.
    2. Device is connected to a control unit; communication between one device and another must go through the central unit.
    3. Node in the ring incorporates a repeater.
    4. Both (a) and (c).
7. The network device that connects two or more networks together and whose sole aim is to trace the best route for the information to travel is .....
    1. Hub
    2. Gateway
    3. Switch
    4. Router
8. Which of the following cables support the highest bandwidth and faster transmission rate?
    1. Fibre optic cable
    2. Coaxial cable
    3. UTP cable
    4. Twisted-pair cable

9. A computer network, which provides long-distance transmission of data, images, and sound over a large distance, is .....
    1. MAN
    2. WAN
    3. LAN
    4. VAN
10. Identify the odd term out
    1. Router
    2. Switch

State True or False

1. LANs operate within a wide area network.
2. A telephone line can only send analog signals.
3. In a bus topology, the communications channel consists of only one path for a message to travel to the destination.
4. Twisted-pair cable is an example of unguided media for data communication.
5. Hub is also called as protocol converter.
6. Phase modulation is the encoding of information into a carrier wave by variation of its phase in accordance with an input signal.
7. Coaxial cable can be easily tapped.
8. Mesh topology involves minimum cabling.
9. Optical fibre use copper wires for data transmission.
10. In full-duplex transmission, both parties can communicate at the same time.

Descriptive Questions

1. Write down the advantages and disadvantages of optical fibre.
2. Differentiate:
    1. Time-division multiplexing and frequency-division multiplexing
    2. Guided and unguided media
    3. Asynchronous and synchronous transmission
3. What is modulation? List its types.
4. Explain any three guided and unguided data transmission media.
5. What do you mean by computer network? Explain various types of computer networks.
6. Explain three basic switching techniques used in the computer network.
7. With the help of a diagram, explain various network topologies used to connect computers in a network.
8. Describe the following network devices:
    1. Router
    2. Hub
    3. Gateway
    4. Bridge

ANSWERS

Fill in the Blanks

1. Data communication
2. Simplex, half duplex, full duplex
3. Bandwidth
4. Amplitude
5. Multiplexing
6. Message switching
7. Metropolitan area
8. Bus, Star
9. Gateway
10. Network interface card

Multiple-choice Questions

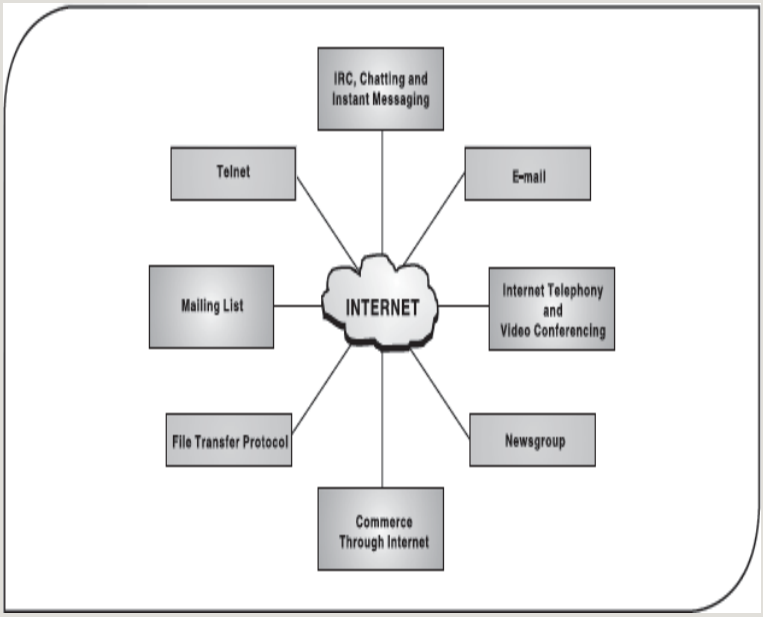
1. (b)
2. (c)
3. (b)
4. (c)
5. (b)
6. (d)
7. (d)
8. (a)
9. (b)
10. (c)

State True or False

1. False
2. True
3. True
4. False
5. False
6. True
7. True
8. False
9. False
10. True

Chapter 17

The Internet



The Internet is the world's largest computer network made up of thousands of networks worldwide. The chapter unravels the mysteries of this widely popular computer concept and helps in building an understanding of the basic Internet structure and processes. The discussion begins with the evolution of the Internet. It is continued with basic Internet terms and how to get connected to it. This chapter also highlights various Internet applications and its languages. The discussion is concluded with the various Internet languages used in the development of Web pages.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- Internet—the world's largest computer network
- The evolution of Internet
- The basic terms of Internet
- The basic requirements for getting connected to Internet
- The various applications of Internet

17.1 INTRODUCTION

In the past two decades, computers have played a significant role in our everyday life. With the growth in popularity and networking of computers, intercommunication between different computers became easier, which led to the evolution of *Internet*. It is the most talked about computer technology in recent years. For a majority of people, who are unaware of computer technology, Internet depicts some enigmatic computer gizmo, which has something to do with games, music, and generally entertainment. However, Internet is a much broader concept than mere entertainment and recreational activities. It is revolutionizing and enhancing the ways of human interaction and communication.

The word Internet is derived from two words: *interconnection* and *networks*. Also referred to as *"the Net,"* Internet is a worldwide system of computer networks, that is, a network of networks, which allows the participants (users) to share information. It consists of thousands of separately administered networks of various sizes and types. Each of these networks comprises tens of thousands of computers. Moreover, the

total number of users of the Internet is known to be in millions. This high level of connectivity encourages an unparalleled degree of communication, resource sharing, and information access. It is probably the most powerful and important technological advancement since the introduction of the desktop computer. In order to benefit from what the net has to offer, a basic understanding of what it is and how it works is helpful.

17.2 EVOLUTION OF INTERNET

The foundation of Internet was laid in 1969 by the Department of Defense (DoD) of United States of America. They wanted to create a computer network that could continue to function in the event of a disaster, such as a nuclear war. Even if a part of the network was damaged or destroyed, the rest of the system would continue to work. That network was known as *ARPANET (Advanced Research Projects Agency Network)*, which linked US scientific and academic researchers. It was the forerunner of today's Internet.

Later in 1980, another agency, the National Science Foundation (NSF), created a new network of computers based on ARPANET, called *NSFNET*, which turned out to be more efficient and capable. Initially, NSFNET was designed to link five super computers situated at the major universities of NSF and allowed only academic research. Over the time, this network expanded to include sites for business, universities, government, etc. and finally becoming a network consisting of millions of computers, now known as the *Internet*.

In the 1980s, Usenet newsgroups and electronic mail (e-mail) came into picture. To keep track of the information on the Internet, indices such as archie and the wide area information server (WAIS) were also created in this decade. To give users an easy-to-use interface to work

with, the University of Minnesota created Gopher, a simple menu-driven system for accessing files.

Internet really became popular in the 1990s after the development of the World Wide Web (WWW). Before that, it was open for a handful of sites only. It is amazing that as recently as June 1993, there were only 130 Web sites, but now there are millions.

With the growth of Internet, the quality, quantity, and variety of information also grew. The Internet today is a repository of every type of information. Nowadays, an Internet user can get all sorts of information ranging from how to add to the design of a functional spaceship to choosing a product for personal use (Figure 17.1).

Figure 17.1 Hypothetical Diagram of the Internet

17.3 BASIC INTERNET TERMS

Now let us familiarize ourselves with some basic Internet terms.

WWW: The *World Wide Web* (abbreviated as the *Web* or *WWW*) is a collection of linked documents or pages, stored on millions of computers and distributed across the world. The concept of the Web began at CERN (the European Center for Nuclear Research), Geneva, Switzerland, in the year 1989. Since then, WWW is the most popularly used Internet subnetwork. One of the main reasons that led to its popularity is that it provides information in multimedia form, that is, in more than one medium such as, text, graphics, video, and audio. Further, it provides a simple and consistent way of accessing the information available on the Internet by using *hypertext* system. In hypertext system, the documents are connected to other related documents on the Internet through links. The Web uses a specific Internet protocol called *HTTP* to support hypertext documents.

Note: *Although the WWW is often referred to as the Internet, they are actually two different concepts. The WWW is one of the most popular services available on Internet.*

Web Page: The hypertext documents on the WWW are known as *Web pages*. A *Web page* is written in a language called *HTML* (*hypertext markup language*) which enables to embed hypertext links (discussed later in this section) in the document. Using these hyperlinks, the user can jump from one Web page to another.

Note: *Web pages are also known as HTML documents.*

Web site: A Web site is a set of related Web pages (linked through hypertext links), published by an organization or an individual. Typically, a Web site contains a home page along with other additional Web pages. Each Web site is accessed by its own address known as *URL* (*uniform resource locator*). All the Web sites on the Internet constitute the WWW.

Home Page: When you browse the WWW, you will often see the term "home page." A home page (also called *index page*) is the starting point or a doorway to the Web site; it refers to the Web page that identifies a Web site and contains the hyperlinks to other Web pages in the Web site. It may contain links to other Web sites also. Like the table of contents of a book, the home page usually provides an overview of what could be found at the Web site. If there is not much information, the home page may be the only page of the Web site. However, usually you will find at least a few other pages.

Browser: A *browser* (short for *Web browser*) is a special software that enables users to read/view Web pages and jump from one Web page to another. It is the software that is needed to find, retrieve, view, and send information over the Internet. The most popular browsers are Microsoft Internet Explorer and Netscape Navigator.

Browsers are of two types:

- **Graphical Browsers:** allow retrieval of text, images, audio, and video. Navigation is accomplished by pointing and clicking with a mouse on highlighted words and graphics. Netscape Navigator, Internet Explorer, and Mozilla are graphical browsers.
- **Text Browsers:** provide access to the Web in text-only mode. Navigation is accomplished by highlighting emphasized words on the screen with the arrow up and down keys, and then pressing the Enter key to follow the link. Lynx is an example of a text-based browser (Figure 17.2).

Figure 17.2 A Web Page on a Graphical Browser

URL: Each Web page has a unique address, called a *URL* (uniform resource locator), which identifies its location on the Internet. Usually, the format of an URL consists of four parts: protocol, name of the Web server (or domain name), path, and filename. Here is an example, *http://www.xyz.com/tutor/start/main.htm*. The structure of this URL is:

- Protocol: http
- Web server name/domain name: www.xyz.com
- Path: tutor/start/
- File Name: main.htm

The first part of the address, the part before the colon, is the protocol. Most of the time you see *http* (hypertext transfer protocol) for accessing a Web page. Apart from http, you could also find other protocols such as *ftp* (file transfer protocol), *news* (news server), and *telnet* (for accessing remote computers). After the protocol, comes the domain name. The protocol and the domain name are separated by colons and slashes (Figure 17.3).

Figure 17.3 Uniform Resource Locator

A domain name is a unique name that identifies a particular Web site and represents the name of the server where the Web pages reside. They are the human-readable addresses and hence are easier to remember. Domain names have two or more parts separated by dots. The letters occurring after the dot (.) identify the domain type. A domain type represents the type of organization or countries to which the Web server (sometimes called *host computer*) belongs.

THINGS TO REMEMBER

Domain Types

Some of the domains that are typical on the Internet are:

com	Commercial and for-profit organizations
edu	Colleges and universities (education providers)
gov	Federal government agencies
mil	US military sites
net	Internet infrastructure and service providers

org	Miscellaneous and non-profit organizations

Dozens of domain types have been assigned to identify and locate files stored on host computers in countries across the world. These are referred to as two-letter Internet country codes and have been standardized by the International Standards Organization (ISO). For example, *uk* for the United Kingdom, *de* for Germany, au for Australia, and in for India. The last part of a URL consists of the path and the file name. Path name specifies the hierarchic location of the said file on the computer. For instance, in *http://www.xyz.com/tutor/start/main.htm, the file main.htm* is located in *start*, which is a subdirectory of tutor.

Hypertext: The *hypertext* is a system that provides a simple and consistent way to organize large data (that includes text, images, pictures videos, etc.) available on the Internet. In hypertext systems, the documents contain links to other related documents on the Internet. These links are called *hyperlinks, hotlink*, or simply *links*. A hyperlink is used to jump to another part of the same page or to load a different Web page. Usually hypertext links appear in a different colour (typically in blue and underlined) from the rest of the text. When the cursor is moved over a text link or over a graphic link, it will change from an arrow to a hand. The hypertext words often provide a hint about the connected page. When the user returns to a page with a link he/she had already visited, the hypertext words will often be in a different colour (usually magenta), indicating that the link has already been used. However, the user can certainly use that link again (Figure 17.4).

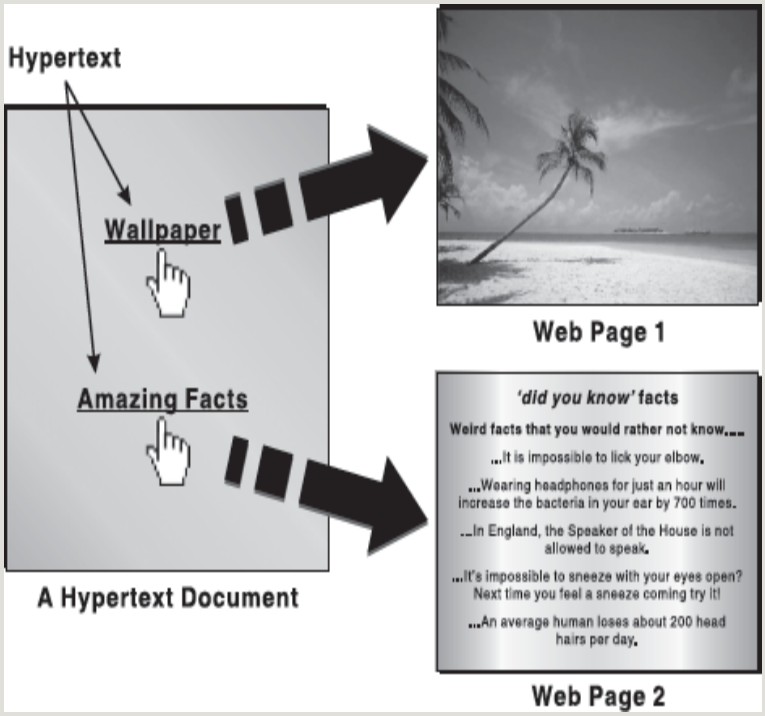


Figure 17.4 Hypertext Links

Hyperlinks are the foundation of the Web and the operation of the Web relies primarily on them as it is a means of information retrieval. Producing hypertext for the Web is accomplished by creating documents with HTML (hypertext markup language).

Internet Service Provider (ISP): An *ISP* is an organization that connects its subscribers' computer using modem to the Internet. The connections can be provided by telephone lines, cable lines, or wireless connections. In order to use the Internet connection, the user has to first log on to the host computer using the username and password provided by the ISP. Some of

the well-known ISPs are America Online (AOL), Microsoft Network (MSN), Yahoo!, etc.

While choosing an ISP, many factors should be kept in mind for subscribing. First, the user should check the speed and consistency of the Internet access. He/she should also check ISP's reliability, that is, whether the ISP provides consistent access to the Internet or does it have frequent downtimes. Apart from these two criteria, the price of the ISP and other additional services provided by it should also be considered. For example, many ISPs provide free Web space so that users can create their own Web site, free e-mail address, and much more. Last but not the least, the user should also check whether the ISP provides a round-the-clock technical assistance.

Web Server: A server is a computer equipped with server software, which provides a specific kind of services to client software running on other computers. It can refer to a particular piece of software, for example, a Web server, or to a machine on which the server software is running. A *Web server* receives request from the clients, processes this request, and sends the result back to the client. In other words, Web server fulfils the client's (browser) requests (Figure 17.5).

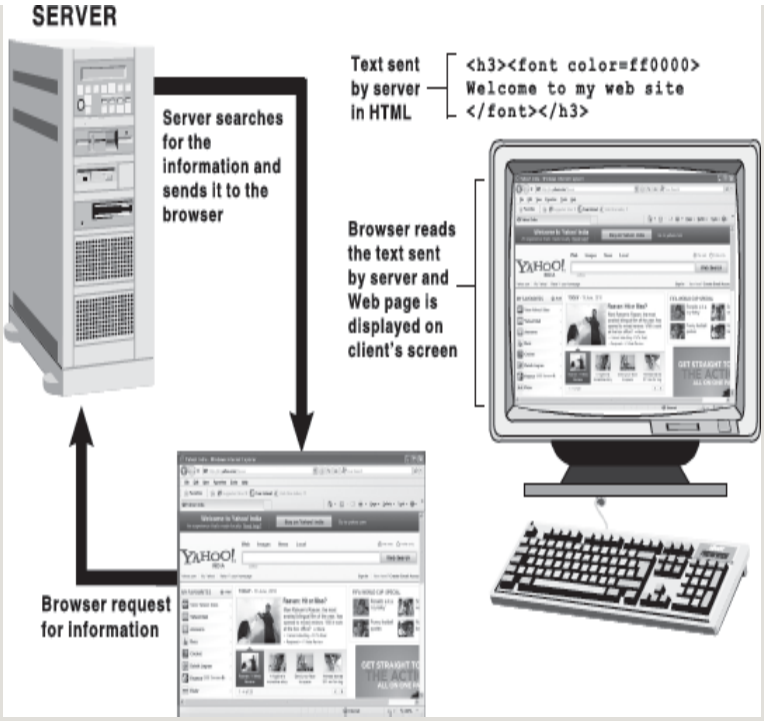


Figure 17.5 Data Transfer from Web Server to Browser

To view a Web site, the browser sends a request to the server. On receiving the request, the server sends the appropriate Web page to the client's machine. The client's machine (browser) receives the information in the form of HTML commands. The browser interprets the HTML, finds all the pictures (or other types of media), and displays the information onto the user's screen.

Download and Upload:Download refers to the act of transmitting data from a remote computer on the Internet or other network to one's own computer. When the user downloads a file, he/she is actually copying a particular file from the remote computer and placing it on his/her hard drive. Downloading allows a user to save

file for later use, thereby reducing the time spent online and the corresponding charges.

Upload is just the opposite of download. *Upload* refers to the act of transmitting data from local computer to any other computer on the Internet or network. For example, a software company may upload a demonstration of its new software onto the Web so that users can then download the software.

Online and Offline: The term *online* is commonly referred to the state of being connected to the networked computer system or Internet. For example, if you are browsing the WWW, your computer is said to be online. Being online also refers to any peripheral device (like printer) that is connected to the system and ready to use.

Offline is just the opposite of online; it refers to the state of not being connected to the remote computer or Internet. A printer or other peripheral that is not ready to use is also considered Offline.

17.4 GETTING CONNECTED TO THE INTERNET

Now you are aware of the basic Internet terms, let us shift our focus on how to connect to the Internet. The basic requirements for getting connected to the Internet (Figure 17.6) are discussed in this section.

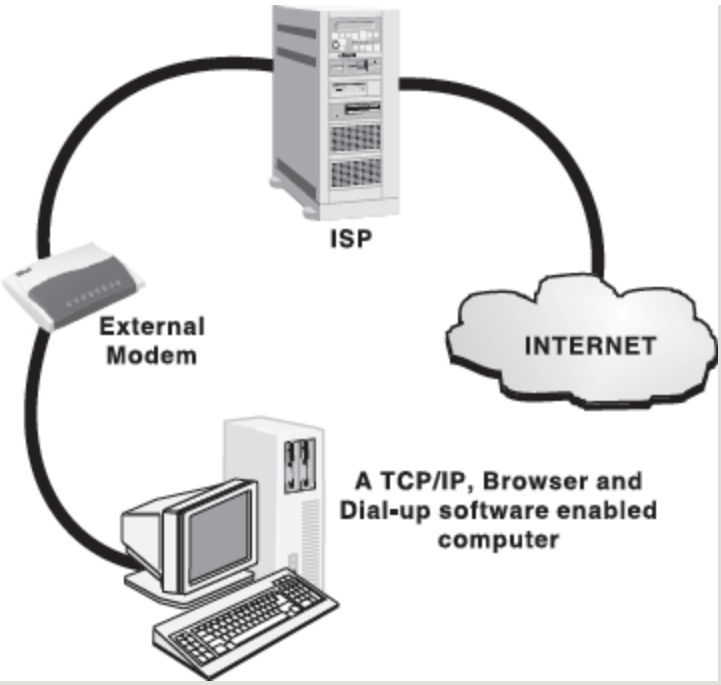


Figure 17.6 Connecting to the Internet

The Computer: The computer must have at least 386-microprocessor chip with a minimum of 16 MB of RAM. For a good browsing experience, use a faster chip (Pentium III/Pentium IV) with more RAM (128/256 MB). In addition, a colour monitor, with at least 640 × 480 resolution and a capability of displaying a minimum of 256 colours, is also vital. The system should also possess a hard disk, with at least 200 MB of free space, to store Internet software and temporary Internet files. Since the Web is a multimedia medium, you can also enjoy sound on your system with a sound card and a pair of speakers.

Modem: Apart from the previously mentioned configuration, a computer system must have a device called *modem*, which enables a computer to transmit data over telephone lines. A *modem* (acronym for

modulator–demodulator) is a hardware, which converts digital data into analog signals (that is, modulation) that can be sent over an analog telephone line and convert the analog signal back into digital data (that is, demodulation). Thus, when the modem receives the data from a Web server via telephone system, it reconstitutes the analog signal into digital form so that the computer can understand it. The most critical aspect of the modem is its speed of operation. The speed of modem is measured in kilobits per second (kbps). Modern modem supports 28 kbps to 56 kbps speeds. Modems are of two types:

- **Internal Modem:** is a card, which is fitted inside the computer with a lead running directly from the computer to the telephone socket.
- **External Modem:** is a small external box wired between the computer and the telephone socket.

Internet Connection: Before connecting to the Internet, you need to buy an Internet connection from an authorized ISP. The speed of accessing the Internet entirely depends upon the type of connection you bought from the ISP. Some of connections available for the users these days are briefed below:

- **Dial-up:** A dial-up connection is the access method that uses telephone lines to connect to the Internet. It is the most common way for individuals who use home computers to connect to the Internet. To connect to the Internet by using dial up, the user needs to specify a username, a password, and a telephone number. After the connection is established, the user can start browsing sites on the Internet.
- **ISDN:** It stands for *integrated services digital network* and is more common in business and commercial use. It was the first high-speed alternative to regular analog telephone modems. ISDN involves the digitization of telephone network so that voice, graphics, text, and other data can be provided to users from a single terminal over existing telephone wiring. Digital connections have fewer errors in transmission, which means speed of downloading graphics, Web pages, sound, and so on is increased to four times faster than with dial-up modems.
- **Cable Modem:** A cable modem connects the user to the Internet through a cable television line. Nowadays, many cable companies offer Internet along with television accessibility. A cable modem will typically have two connections, one to the television outlet and the other to the computer. Cable modems not only provide a faster Internet access (10 to 100 times as fast as the dial-up modem), but interactivity to the television also.

- **Leased Line:** This facility provides reliable, high-speed Internet access ranging from 2.4 kbps to 45 Mbps. A leased line connection is an affordable way to link two or more sites for a fixed monthly charge. Usually leased lines facility can be provided via a fibre optic or copper lines. Leased line service provides a consistent amount of bandwidth and is an excellent way to provide data, voice, and video links between sites.
- **DSL:** DSL or *digital subscriber line* service is provided through an existing telephone line, but it works differently than regular analog modem dial-up access. DSL operates over normal telephone lines and it can be used simultaneously with the telephone. DSL can increase the connection speed by as much as ten-fold from a standard dial-up modem.
- **Broadband:** This type of access is good for remote locations, where ISDN, cable, or DSL are not available. It gives a decent download speed, but to upload data, the user still needs a regular analog modem to dial in, via a telephone line. Satellite connection can be either a two-way service or a one-way service. In case of two-way satellite service, the data are transmitted via satellite to a dish antenna at the user's house. In one-way system, the user needs a conventional modem and telephone link to an ISP. Satellite connection is expensive but sometimes is the only fast option for people who are beyond the service area of cable and DSL providers.

Before deciding the type of Internet connectivity, one should first decide what his requirements are. If the user wants the Internet mainly for sending e-mail, occasional chats, and infrequent browsing, then he should work with regular modem dial-up access. In this case, user should invest in a good-quality 56 kbps modem and a good ISP. If the user is using the Internet frequently for research, downloading or uploading a fair amount of data, to play multi-player games or to use the Internet's multimedia capabilities like live audio or video streaming, then he should look into other high-speed accesses such as cable modem or ISDN.

Internet Software: To connect to the Internet, a computer requires mainly three different kinds of software:

- **TCP/IP:** TCP/IP, or *transmission control protocol/Internet protocol*, is the basic communication protocol of the Internet. It allows programs on user's computer to communicate over the Internet. Usually, when you are set up with direct access to the Internet, your computer is provided with a copy of the TCP/IP. Similarly, every other computer that you may send messages to or get information from also has a copy of TCP/IP.

- **Dialler Software:** This software is provided by the ISP to instruct the modem to dial the telephone number and to identify the user's machine to the access provider's system for access to the network.
- **Browser:** To use the Internet, a Web browser is essential which allows the user to view the information available on the WWW.

17.5 INTERNET APPLICATIONS

Today, the Internet has grown from a small network to biggest network of computers. Due to this colossal growth, the Internet has become the biggest domain of information. Nowadays, an Internet user has access to a wide variety of services such as e-mail, file transfer, vast information resources, interest group membership, interactive collaboration, multimedia displays, real-time broadcasting, shopping opportunities, and much more. Some of the important services provided by the Internet are briefed in the following few sections (Figure 17.7).

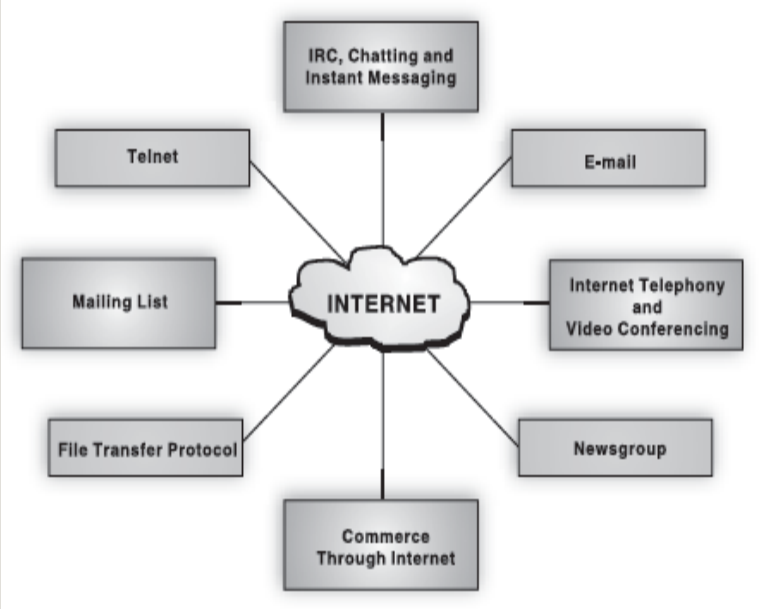


Figure 17.7 Internet Applications

17.5.1 Electronic Mail (E-mail)

Electronic mail, or e-mail, is a fast, easy, and inexpensive way to communicate with other Internet users around the world. It is one of the basic and earliest services of the Internet and the most used application on the Internet too. E-mail overcomes most of the problems and delays of getting a physical document from one person to another. Rather, it has the dual advantage of not only being faster but cheaper to use as well.

Each user of e-mail has a mailbox address to which messages are sent and the mailbox can be accessed from anywhere and at anytime. It does not matter what type of connectivity the user is having, he will always be able to use this service. One can receive and send e-mails from all systems whether it is a Windows PC or a UNIX machine. Not only this, e-mail can also be used to send documents, images, audio, video, and so on as an attachment along with the mail.

17.5.2 File Transfer Protocol (FTP)

The *FTP* is a set of rules that enables a user to transfer files from one system to another. It was the first service developed for the Internet so that government and educational institutions could easily exchange files. It allows the user to get access to the files stored in the directory of a remote computer that is connected to the Internet. Using FTP, one can upload and download files from the remote computer (known as *FTP servers*), if he/she has access permission on the remote machine. Though files can be transferred using e-mail, this is not a good choice if the file size is large or when there are many files. FTP sites can consist of hundreds to thousands of files with information on books, music, software, games, images, sounds, multimedia, course ware, etc. FTP can be done using the command prompt,

browsers, and various GUI-based FTP software such as CuteFTP and WS_FTP (Figure 17.8).

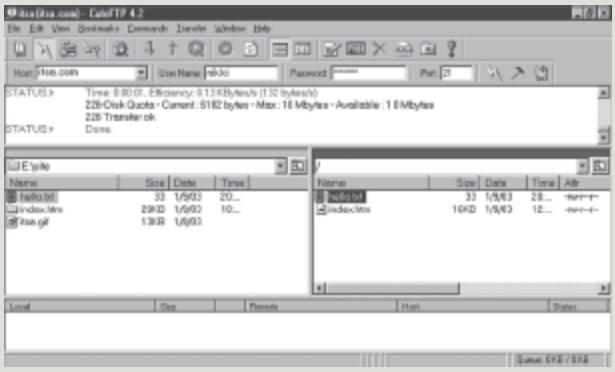


Figure 17.8 FTP Using Graphical FTP Software

17.5.3 Telnet

The word "telnet" is derived from **tel**ecommunications and **net**work and is a protocol that allows a user to log on to a remote computer. Telnet is also known as *remote login*, which means connecting one machine to another in such a way that a person may interact with another machine as if it is being used locally. It means that someone in New Delhi could connect to a computer in the New York City Public Library and search the card catalogue the same way as someone located at a terminal in the library. Once connected, the user's computer emulates the remote computer. When the user types in commands, they are executed on the remote computer. The user's monitor displays what is taking place on the remote computer during the telnet session (Figure 17.9).

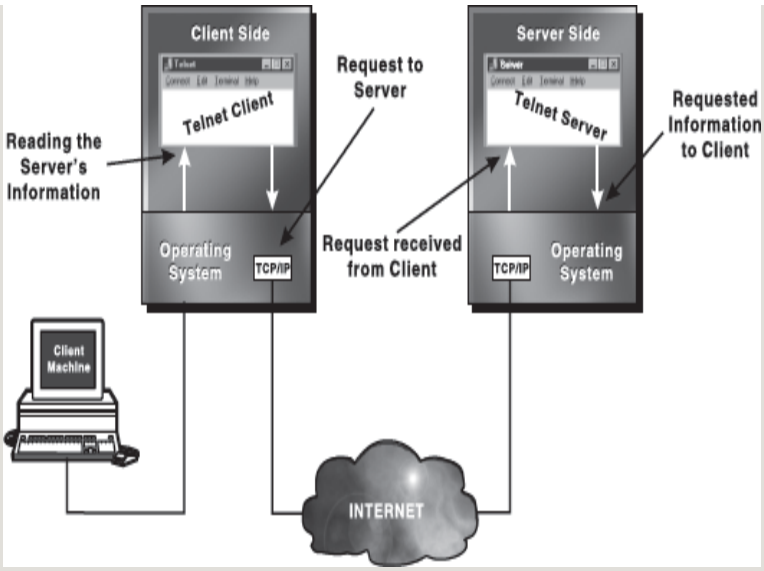


Figure 17.9 Remote Login through Telnet

THINGS TO REMEMBER

Telnet

The telnet client computer must run a telnet client application and the telnet server must run a telnet server application. Moreover, the user's machine has to be connected to the Internet or to a network in order to use telnet. In addition, to transmit information between the telnet client and the telnet server, the TCP/IP protocol is used. Note that unlike FTP, telnet does not have file transfer capability. It can only be used to access information, run programs, and edit files with the help of certain telnet commands.

The user's computer, which initiates the connection, is referred to as the local computer or telnet client, and the machine being connected to, which accepts the

connection, is referred to as the remote computer or telnet server.

17.5.4 Internet Relay Chat (IRC)

IRC or *Internet Relay Chat* is a service on the Internet that allows people to communicate in real time and carry on conversations via the computer with one or more people. It provides the user with the facility to engage in simultaneous (synchronous) online "conversations" with other users from anywhere in the world. Unlike other forms of group communication on the Internet, chat is an instant communication. However, both parties must be connected to the Internet and that too at the same time, running the right software and actively participating in order to chat (Figure 17.10).

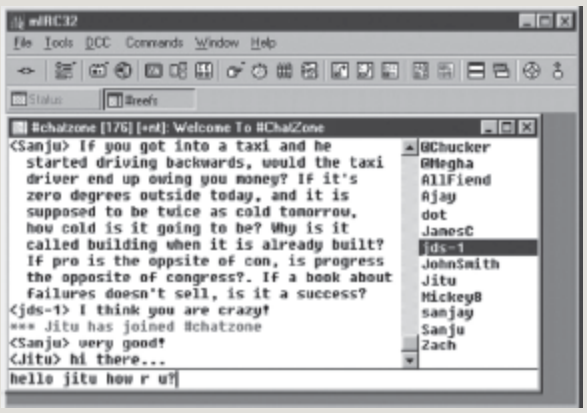


Figure 17.10 Internet Relay Chat

Generally, the user runs a program (called *"client"*) to connect to an IRC server. The server relays information to and from other servers on the same net. Popular IRC client applications are mIRC, Virc, and LeafChat. Once connected to an IRC server on an IRC network, the user

will usually join one or more “channels” (also known as *chat rooms*) and converse with others there. Each channel is devoted to a different topic. Conversations may be public (where everyone in a channel can see what the user types) or private (messages between only two people, who may or may not be on the same channel).

17.5.5 Chatting and Instant Messaging

Chat programs allow users on the Internet to communicate with each other by typing in real time. They are sometimes included as a feature of a Web site, where users can log into chat rooms to exchange comments and information about the topics addressed on the site. For example, America Online is well known for sponsoring a number of topical chat rooms. Chatting has become one of the “killer applications” of Internet and has become quite popular with all kinds of Internet users. It is quite cheap to chat online. Although e-mail is also a cheap source of communication, it is not done on “real time” and the response to the message solely depends on the recipient. Note that even though chatting is based on standardized IRC system, but unlike IRC, the user does not need to have any special software to connect to any chat rooms.

A variation of chat is instant messaging where a user on the Web can contact another user currently logged in and type a conversation. To avail this Internet service, the user must have software called *instant messenger* installed on the system (Figure 17.11). Instant messaging is just as if making a text-only telephone call over the Internet, that is, the user is “talking” to the other person (possibly more) instantly. There are various products out there for instant messaging such as MSN Messenger and Yahoo Messenger.

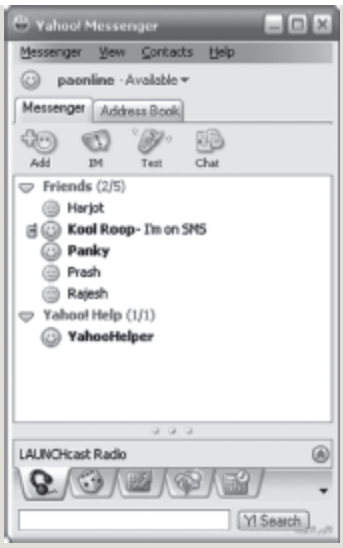


Figure 17.11 Instant Messenger

17.5.6 Internet Telephony

Internet telephony is the use of the Internet rather than the traditional telephone company infrastructure, to exchange spoken or other telephonic information. It consists of hardware and software that enable people to use the Internet as a transmission medium for telephone calls (Figure 17.12).There are many Internet telephony applications available. Some applications such as CoolTalk and NetMeeting come bundled with popular Web browsers.

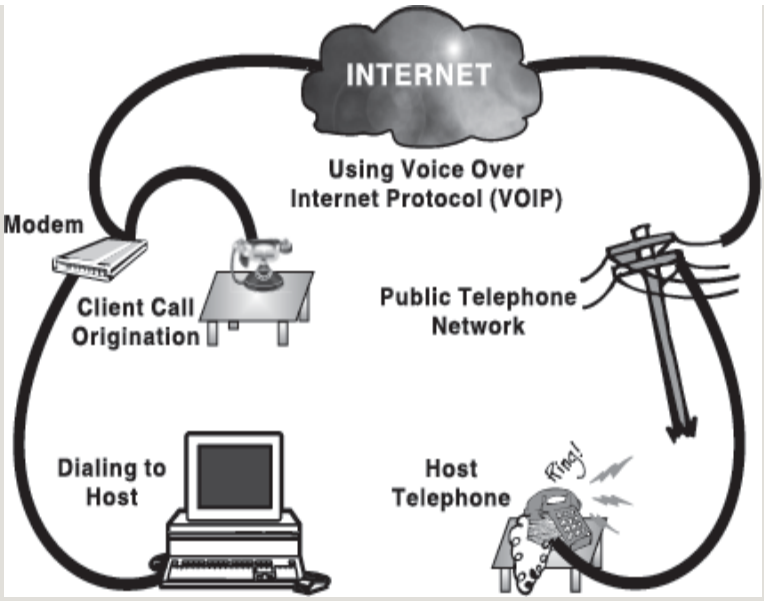


Figure 17.12 Internet Telephony

The required hardware for Internet telephony generally consists of end devices (either traditional telephones or audio-equipped personal computers) and gatekeepers that provide call admission control, bandwidth management, address translation, authentication, and user location. They communicate with each other using the voice over internet protocol (VOIP). For the cost of a local connection to their service providers, users can call anywhere in the world. Note that there is no single Internet telephone standard at this time. Therefore, participants need to use the same software at each end of the connection.

17.5.7 Video Conferencing

Video conferencing uses the same technology as IRC, but also provides sound and video pictures. It enables direct face-to-face communication across networks. A video conferencing system has two or more parties in different

locations, which have the ability to communicate using a combination of video, audio, and data. A video conference can be person to person (referred to as *"point-to-point"*) or can involve more than two people (referred to as *"multipoint"*), and the video conferencing terminals are often referred to as *"endpoints"* (Figure 17.13).

Figure 17.13 Video Conferencing

In this form of meeting, participants in remote locations can view each other and carry on discussions via Web cameras, microphones, and other communication tools. The following five elements are common to all video conferencing endpoints:

- **Web Camera:** It captures live images to send across the network.
- **Visual Display:** It displays the images of the people taking part in the video conference.
- **Audio System:** It includes both microphones to capture audio from the endpoint and loudspeakers to play back the audio received from

other endpoints across the network connection.

- **Compression:** Videos are very bandwidth intensive and they take a long time to load. Therefore, video systems include technologies, often referred to as codecs, to compress and decompress video and audio data, allowing transmission across a network connection in near-real time.
- **User Interface and Control System:** The user interface allows the users to control interactions, for example, placing calls, storing and locating numbers, and adjusting environment settings such as volume. The control system handles the underlying communication that takes place between endpoints.

Video conferencing has many benefits, as a tool for both teaching and learning. A key factor is that it provides real-time, visual communication, unlike other communications methods such as e-mail. Video conferencing technology is still in its infancy, and one of its major limitations is the bandwidth (the volume of information per unit time that a computer or transmission medium can handle) available on the Internet. As the protocols and applications for video conferencing develop higher resolution and improved speed, participation will increase.

17.5.8 Commerce through the Internet

Today business is taking place through electronic telecommunication media. Nowadays, various organizations offer the facility of shopping online. This type of business model is known as *electronic commerce* or *E-commerce*. E-commerce refers to buying and selling goods and services online. It integrates communication, data management, and security capabilities, which allow organizations to exchange information about the sales of goods and services. Essentially, it is the end-to-end digital exchange of information needed to conduct business including *electronic data interchange (EDI)* and *electronic funds transfer (EFT)*. EDI is the computer-to-computer exchange of routine business transactions including payments, information exchange, and purchase order requests. EFT is a technology that

allows the transfer of funds from the bank account of one person or organization to that of another (Figure 17.14).

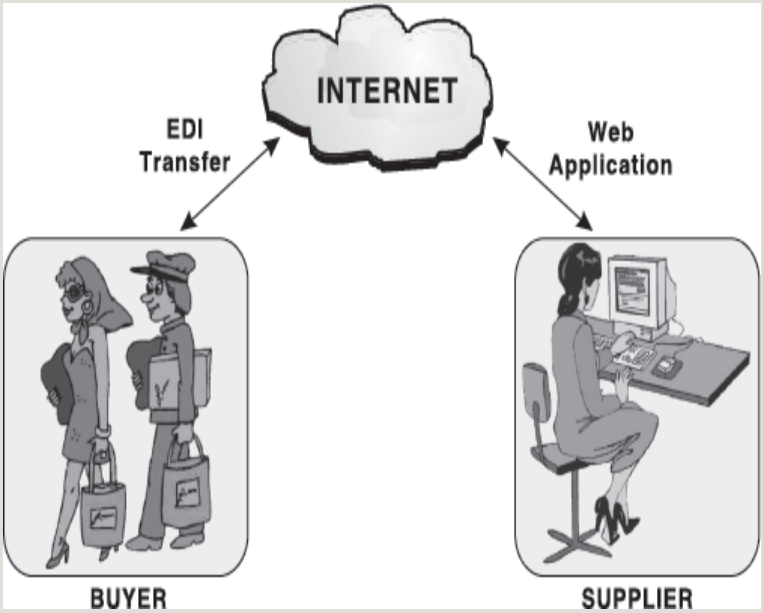


Figure 17.14 Electronic Commerce

While discussing about e-commerce, we normally talk about virtual shopping malls, up front store where the user virtually visits and selects the product(s) of his/her choice. After selecting the desired product(s), the payment for the purchase is done. Online payment can be done through credit card. Techniques for cash transactions have also been developed in the form of Internet money such as Cyber cash and Digital money. After transfer of money, the items are received at the doorstep, which means delivery of the product is done at the address desired by the buyer.

Through the development of e-commerce, the Internet has established an electronic global worldwide market

that operates 24 hours and 365 days a year. Organizations and individuals are able to market goods and services, while purchasing is supported by secure facilities such as digital signatures (the process that operates on a message to assure message source authenticity and integrity) and encryption (a method of ensuring data secrecy). Developments in software mean that an increased number of facilities are becoming available all the time. Some of the well-known e-commerce sites are http://www.amazon.com and http://www.ebay.com.

The increased use of mobile devices such as PDAs (personal digital assistants) and mobile phones has opened a new horizon of proliferation in the electronic market and has coined a new term "mobile commerce." *Mobile commerce* or *M-commerce* refers to transactions through a mobile phone network and data connection that result in the transfer of value (monetary or otherwise) in exchange for goods and services. Mobile commerce has evolved as a new and simpler concept to operate financial transaction through mobile phones and it includes services such as banking, payment, and ticketing. M-commerce users may view and interact with information that appears through SMS (short message service), WAP (wireless application protocol), or a standard HTML browser on a PDA, Web-enabled mobile phone or laptop computer connected to a mobile phone on a wireless network.

17.5.9 Newsgroups (Usenet)

Newsgroups are international discussion groups that focus on a particular topic and helps in gathering information about that topic. The topics discussed here cover all the fields such as politics, computers, technology, and many more. The information or articles that make up the "news" are written by people interested in a specific topic. These articles are posted to the

newsgroup so that others can read, reply, and comment on them. To read and post articles in a newsgroup, newsreader software such as Microsoft Outlook Express or Netscape News is required. Some newsgroups are devoted to current events and headlines. However, newsgroups usually describe topical discussion groups, not "the news" in the traditional sense. Most newsgroups exist on a network known as *"Usenet,"* which contains thousands of newsgroups (Figure 17.15).

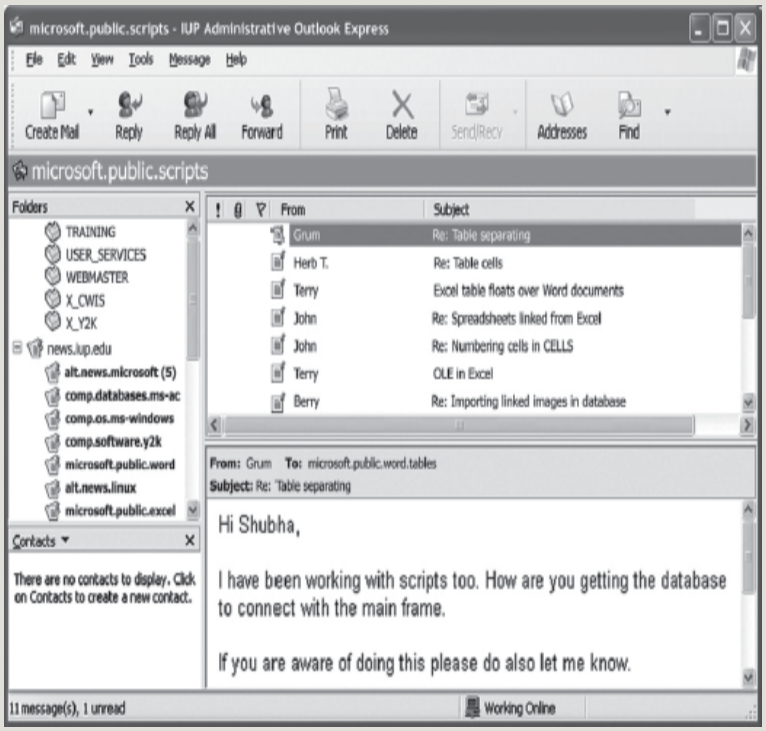


Figure 17.15 Reading News with Newsreader

Newsgroups provide a source of information and a medium through which users can ask questions from the Internet community. They provide a forum for users to discuss topics of mutual interest. It is normal to post the

rules, restrictions, and subject matter for the newsgroup in a message called *frequently asked questions* (*FAQ*). To simplify the selection of suitable newsgroups, they are divided into subject classifications known as top-level categories. These top-level categories are then further sub-divided and this continues down the hierarchy. There are 20 top-level categories including those shown in Table 17.1.

Table 17.1 Newsgroups Prefixes and Forums

Prefixes	Forums
Biz	Business
Comp	Computers and software
News	News and topical subjects
Rec	Recreation and entertainment
Sci	Science
Soc	Social issues and socializing
Talk	Debate and discussion
Misc	Anything else

17.5.10 Mailing Lists (Listserver)

The Internet is home to a large community of individuals who carry out active discussions, organized around topic-oriented forums that are distributed via e-mail. This method of Internet communication is known as *mailing list* and it enables people with similar interests from all over the world to communicate and share information with each other. The mailing list groups are administered by a software program known as *listserver*. There are varieties of programs available such as Listserv and Majordomo. A listserver watches for incoming mail on a certain mailbox continuously and forwards any message to a list of other addresses. Such a mailing list

can bring people together with a common interest (Figure 17.16).

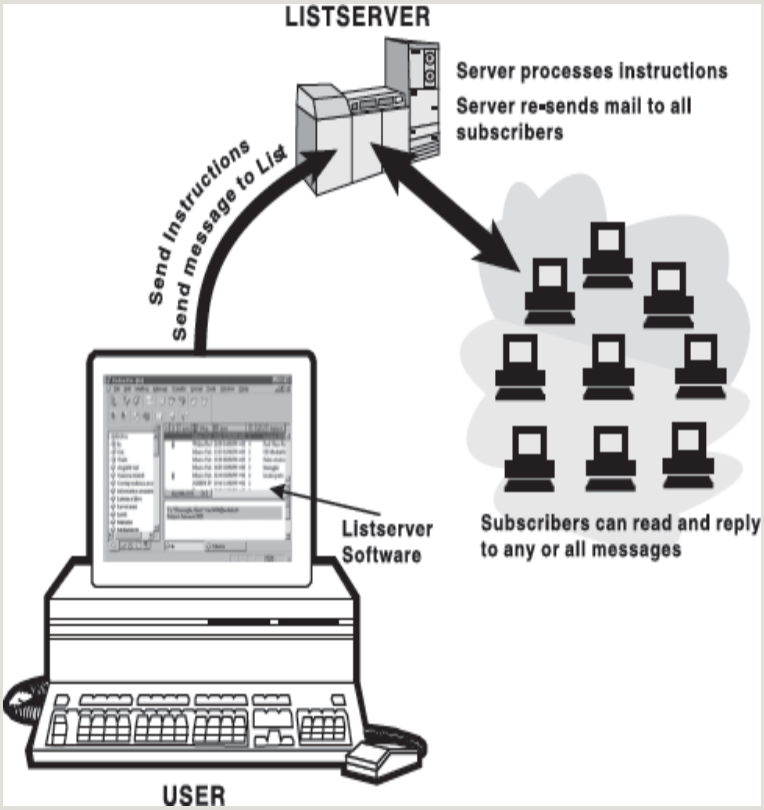


Figure 17.16 Mailing List

THINGS TO REMEMBER

Mailing List

There are hundreds of mailing lists, each for a different subject. Subscribers can send special messages to a separate address to retrieve files from the archive or to get other information about the mailing list. When the user subscribes to a listserver, messages from the other subscribers are

automatically sent to their electronic mailbox (e-mail address). Note that to participate in a listserver discussion, the user must have an e-mail account. To subscribe to a listserver, the user has to send an e-mail message to the desired listserver located on computer networks throughout the world.

17.6 DATA OVER THE INTERNET

The data, in terms of Internet, primarily comprises the hypertext documents containing links that connects to other documents or files. These documents may even include animations, and multimedia components. Depending upon how information is displayed in the browser, we can classify the displayed data into two categories: *Static Component* and *Dynamic Component*.

17.6.1 Static Component

Static component includes the information displayed on the browser that does not interact with the user. For example, the text written in the form of a paragraph on the Web page is static component. Producing static component for the Web page is accomplished by creating documents in HTML.

17.6.2 Dynamic Component

Dynamic component imparts life to a Web page. This component refers to a specification by which a user can interact with a Web server. For example, when a user fills a form on the Web page, the text filled in various fields is sent back to the originating point of the Web page. This way a user is able to interact with the server hosting that Web page. To impart dynamism, a special program (often called scripts) is written in any programming language including Perl, VBScript, JavaScript, and so on. Some of the programming languages used to provide dynamism to a Web page are listed below:

Perl: Web server uses external programs, called *CGI* scripts or *CGI* program, usually written in Perl to communicate with other programs such as database and mail running on the server. CGI programs are generally used to process the information from an interactive form on a Web page. A typical application would be to search a database. For example, a user might search for the word "white papers" on a Web site. The server receives the request, passes it to the CGI script, which looks up a database, and responds with the appropriate result formatted in HTML.

JavaScript and JScript: JavaScript is a programming language, which runs on the browser. It was developed by Sun Microsystems and was first implemented in Version 2 of the Netscape browser. JScript is Microsoft's implementation of JavaScript for the Internet Explorer.

JavaScript runs on the browser (client) and does not require any server software. Therefore, it is often called client-side scripting language. As all execution takes place on the browser, JavaScript is responsible for most of the interactivity on a Web page. Effects like changing of images, text colours, and other things can be made possible through JavaScript. The language has also been widely used for basic form validation.

VBScript: Like JavaScript, VBScript is also a client-based language that has been developed by Microsoft and it runs only on the Internet Explorer. Although JavaScript runs on both Netscape and Internet Explorer, VBScript is easier to code when developing active server pages.

Active Server Pages (ASP): ASP is a technology promoted by Microsoft. It uses some special tags, which can be embedded in the HTML code to generate dynamic Web pages. ASP scripts run on the server side. A typical

user can use VBScript, JavaScript/ Jscript, or a combination of both to write ASP pages. The great advantage in using ASP is the ease of maintenance of the Web site.

PHP: PHP is an open-source development environment to develop dynamic Web pages. PHP works almost like an ASP, but it is more secure and handles databases more easily. PHP code is embedded inside the HTML and can be linked to databases for generating dynamic content.

XML (eXtensible Markup Language): XML is a markup language that enables programmers to create customized tags. These tags can provide the much-needed functionality presently unavailable with HTML. XML documents can be accessed using ASP, PHP, and so on.

Java: Java is a very powerful and object-oriented language developed by Sun Microsystems. Web-based Java applications are generally in the form of "Java applets." They are small programs written in Java, which are called by an HTML page from a Web server and run on a Java-compatible Web browser.

LET US SUMMARIZE

1. *Internet* is a worldwide system of computer networks, which allows the participants to share information. The foundation of Internet was laid in 1969 by the DoD of United States of America.
2. The *World Wide Web* (abbreviated as the *Web* or *WWW*) is a collection of linked documents or pages, stored on millions of computers and distributed across the world.
3. A *Web page* is a document on the WWW, which is created with HTML. The Web pages are linked together through hyperlinks, which enables the user to jump from one Web page to another by clicking on a link.
4. A *Web site* is a set of related (linked through hypertext links) Web pages, published by an organization or an individual. All the Web sites on the Internet constitute the WWW.
5. A *home page* (also called *index page*) is the starting point or a doorway to the Web site; it refers to the Web page that identifies

a Web site and contains the hyperlinks to other Web pages in the Web site.

6. A *browser* (short for *Web browser*) is a special software that enables users to read/view Web pages and jump from one Web page to another. It is the software that is needed to find, retrieve, view, and send information over the Internet.
7. Each Web page has a unique address, called a *URL*, which identifies its location on the Internet.
8. The *hypertext* is a system that provides a simple and consistent way to organize large data (that includes text, images, pictures videos, etc.) available on the Internet. In hypertext systems, the documents contain links to other related documents on the Internet. These links are called *hyperlinks, hotlink*, or simply *links*.
9. An *ISP* is an organization that connects its subscribers' computer using modem to the Internet. Some of the well-known ISPs are America Online (AOL), Microsoft Network (MSN), Yahoo!, etc.
10. A *Web server* receives request from the clients, processes this request, and sends the result back to the client. In other words, Web server fulfils the client's (browser) requests.
11. *Download refers* to the act of transmitting data from a remote computer on the Internet or other network to one's own computer. On the other hand, *upload* refers to the act of transmitting data from local computer to any other computer on the Internet or network.
12. *Modem* is a device that allows computers to communicate over telephone lines. It converts digital data into analog signals that can be sent over an analog telephone line and converts the analog signal back into digital data.
13. *TCP/IP* is the basic *communication* protocol of the Internet. It allows programs on user's computer to communicate over the Internet.
14. The *FTP* is a set of rules that enables a user to transfer files from one system to another.
15. *Telnet* is a protocol that allows a user to log on to a remote computer. Also known as remote login, it connects the user's machine to another in such a way that the user may interact with another machine as if it is being used locally.
16. *IRC* is a service on the Internet that allows people to communicate in real time and carry on conversations via the computer with one or more people.
17. A *chat room* is a virtual meeting room where computer users can communicate with each other in real time. People from all over the world can come together in a chat room to discuss topics of mutual interest.
18. *Internet telephony* is the use of Internet instead of traditional telephone infrastructure to exchange spoken or other telephonic information. It consists of hardware and software that enable people to use the Internet as a transmission medium for telephone calls.
19. *Video conferencing* is a form of meeting in which participants in remote locations can view each other and carry on discussions via Web cameras, microphones, and other communication tools.
20. *E-commerce* refers to buying and selling goods and services online and if the transaction is done through a mobile phone network, it is known as mobile commerce or m-commerce.
21. *Newsgroups* are international discussion groups that focus on a particular topic and helps in gathering information about that

topic. To read and post articles in a newsgroup, newsreader software such as Microsoft Outlook Express or Netscape News is required.

22. *Mailing list* enables people with similar interests from all over the world to communicate and share information with each other.
23. Depending upon how information is displayed in the browser, we can classify the displayed data into two categories: *static component* and *dynamic component*.

EXERCISES

Fill in the Blanks

1. A computer on a network that requests files from another computer is known as .....
2. Each Web page has a unique address, called ..... that identifies its location on the Internet.
3. The ..... is a collection of linked documents or pages, stored on millions of computers and distributed across the world.
4. A program that enables users to read/view Web pages and jump from one Web page to another is known as .....
5. ..... refers to the act of transmitting data from a remote computer on the Internet or other network to one's own computer.
6. The US Department of Defense laid the foundation of the Internet with a network called .....
7. AltaVista, Google, and HotBot are the examples of .....
8. The Web uses a specific Internet protocol called ..... to support hypertext documents
9. The technique for ..... is known as Telnet.
10. The mailing list groups, administered by a software program, are known as .....

Multiple-choice Questions

1. URL is .....

    1. a computer software program.
    2. a type of Web server.
    3. the address of a document or “page” on the WWW.
    4. an acronym for unlimited resources for learning.

2. What would you use for immediate, real-time communication with a friend?

    1. e-mail
    2. IRC
    3. Usenet
    4. Mailing list

3. What does the URL http://www.uni.edu tell you about the source?

    1. It is associated with an educational institution
    2. It is associated with a military establishment
    3. It is associated with a commercial organization
    4. None of the above

4. When a user subscribes to a newsgroup, ...

    1. All new posts are e-mailed to the user automatically.
    2. The user must agree with everything said in that newsgroup.
    3. The user is billed annually for the subscription.
    4. The user is barraged with spam.
5. The Internet is owned by .....

    1. The US government
    2. A consortium of telecommunications companies
    3. The IETF
    4. None of the above

6. A home page is .....

    1. a Web page created by an individual home user, as opposed to the one created by a corporation.
    2. the Web page you like to visit the most.
    3. the Web page that identifies a Web site.
    4. a Web page you eventually reach by clicking a series of hyperlinks.

7. Which of the following is not a type of Internet connection?

    1. ISDN
    2. Dial up
    3. Leased line
    4. ISP

8. E-mail is .....

    1. Mail concerning electronic devices
    2. Transaction of letters, messages, and memos over a communication network
    3. Transaction of messages within a computer
    4. None of the above

9. Which one is not the language of Internet?

    1. Perl
    2. HTML
    3. XML
    4. COBOL

10. Choose the odd one out.

    1. Web page
    2. Home page
    3. Index page
    4. Cover page

State True or False

1. WWW cannot incorporate sound.
2. All URLs are unique.
3. Telnet cannot be used for remote login.
4. .EDU is a valid domain type.
5. Internet Explorer is a character-based browser.
6. Messengers can be used to send e-mails.
7. Dial up is the fastest possible Internet connection.
8. Microsoft owns Internet.
9. Outlook Express can be used for reading news.
10. Listserv is a Web browser.

Descriptive Questions

1. Explain a typical structure of a URL.
2. List the various services provided by the Internet.
3. What is a newsgroup? How can you subscribe to it?
4. Distinguish between Internet and WWW.
5. Define the following

1. Web page
2. Home page
3. Web site

6. Write short notes on the following:

    1. Chatting
    2. Instant messaging
    3. Internet telephony
    4. Video conferencing

ANSWERS

Fill in the Blanks

1. Client
2. URL
3. WWW
4. Web browser
5. Downloading
6. ARPANET
7. Search engines
8. Hypertext transfer protocol
9. Remote login
10. Listserver

Multiple-choice Questions

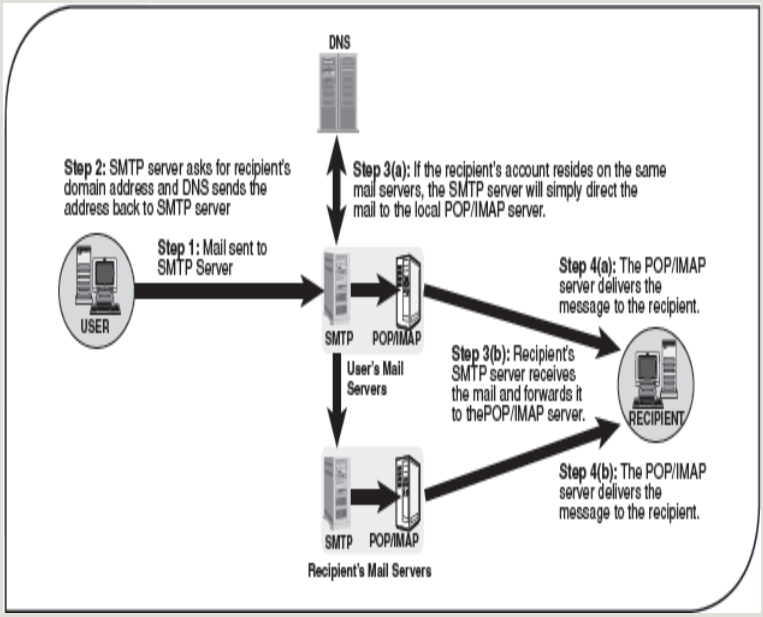
1. (c)
2. (b)
3. (a)
4. (a)
5. (d)
6. (c)
7. (d)
8. (b)
9. (d)
10. (d)

State True or False

1. False
2. True
3. False
4. True
5. False
6. True
7. False
8. False
9. True
10. False

Chapter 18

Internet Tools



This chapter introduces the reader to different Internet tools and how these are used for different purposes such as browsing, searching and sending and receiving e-mails. The chapter mainly focuses on the browsing and searching the internet and using it for sending and receiving e-mails. The web browser tool is discussed with its functions and how it helps in quick searching. An important searching tool–search engine is discussed along with its working. The second part of the chapter discusses e-mail client to send and receive e-mails offline. The latter part of the chapter describes the messenger, which is used for sending instant messages from one computer to another.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The Internet tools, which are versatile tools that enable a user to collect information and interact with people across the world
- Web browser—a software application, which provides the user an interface to navigate the Internet
- How to browse the Internet using Internet explorer
- The e-mail client, which is used to send and receive e-mails over the Internet
- Various search engines used to access information on topics desired by the user
- Instant messaging—a web-based service used to exchange messages in real time between two or more people over the Internet

18.1 INTRODUCTION

In recent years, Internet has become one of the most popular means of communication. Whether one wants to know about computers, literature, commerce, or other such subject matters, Internet is probably the best solution that provides all kinds of information at a single place. We have discussed some of the most important terms related to Internet in the previous chapter, such as e-mail, browsing, instant messaging, etc. Now let us study those concepts from a practical perspective. This chapter focuses on various Internet tools like browser, e-mail client, and messenger. All these versatile tools enable the user to collect information and interact with people across the world.

A web browser is the most essential tool for accessing the Internet. Through a browser, one can browse, search, and collect information from the Web. It is also used to

send and receive e-mails in online mode. If users want to manage their e-mails offline, they can use an e-mail client, which is specially designed for the purpose of sending and receiving messages without the need to be online. Apart from these tools, many websites provide a special application known as the instant messenger which permits users to send and receive instant messages, transfer files, conduct audio/visual chatting, listen to Internet radio stations, and many such useful services.

18.2 WEB BROWSER

A *web browser* (or simply browser) is a software application, which provides a graphical user interface (GUI) so that the user can navigate the Internet easily by clicking on menus, icons, or buttons instead of learning difficult keyboard commands. A web browser uses the HTTP protocol to request web pages from the web server. These pages contain special instructions (written in HTML) that tell the browser how to display the web contents on the user's screen. The instructions may include hyperlinks to other web pages, information about text formatting and colour, position information for images contained in the document, and other such things. Most browsers natively support a variety of formats in addition to HTML, such as the JPEG, PNG, and GIF image formats, and can be extended to support more through the use of *plug-ins* (an add-on piece of software that extends the features or functionality of a larger application). The two most popular web browsers are Netscape Navigator and Microsoft Internet Explorer.

18.2.1 Web Browser Features

There are many web browsers available having different features. However, every browser shares some of the basic features mentioned below:

- A browser handles requests for HTML files, interprets links, and deals with embedded images, audio, and video elements.
- A browser keeps the history of the websites visited.
- A browser lets a user to save a collection of pages, allowing for later retrieval.
- A browser provides a row of buttons at the top of the browser window for browsing the Internet conveniently.
- A browser connects to an e-mail program for importing favourites/bookmarks and sending and receiving e-mails from the most commonly used applications and formats.
- A browser supports web standards currently in use such as HTML, HTTP, JavaScript, and Unicode.
- A browser supports multimedia data in three ways.
    - **Native Support**: A browser supports some of the image formats natively. In other words, a browser presents the image as a part of the web page in the browser window.
    - **Plug-ins:** A plug-in is a program that adds functions to the web browser, such as an audio player or a compression utility. A plug-in is a set-up within the browser through which it functions properly and correctly.
    - **Helper Programs**: Helper programs are external applications launched by a web browser or Internet client program to view non-native file formats and data types. These programs are not integrated with the browser.

Note: *In this book, we have used Internet Explorer 8.0 (IE 8.0) for explaining web browser concepts Internet Explorer 8.0 is freely available from Microsoft website (http://www.microsoft.com). It is also bundled with the Windows 7 operating system.*

18.2.2 Opening Internet Explorer

To open Internet Explorer, perform *any* one of the following steps:

- Double-click the **Internet Explorer** icon located on the desktop.
- Click the **Internet Explorer** icon from Quick Launch.
- Click **Start** and select **Internet Explorer**.
- Click **Start**, select **All Programs**, and then select **Internet Explorer** (Figure 18.1).

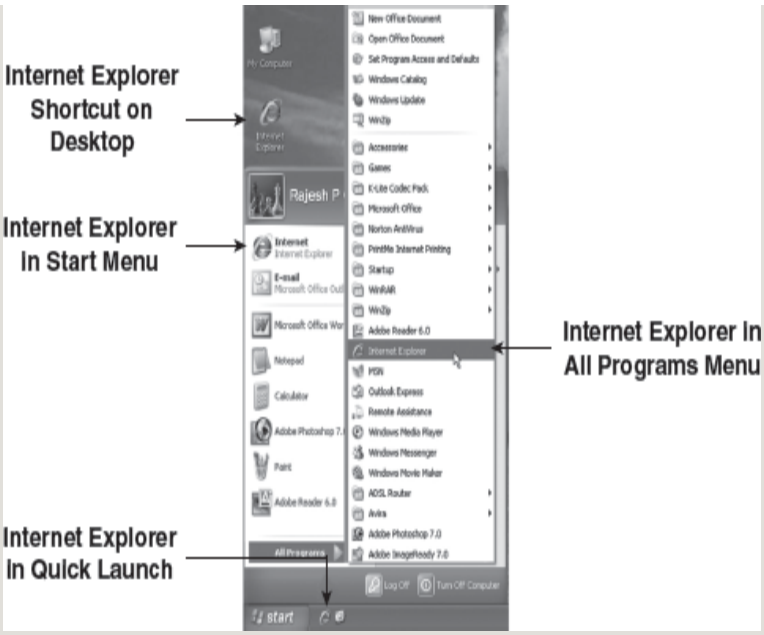


Figure 18.1 Opening Internet Explorer

18.2.3 Internet Explorer Environment

When Internet Explorer is opened, the main screen of the Internet Explorer is displayed. This main window has many parts, which are described in detail below (Figure 18.2).

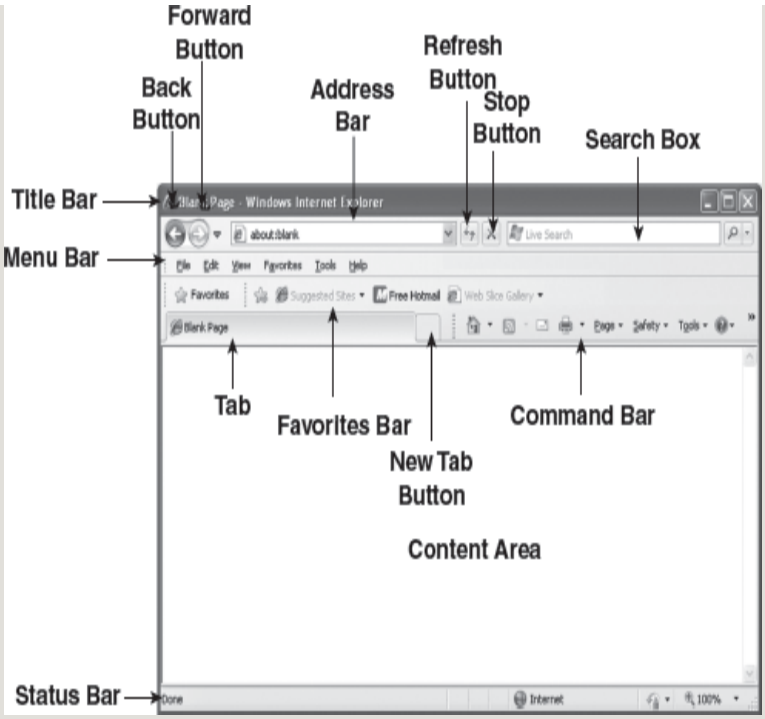


Figure 18.2 Main Internet Explorer Browser Window

Title Bar: The title bar is located at the very top of the window and tells you the title of the page you are viewing. The title bar also tells you which Internet Explorer application is currently active. In addition, like all other Microsoft applications, it has the **Minimize, Restore/Maximize**, and **Close** buttons.

Tabs: Internet Explorer 8 allows opening multiple tabs in a single browser window with each tab displaying a web page. You can switch between the tabs to view web page of your choice by clicking on the respective tab. Tabbed browsing helps reducing the number of items in the taskbar. To open a new tab either click the **New Tab** button or press **CTRL+T**.

Back and Forward Buttons: The *Back button* allows the user to move to the most recently displayed web page. On the other hand, the *Forward button* allows moving to the page the user most recently has moved back from. This way the user can navigate through the most recently viewed websites easily.

Address Bar: The *Address bar* displays the Internet address (URL) of the page currently being displayed. You can access a site by typing its URL into the address box and pressing Enter (pressing **ALT+Enter** opens the web page in a new tab). In addition, you can search the Web by typing **Find, Go**, or ? followed by the keywords, phrase, or website name in the Address bar and pressing **Enter**. The Address bar uses the default search provider for searching.

Refresh Button: The *Refresh button* reloads the current document that the user is viewing. Reloading the current page is useful when the page is updating very frequently as the user can view these changes as soon as they are available.

Stop Button: The *Stop button* stops the operation that is currently being performed by the Internet Explorer. For example, if it is pressed before a page has finished loading, the page will display only those elements that had been loaded before the button was pressed.

Search Box: The *search box* is used to search the Web. Simply type the keyword or phrase that you want to search and then press **Enter** to open the search result in the same tab or **ALT+Enter** to open the result in a new tab. To go to the search box, you can either click in the search box or press **CTRL+E**.

Menu Bar: The *menu* bar has many different sub-menus, which control all options, functions, and

commands for the entire Internet Explorer. Some of the browsing controls can also be found in these sub-menus. Each command in the menu bar represents a menu and is activated by pressing **ALT**+ the underlined letter of the menu from the keyboard. Different menus include **File, Edit, View**, and so on.

Favorites Bar: The *Favorite bar* replaces the Links toolbar available in previous versions of the Internet Explorer. Like Links toolbar, it enables you to keep the links of your favourite websites so that they can be accessed with just a click. Now, you not only can drag links from the address bar, but also from web pages you are viewing. In addition, the Favorites bar can hosts RSS feeds and Web slices. *Web slice* is a new feature that lets you notice about updates to the contents on your favourite websites.

Command Bar: The *Command bar* hosts many buttons that are used to perform frequently used actions. Buttons that appear on the command bar are **Home** button, **Read Mail** button, Print button, **Help** button,etc. You can customize the command bar to add or remove buttons from it.

Content Area: The *content area* is the portion of window that holds the web page present in the current tab. Text, images, animation, and links available on the web page appear in this area. It is quite common that a web page is too large to fit in the browser window at once. In such situation, scroll bars located on the right or bottom of the window are used to view the page's area that is beyond the window region.

Status Bar: The *Status bar* displays the current state of activity on the web page. The information that appears on the status bar includes the item currently being downloaded, the current web page's download progress,

error (if any) in downloading, and zoom level of web page.

18.3 BROWSING INTERNET USING INTERNET EXPLORER

Browsing is an act of looking for information on the Internet by repeatedly going through several web pages usually with the help of hyperlinks. It is also known as *navigating or surfing* the Web. Browsing the Internet is the same as walking through a shop with an intention to purchase something or just for fun. Let us consider an example to illustrate how to browse Internet.

THINGS TO REMEMBER

Browsing the Internet

There is no standard way for browsing the Internet. However, usually when a user opens a website through the web browser, it contains several hyperlinks. With the help of these hyperlinks, a user can jump to a new web document in the same or a new window. Remember that efficiently browsing the web is just like any other task in life, it requires practice to be good at it.

Opening and Navigating a Website: In the **Address** bar, enter the website's address (URL) and press **Enter** to open the website. When the website opens, its home page is displayed. Since most web pages contain more information than can be displayed in one window, you need to know to move through a page to view all the information. Some common methods that can be used are as follows:

- Use the vertical or horizontal scroll bars to move around or through a document.
- Press the **Up** and **Down** arrow keys to move up or down one line. Pressing the Page **Up** (or **PgUp**) key moves up one window length and

pressing **Page Down** (or **PgDn**) moves down one window length. Pressing **Ctrl+Home** takes you to the beginning of the document and pressing **Ctrl+End** takes you to its end.

Usually a web page contains many hyperlinks, in the form of both text and images. You can always tell when the mouse pointer is over a hyperlink because it will change to hand with a pointing index finger ( ). To open a hyperlink, simply click on it and the related web page opens in the same window of the Internet Explorer. You can also open a hyperlink by right-clicking it and selecting the **Open** option from the short-cut menu. In addition, you can select the **Open in New Tab or Open in New Window** option from the short-cut menu (see Figure 18.3) to open the link in a new tab or a new window, respectively.

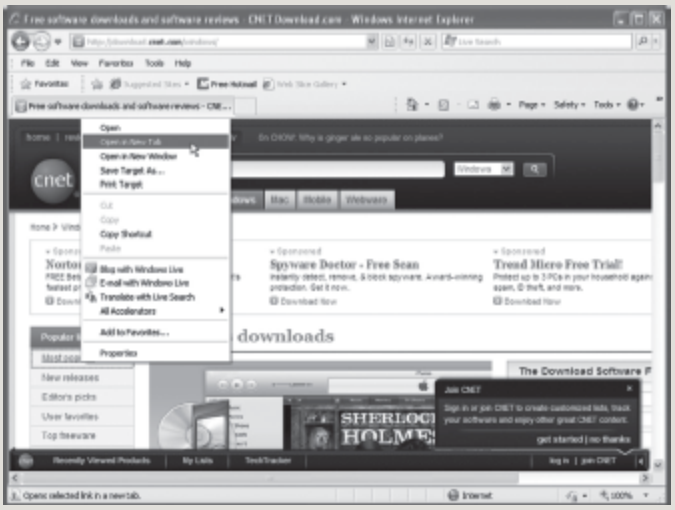


Figure 18.3 Web Page in Internet Explorer

If you have browsed through several pages in a window, you can move backward and forward by clicking the **Back** and **Forward** button at the top of Internet

Explorer. You can go back or forward one page at a time or you can select from a list of pages. To do this, click the down arrow ( ) beside the Forward button and click on any site from the list of visited sites (see Figure 18.4).

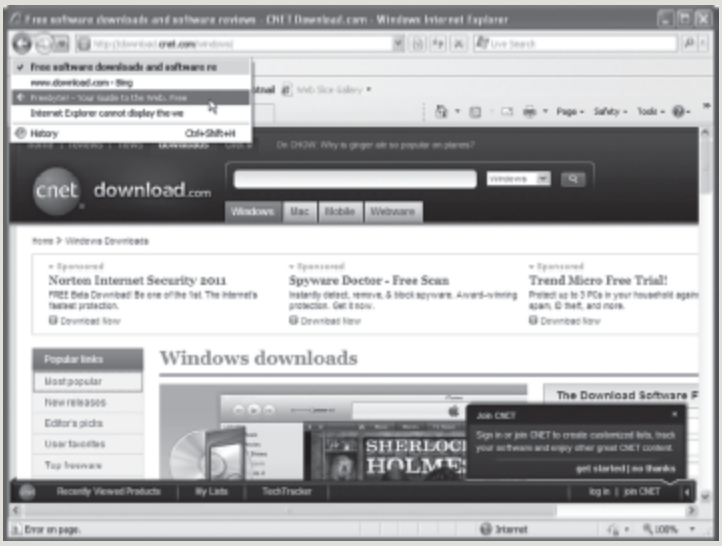


Figure 18.4 Moving Backward and Forward

Saving Text and Images: If you wish to save any text from the Internet, first select the text on the web page. After selecting the desired text, right-click it to open a pop-up menu, which contains options like **Cut, Copy, Paste, Select All,** and **Print**. Note that, often the **Cut** and **Paste** options are inactive to prevent users from deleting the important data from the website. Once you copy the text from a web page, you can paste it in any text-editing software like the Notepad or WordPad (Figure 18.5).

Figure 18.5 Copying Text

You often come across pictures on the Web that you would like to save, print, e-mail to others, or save them as wallpapers. To perform these activities, right-click the image and select the Save Picture As option. This will display the **Save Picture** As dialog box where you can select the location and enter a name for the picture.

In case you want to save the entire web page for future reference, click on the **File** menu and select the **Save As** option to display the Save Webpage dialog box. Now choose where you want the file to be saved, under what name, and in what format and then click the Save button (Figure 18.6).

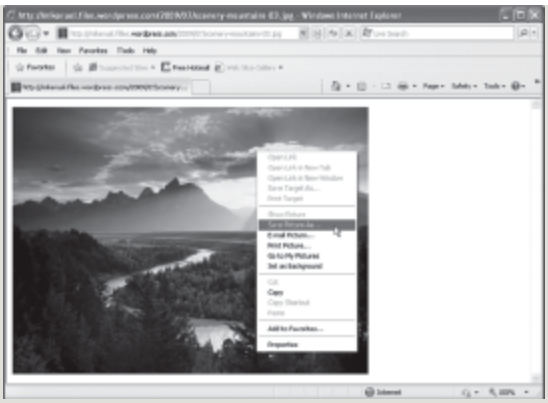


Figure 18.6 Saving an Image

Downloading Files from Links: Many web pages have links to files (like .ZIP or .EXE files) that are not natively displayed in the Internet Explorer content area. Such files have to be copied on the hard disk so that the user can use them. To save such files, click on the link. Internet Explorer will ask you whether to save the file to the hard disk. If you intend to save the file then click the **Save** button to display the Save As dialog box so that you can provide the location on the hard disk to save the file. Another way of saving a file is to right-click the link and selecting the **Save Target As** option from the short-cut menu. Again, the Save As dialog box will appear where you can name the file and provide the location to where it will be downloaded (Figure 18.7).

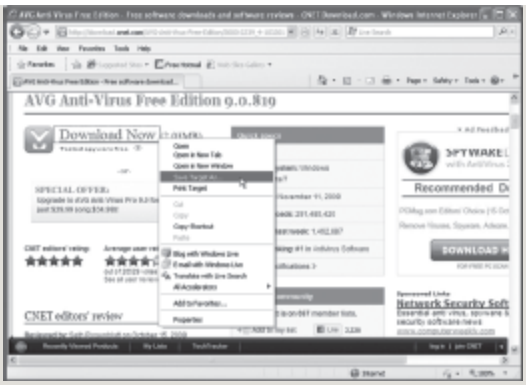


Figure 18.7 Downloading a File

Add to Favorites: Internet Explorer uses a feature called *Favorites* to put websites, one visits often,within easy reach. That way, a website can be opened simply by a click without the user having to remember or type anything. It is similar to a bookmark in a book. It lets users find their place without checking the table of contents or the index or having to write down the page number. To learn how to use Favorites, follow the steps given below:

1. Go to the web page that you want to add to your Favorites.
2. On the **Favorites** menu, click **Add to Favorites**.
3. In the Add a Favorite dialog box, type a new name for the page if you want, and click **OK** (Figure 18.8).

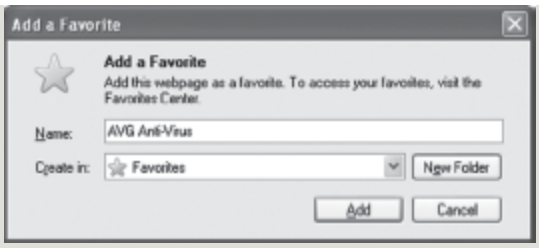


Figure 18.8 Add Favorite Dialog Box

Once a website is added to the Favorites list, simply click the link to the web page you want to visit (Figure 18.9).

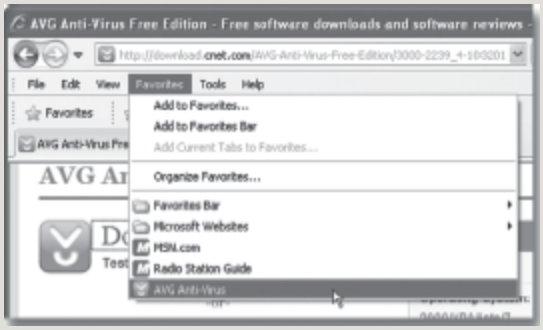


Figure 18.9 Favorite Menu

18.4 ELECTRONIC MAIL (E-MAIL)

The Internet was developed mainly to communicate and interchange knowledge irrespective of the distance limitation. Hence, from the very beginning, users used Internet primarily for sending and receiving messages. Even today, communicating through the Internet is the main application of the Internet. Since these messages are communicated electronically, they are known as electronic mail or e-mail. Thus, *electronic mail* can be defined as the process of exchanging messages

electronically, via a communications network, using the computer. E-mails allow users to communicate with each other in less time and at nominal cost as compared to traditional telephone or mail services. Apart from a textual message, e-mails can also consist of other data formats such as pictures, sound, and video. E-mails can be sent anywhere in the world using your computer and a modem. Its delivery is almost instant and is very economical to use especially for international messages, costing only the price of the telephone connection to the Internet service provider (ISP). You may send many messages at one time or just one to a designated location.

18.4.1 E-mail Address Structure

In order to use e-mail, one must have access to the Internet and an e-mail account. An e-mail account is a service that allows the user to send and receive e-mails through the Internet. Usually, e-mail accounts are provided as part of ISP's monthly packages. One can also obtain free e-mail accounts such as Yahoo and Hotmail on the Web. An e-mail account provides a unique e-mail address and a mailbox where the user can save all his/her mails.

Generally, there are two parts of an e-mail address: *the logon identity* and the *identity of the e-mail server*. These are separated by the symbol @ (pronounced as at the rate). For example, a typical e-mail address would look like:

The first part of the address indicates the identifiable name of the user. It is just like a home address (and it is

unique) so that the mails could be sent to that address. The user name is a name, which was assigned to or desired by the user, while signing up for e-mail address. For example, *feedback@itlesl.com*. The @ symbol in the address is used to separate the user name from the rest of the address. Next comes the host name (itlesl), also called the *domain name*. This refers to the mail server, the computer where the recipient has an electronic mailbox. It is usually the name of a company or an organization. The end of the domain name consists of a dot (.) followed by three or more letters (such as .com or .net). This part of the domain name indicates the type of organization or the country where the host server is located. Note that you may find variations in e-mail addresses, but these common elements will always remain the same.

Note: *There are no spaces in an e-mail address and it is usually (but not always) in lower-case letters.*

18.4.2 Checking E-mails

To send and receive e-mails, the user must have an e-mail account. In case, you do not have a mail account, you must sign up with e-mail provider. Let us assume that we have an e-mail account. Now, follow the steps given below to check mails.

1. Open Internet Explorer and type **www.yahoo.com** into the address bar and press the **Enter** key to display Yahoo's home page. Now click the **Yahoo Mail** link under **MY FAVORITES** section as shown in Figure 18.10.

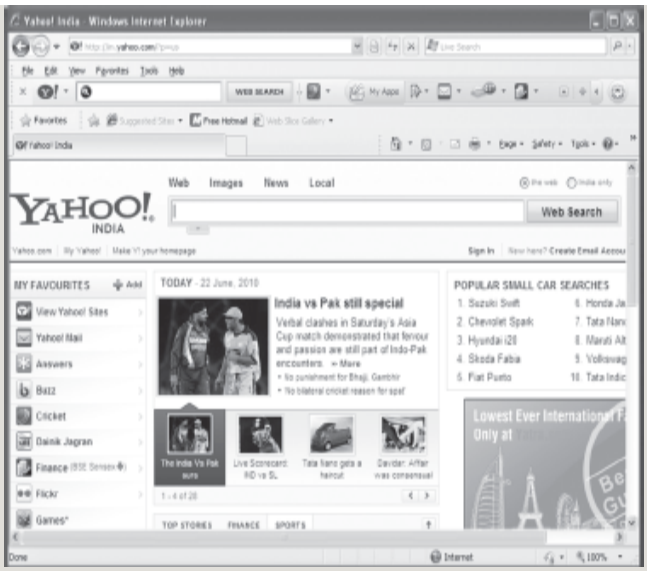


Figure 18.10 Home Page of Yahoo

2. When the **Yahoo Mail** link is clicked, Yahoo loads a new web page, which allows users to log on to their mail accounts. To log in, you have to enter your username (e-mail address) and the password. After entering the information, either press the **Enter** key or click the **Sign In** button (Figure 18.11).

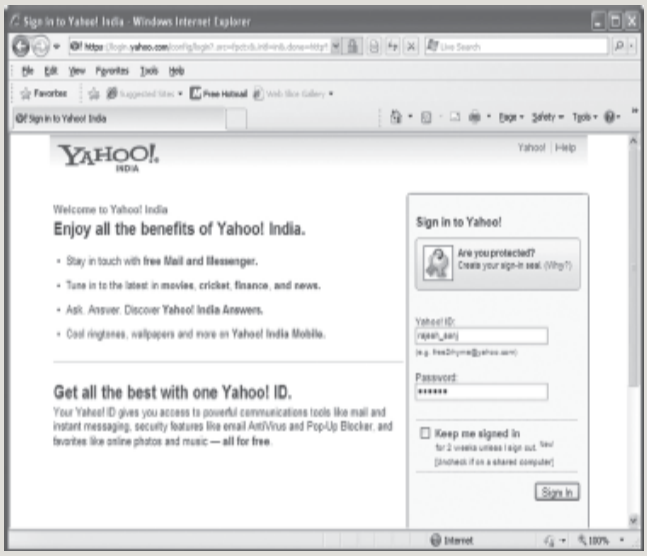


Figure 18.11 Login Page

3. To check for new e-mail(s), click the **Inbox** link. The screen would look as shown in Figure 18.12.

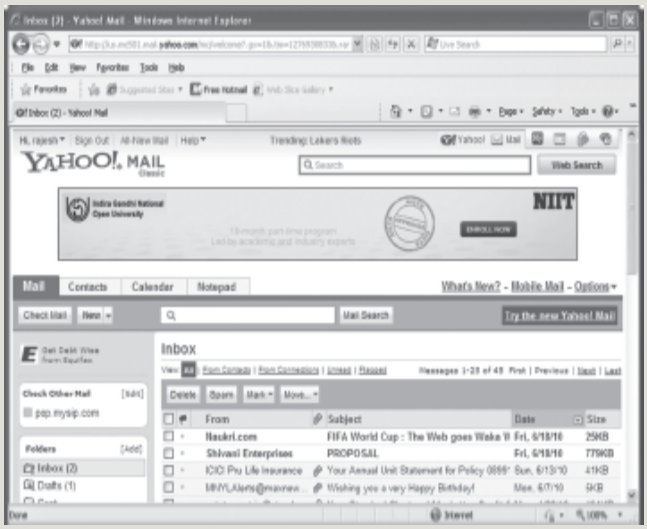


Figure 18.12 Inbox Page

4. To read e-mail, simply click on the subject of the e-mail. The screen would look as shown in Figure 18.13.

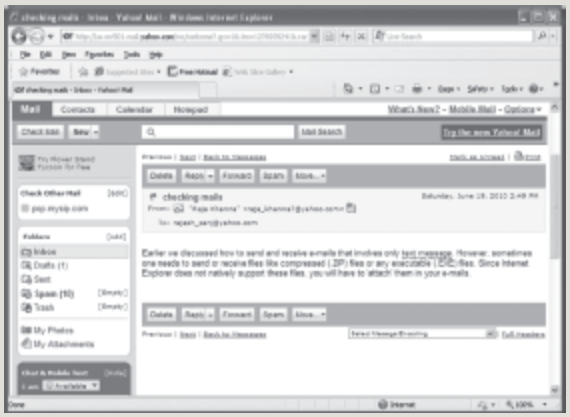


Figure 18.13 Reading E-mail

5. After reading the mail, you can reply to the mail or forward the same mail to other persons by clicking the **Reply** or **Forward**

button, respectively. In case, you do not want to keep the mail in the mailbox, click the **Delete** button to delete it.

18.4.3 Sending E-mails

One of the important functions of an e-mail service is to provide a platform to send e-mails. A user can reply to the received e-mail or send a new e-mail. To send an e-mail, follow the steps given below.

1. Click on the **New** button so that you can create and send e-mail to the desired address. When the **New** button is clicked, the **Compose Mail** page will be displayed as shown in Figure 18.14.

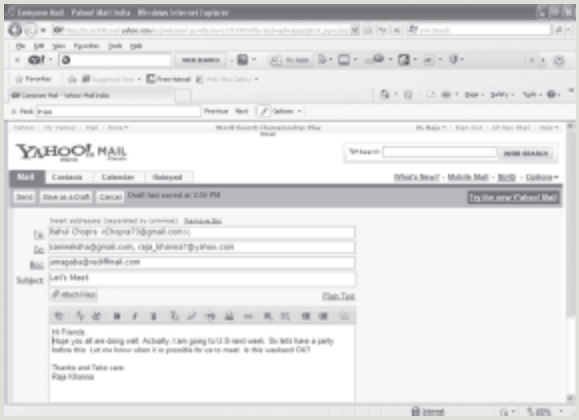


Figure 18.14 Compose Mail Page

The key elements of the Compose Mail page are listed below:

1. **To**: It denotes to whom the mail is to be sent. The user has to write the recipient's e-mail address.
2. **Cc**: Cc or Carbon copy is used to specify the addresses of all the recipients who will also receive copies of the same mail. It is used to send the same message to several people (all the addresses are separated with commas).
3. **Bcc**: Bcc or Blind carbon copy is used to send message to several addresses without showing everyone all the addresses.
4. **Subject**: It denotes the subject of the message as specified by the sender.
5. **Attach Files**: Clicking this button opens another page where user can attach files (created with programs like MS-Word or Zip files) that are to be sent with the e-mail.

2. Now write the desired contents in the text area and click the **Send** button.

18.4.4 E-mail Attachments

Earlier we discussed how to send and receive e-mails that involve only text message. However, sometimes one needs to send or receive files like compressed (.ZIP) files or any executable (.EXE) files. Since Internet Explorer does not natively support these files, you will have to "attach" them in your e-mails.

Receiving and sending attachment is similar to receiving and sending normal mails, with some extra steps. To view an e-mail containing an attached file, follow the steps given below:

1. When you open an e-mail that contains an attachment, it will be displayed as an icon as shown in Figure 18.15.

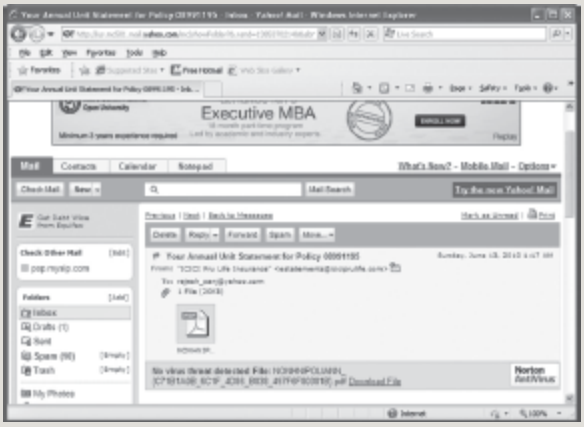


Figure 18.15 Attachment

2. Click on the file icon. As you click, the file will be scanned for viruses and the scan report is shown to you below the file icon.
3. Finally to download the file, click the **Download File** link, which displays the File Download dialog box. Click Save or Open button to save or open the file, respectively. If you click Save, the Save As dialog box appears, which prompts you to save the attached file onto your hard disk.

Note: *Always scan the attached file(s) for viruses. In addition, it is a good practise to avoid unsolicited attachments because they may contain virus.*

To send an attachment with the mail, follow the steps given below:

1. When you compose a mail, click the **Attach Files** button (Figure 18.14), which will display the page as shown in Figure 18.16.

Figure 18.16 Attach Files Page

2. Click on the **Browse** button to display the Choose File to Upload dialog box, which allows you to select the desired file to be sent as an attachment. Repeat this process to attach multiple files.
3. Click on the **Attach Files** button to attach the file(s) to your message and return to the Compose Mail page, which will now display all the attached file's names over the **Attach More** button (Figure 18.17).

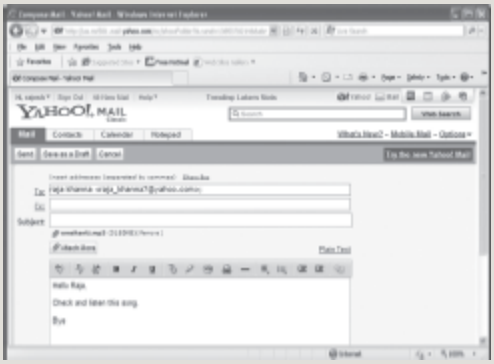


Figure 18.17 Sending E-mail with an Attachment

4. Finally, write the desired contents in the text area and click the **Send** button to send the mail along with the attachments.

18.4.5 How E-mail Works

To send e-mail, one needs a connection to the Internet and access to a mail server, which forwards the mail. The standard protocol used for sending e-mail is called *SMTP (simple mail transfer protocol)*. It works in conjunction with *POP (post office protocol)* and *IMAP (Internet mail access protocol)* servers.

When an e-mail is sent to a person, who has an e-mail address like *xyz@hisdomain.com*, it is broken down into two parts: xyz (the recipient's account name) and *hisdomain.com* (the recipient's domain name). The SMTP server contacts a DNS (domain name service) server and asks for the location of *hisdomain.com*. The DNS server sends the address back to the SMTP server. The SMTP server then sends the e-mail message to the SMTP server where *hisdomain*. com is located. This SMTP server delivers the e-mail message to xyz's account on the POP or IMAP server. Finally, when xyz logs on to his computer and opens his e-mail client, his e-mail client requests the POP or IMAP server to send all mails from the account to his computer. Note that if you are sending mail to someone whose account resides on the same mail servers, the SMTP server will simply direct the mail to the local POP or IMAP server, where it will be delivered to the appropriate account (Figure 18.18).

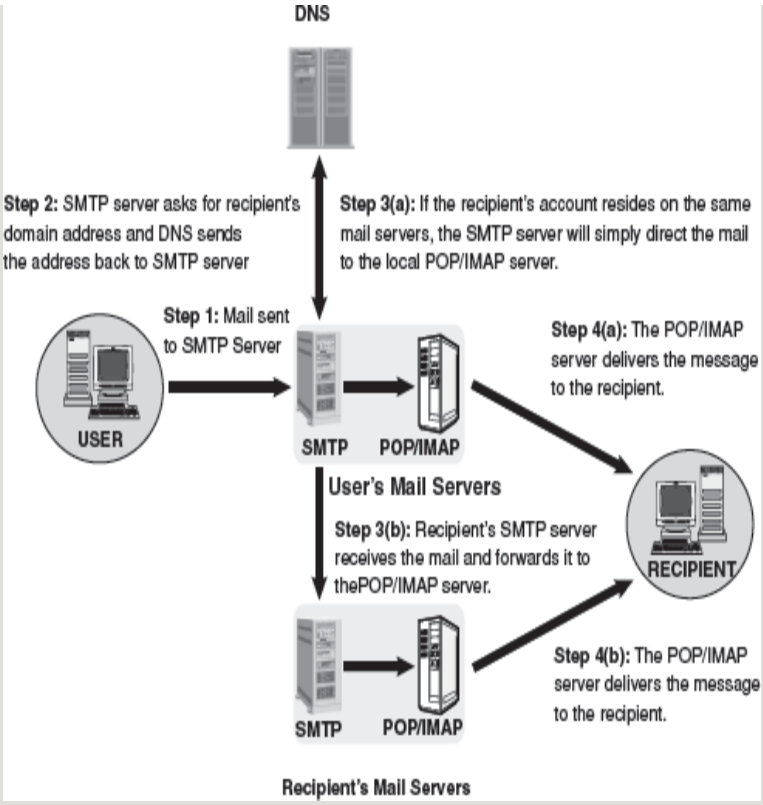


Figure 18.18 Working of E-Mail System

18.4.6 Netiquette

Communicating online, whether by e-mail, chat rooms, or message boards, is informal, quick, and easy. However, there are certain unsaid rules that one should follow. On the Web, you are what you write, so take some pride in your work and do it correctly. The standard rules of a polite Internet behaviour are called *netiquette*. Like e-mail and other online communication, the rules of netiquette are short and to the point. Some of these rules are briefed below:

- Never do anything online that you would not do in real life. Even while communicating online, you are dealing with real people. Therefore, do not say anything over the Internet that you would not say on their faces.
- Good discussion should always be welcomed but do not waste other people's time with useless queries. For example, if you respond to a

listserver to send a private message to one person, you have wasted a lot of people's time. Moreover, before participating in a newsgroup discussion, one should read the FAQs (frequently asked questions).

- Do not use all upper-case letters because using all caps equates to yelling. A single word in upper case (to emphasize) is acceptable, but use upper-case words sparingly.
- Never send unsolicited junk mail and do not forward junk mail or chain letters.
- Respect other people's privacy and copyrights. Treat the mails that you receive as confidential unless the sender permits you to share it with others. Moreover, if you are quoting something from an online book or paper, always mention the source.
- When posting a message to a public bulletin board, forum, or newsgroup, stick to the topic. Posting unrelated comments to a newsgroup is known as *spamming*, which leads to another unpleasant Internet practice known as falming. *Flaming* is the practice of attacking people on a personal level.
- Respect other people's time and bandwidth. While mailing or posting to a discussion group, you are taking up other people's time. It is your responsibility to ensure that the time they spend reading your posting is not wasted. Hence, be concise and brief while mailing and posting information online.
- When someone makes a mistake, whether it is a spelling error, an unnecessarily long answer, be kind about it. If it is a minor error, try to avoid it and even if you feel strongly about it, be polite and point it preferably by private e-mail rather than in public.

18.4.7 Advantages and Disadvantages of E-mail

E-mail can be sent anywhere in the world to anyone who has an e-mail address. It can take days to send a letter across the country and weeks to go around the world. To save time and money, more and more people are using e-mail. However, the slightest error in the address can deliver the mail to the wrong address. Hence, like there are two faces of a coin, e-mail also comes with its own share of benefits and limitations (Table 18.1).

Table 18.1 Advantages and Disadvantages of E-mail

Advantages	Disadvantages
• The delivery of messages is very fast, sometimes almost instantaneous, even though the message is meant for overseas or just to a friend next door	• Although e-mail is delivered instantly, the recipient may or may not read his/her mail on time. That defeats the quickness of electronic mailing
• The cost of e-mailing is almost free as it involves negligible amount of telephone and ISP charges.	• The user must stay online to read and write more than one mail. In addition, most webmail either display advertisements during use or append them to mails sent. It results in increased size of the original mail, which brings a significant decrease in speed of use
• Multiple copies of the same message can be sent to a group of people at the same time and can be sent as easily to a single person	• Since e-mail passes through a network, it may be intercepted in between
• Pictures, documents, and other files can also be attached to messages	• The slightest error in the address or a failure in one of the links between sender and receiver is enough to prevent a delivery

18.5 E-MAIL CLIENT

In the previous section, we discussed how to send and receive e-mails using a web-based e-mail service like Yahoo. This kind of e-mailing is known as *webmail*, which is a World Wide Web interface that allows users to read and write e-mails using a web browser. Webmail has certain limitations like most webmail providers offer only a limited amount of space to save mails and the user has to stay online to manage all his/her e-mails. To solve such problems, *e-mail clients* are used. Some examples of e-mail clients include Microsoft Outlook Express, Eudora, and Pegasus Mail.

18.5.1 Features of E-mail Client

The main advantage of using an e-mail client is that it is generally used offline. Users have to connect to the Internet only when they are ready with all their mails. For example, if a user wants to send a long mail, he/she can compose and send it without going to the website

and wasting precious Internet access time. Furthermore, once all the mails are downloaded, they can be managed into separate folders, again in offline mode. This is simply not possible in webmail because, the users cannot organize them in offline mode even if they can copy–paste mails. Some of the most fundamental features that an e-mail client provides are given below.

- **Offline Mode of Operation**: User can read and compose e-mails without connecting to the Internet.
- **Filtering E-mails**: One can categorize the new, read, and the replied e-mails according to the sender, date, subject of the e-mail, and ascending or descending order of the received e-mails. Additionally, there is a facility to print or delete the chosen set of e-mails.
- **Managing Received E-mails**: The e-mails received by a user can be saved in different folders with suitable names. In this way, the folders can be easily searched for the subjects or items for the contents of e-mails.
- **Managing Multiple E-mail and Newsgroup Accounts**: An e-mail client manages several e-mail and newsgroup accounts so that the user can work with all of them in a single environment.

Note: *In this book, we have explained e-mail client with the help of Microsoft Outlook Express 6.0, which is freely available from Microsoft (http://www.microsoft.com).*

18.5.2 Opening Outlook Express

To open Outlook Express, perform any one of the following steps:

- Double-click the **Outlook Express** icon located on the desktop.
- Click the **Outlook Express** icon from Quick Launch.
- Click **Start** and select Outlook Express.
- Click **Start**, select All Programs, and then select **Outlook Express** (Figure 18.19).

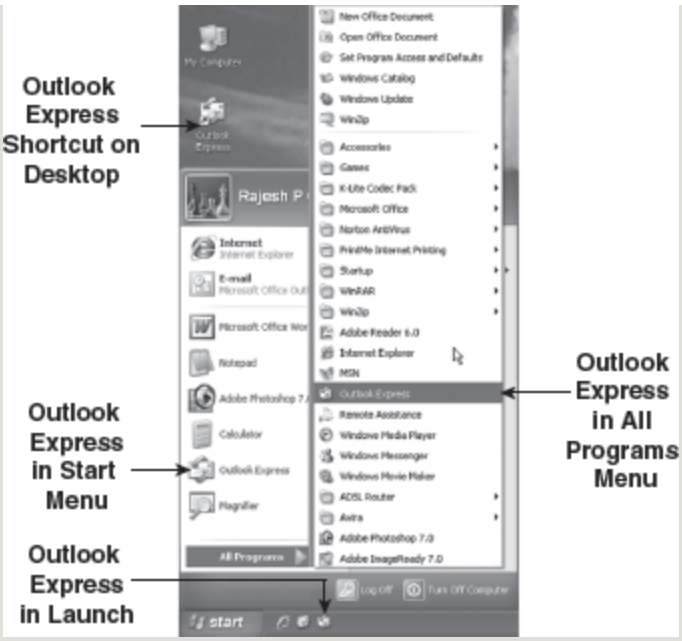


Figure 18.19 Opening Outlook Express

When Outlook Express is opened, the main screen of the program appears, as shown in Figure 18.20. Notice that the screen is divided into four panes: *Folders, Folder Contents, Contacts*, and *Preview Pane*. The Folders pane displays different folders where messages are stored. The Contacts pane displays the names in your address book. As the name suggests, the Folder Contents area of the Outlook Express screen will display the content of various folders in the Folder pane, and the Preview pane lets you glimpse your messages.

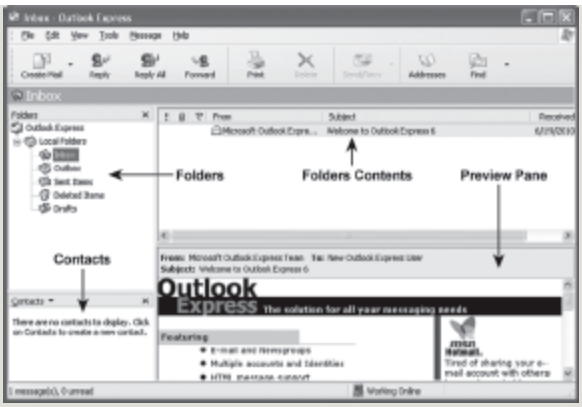


Figure 18.20 Outlook Express Environment

18.5.3 Creating an Account

Before you can send or receive e-mail using the Outlook Express, you need to configure it to communicate with incoming and outgoing mail servers. However, before setting up an account, you must know the following information:

1. Your e-mail address, account name, and password.
2. Information about the incoming and outgoing e-mail servers (you may need to contact your Internet Service Provider to get it).

Note: In this book, we focus on the most popular configuration—an Internet-standard SMTP server that supports POP3 connections.

To create a new account, follow the steps given below:

1. Start Outlook Express and select **Accounts** from the **Tools** menu.
2. To create a new account, click on the **Add** button, and then click **Mail** from the pop-up menu to open the Internet Connection wizard (Figure 18.21).

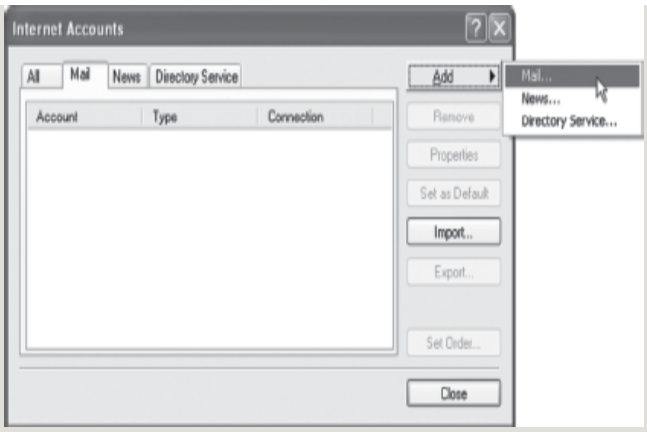


Figure 18.21 Selecting Mail Option from the Add Button

3. Type your name (or nickname) in the **Display name** textbox, as you want it to appear to everyone who gets e-mail from you, and then click **Next** (Figure 18.22).

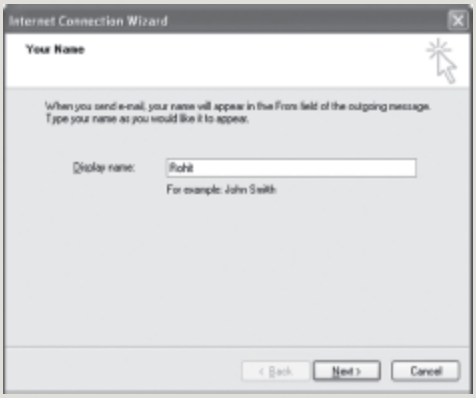


Figure 18.22 Entering Display Name

4. Type your e-mail address in the **E-mail address** textbox, the same address that others will use to send you e-mail, and then click **Next** (Figure 18.23).

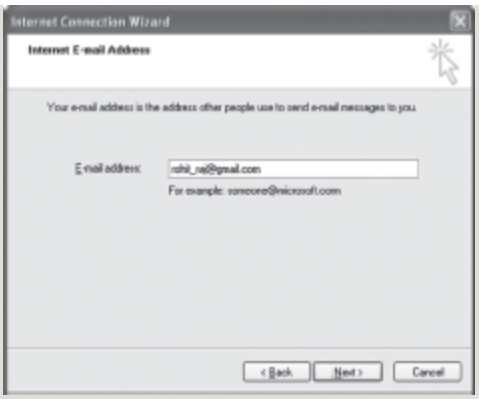


Figure 18.23 Entering E-mail Address

5. Select the type of mail server that you will be accessing. For regular Internet service, you will most likely be using a POP3 account. After that, type the names of the incoming and outgoing mail servers (usually they are same). If required, check with your ISP to ascertain the name of incoming and outgoing mail servers. If you choose HTTP as your incoming e-mail server —as for a Hotmail or MSN account—this wizard page changes slightly so you can identify your HTTP mail service provider. After providing the required details, click the **Next** button (Figure 18.24).

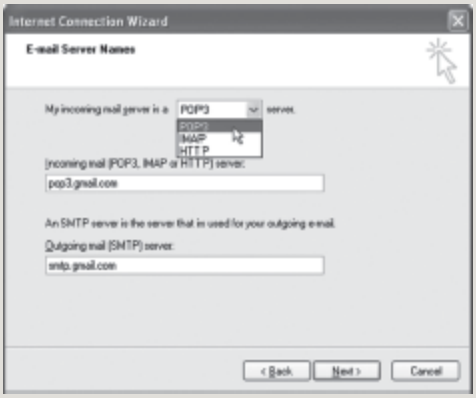


Figure 18.24 Entering Mail Server Names

6. Enter your account name, the name that appears before the @ symbol in your e-mail address, along with your e-mail password (if necessary, obtain it from your ISP) and click **Next** (Figure 18.25).

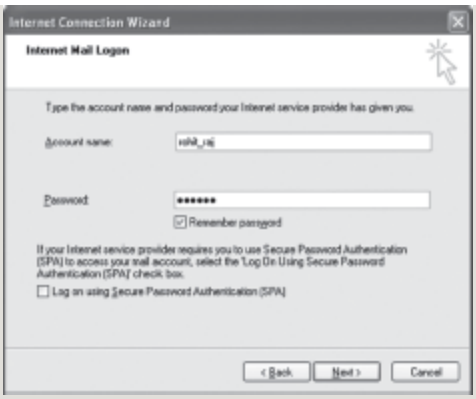


Figure 18.25 Entering Account Name and Password

7. Click **Finish** to save the settings (Figure 18.26).

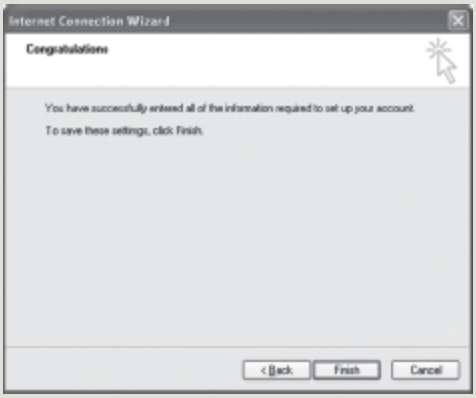


Figure 18.26 Saving the Settings

As you click on Finish, Internet Accounts dialog box re-appears containing the newly created mail account in the list as shown in Figure 18.27.

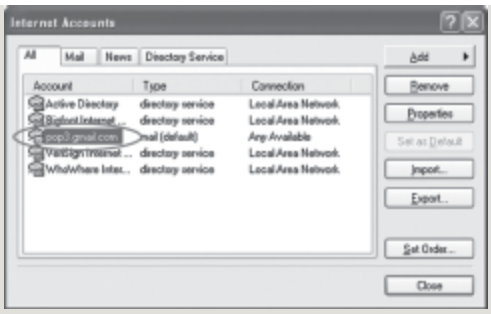


Figure 18.27 Newly Created Mail Account

18.5.4 Sending E-mails Using Outlook Express

Sending mails using Outlook Express is similar to sending mail using any web-based mail accounts. To send e-mail, follow the steps given below:

1. Click on the **Create Mail** button ( ) on the toolbar, or select **New** from the **File** menu, and then select Mail Message option to open a new message window (Figure 18.28).

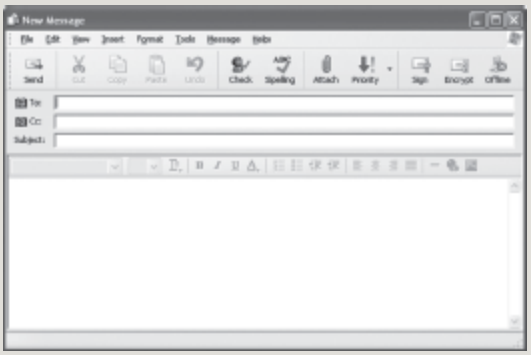


Figure 18.28 New Message Window

2. Now compose your mail in similar fashion, as you would do in any web-based mail (Section 18.4.3). If you want to send an attachment, either select **File Attachment** option from the **Insert** menu or just click **Attach** button ( ) from the toolbar. This will display the Insert Attachment dialog box, using which you can choose the desired file(s).

3. After composing the mail, simply click the **Send** button ( ) on the toolbar.
4. After the message has been sent successfully, an acknowledgment appears in Outlook Express informing that the message has been sent to their corresponding destinations. Note that all the sent messages are also automatically saved in the Sent Items folder.

18.5.5 Checking E-mails Using Outlook Express

To open the received e-mails in Outlook Express, there is no requirement to open a web browser. Instead, Outlook Express itself downloads the received e-mails. The e-mails sent to you are stored on the mail server of your ISP until you retrieve them. For checking e-mails in Outlook Express, the following steps are required:

1. Click the **Send/Recv** button ( ) on the toolbar, or select **Send and Receive** from the **Tools** menu, and then select **Receive All** option. Outlook Express then requests the mail server to transfer e-mails, if any, into the application and then the downloading of e-mails begins (Figure 18.29).

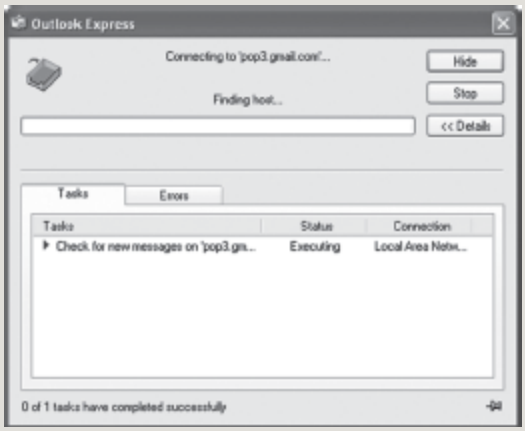


Figure 18.29 Connecting to Mail Server

2. In the **Folders** pane, click the **Inbox** folder. New and unread messages appear in bold in the Folders Contents pane, as shown in Figure 18.30.

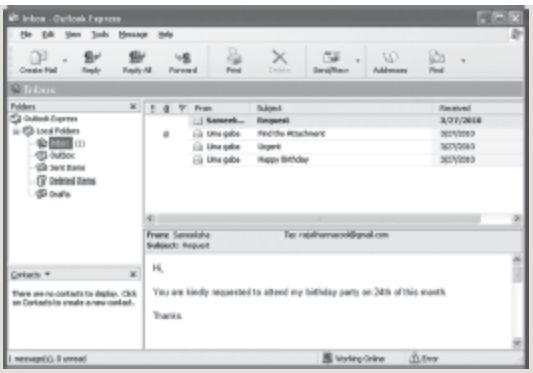


Figure 18.30 Checking for New Mails

3. To read a new message, click it in the Folders Contents pane. The contents of the message are displayed in the Preview pane, as shown in Figure 18.30. You can also doubleclick the message to open the message.
4. If there is an attached file with the message, a symbol ( ) appears before it. To open an attachment, doubleclick its file icon. The attachment opens in the appropriate program (Figure 18.31).

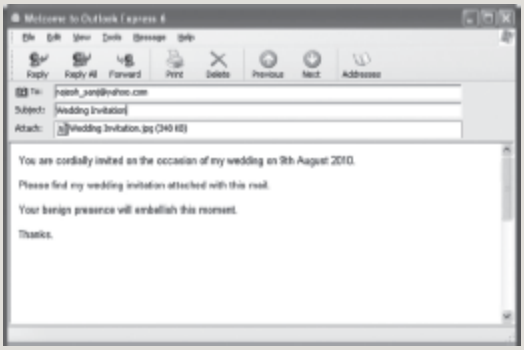


Figure 18.31 Viewing a Mail with Attachment

18.5.6 Creating Newsgroup Account

A newsgroup is a virtual place where people with common interests can ask questions and get answers on just about any imaginable topic ranging from sports to

space. Newsgroups generate huge volumes of messages for which you need a *newsreader* to read. That is where Outlook Express comes in. Not only Outlook Express is an e-mail client, it is also used to gain access to newsgroups. To gain access to a newsgroup, your first step will be to set up a news server account.

As with an e-mail account, the Internet Connection Wizard provides an easy-to-use interface to create a newsgroup account. The steps involved in creating a newsgroup account are given below:

1. Start Outlook Express and select Accounts from the Tools menu.
2. To create a new account, click on the Add button, and then click News from the pop-up menu to open the Internet Connection wizard (Figure 18.32).

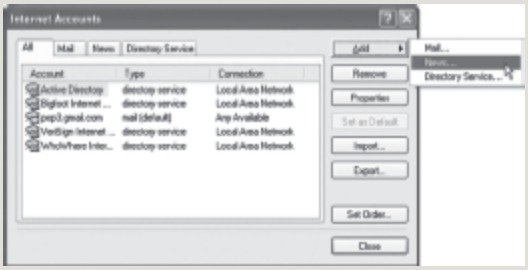


Figure 18.32 Selecting News Options from the Add Button

3. On the Your Name page of the wizard, type your name in the Display name textbox, as you want it to appear when you post messages to a newsgroup, and then click Next.
4. On the Internet News e-mail Address page, type your e-mail address, and then click Next (similar to Figure 18.23).

Figure 18.33 Entering News Server Name

5. On the Internet News Server Name page, type the name of the news server for which you have signed up, and then click Next (Figure 18.33).
6. Click **Finish**, and the new newsgroup account appears in the list, as shown in Figure 18.34. Click Close to close the dialog box.

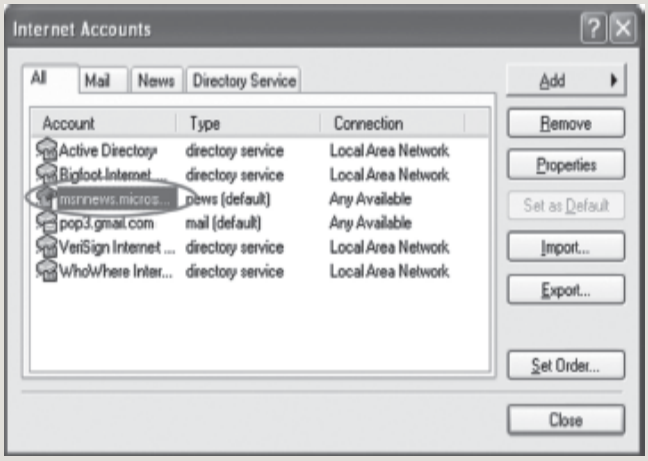


Figure 18.34 Showing Newly Created News Account

7. Now Outlook Express will ask you whether to download newsgroups. If you want to download, click Yes.

Once you subscribe to a newsgroup, it will be displayed in the Folders pane, as shown in Figure 18.35. In a

newsgroup, engaging with people works a lot like e-mail. People correspond by posting messages back and forth, which you can read and reply to, adding your thoughts to the discussion.

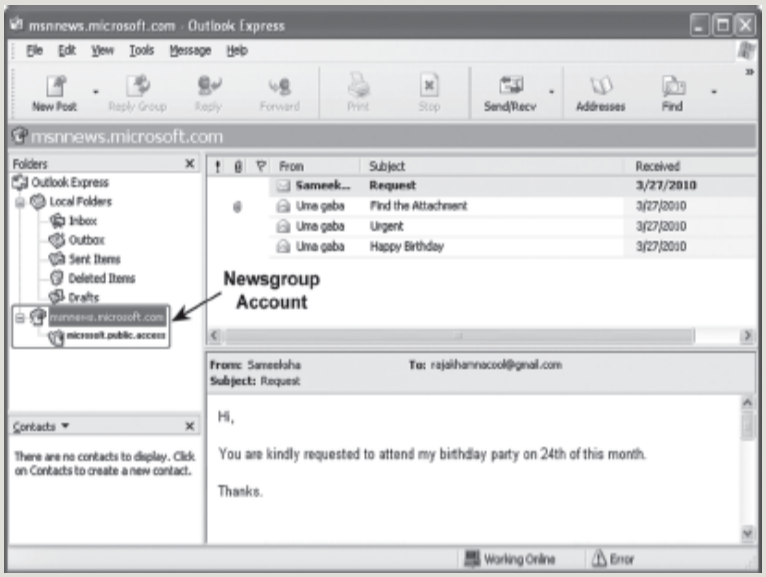


Figure 18.35 Reading News

18.6 SEARCH ENGINES

The Internet provides access to a wealth of information on countless topics contributed by people throughout the world. However, the Internet is not a library in which all its available items are identified and can be retrieved by a single catalogue. In fact, no one knows how many individual files (could be in billions) reside on the Internet. Hence, to conduct a search on the Internet, a special search tool known as *search engines* is used. *Search engine* searches a database of Internet files collected by a computer program called a *wanderer, crawler*, or *spider*. It allows the user to enter keywords relating to particular topics and retrieve information

about the Internet sites containing those keywords. As such, a search engine consists of four components:

1. **Spider**: Program that traverses the Web from link to link, identifying and reading pages.
2. **Indexing Software**: Program that analyses web pages that are downloaded by spiders.
3. **Database**: Warehouse of the web pages downloaded and processed.
4. **Search Engine Mechanism**: Software that enables users to query the index and that usually returns results in term relevancy ranked order (Figure 18.36).

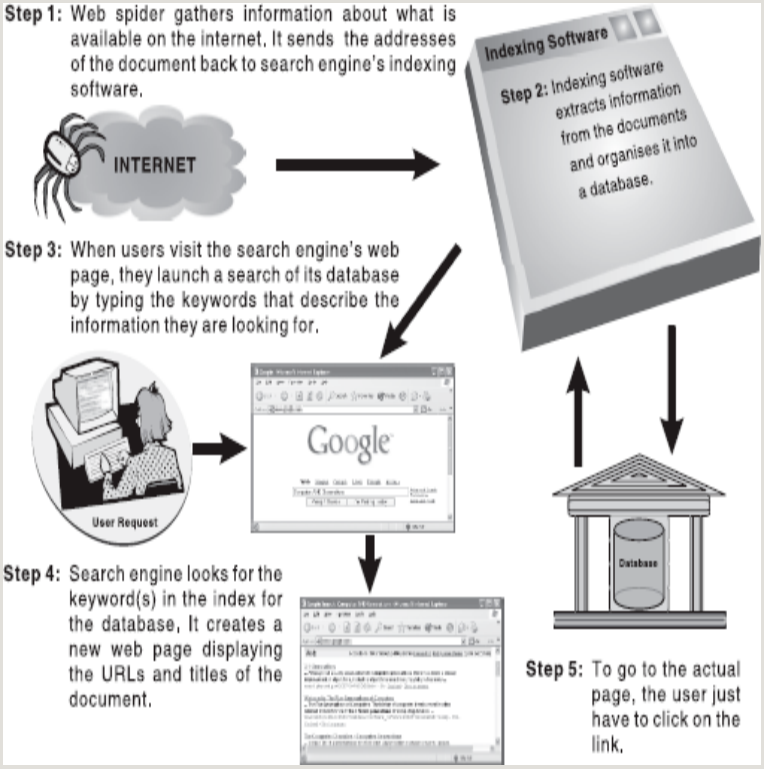


Figure 18.36 Working of a Search Engine

A search engine does not really search the Web directly. To find information on the millions of web pages, a search engine employs special software, called spiders.

After spiders find pages, they pass them on to another computer program for *indexing*. This program identifies the text, links, and other content in the page and stores it in the search engine database's files so that the database can be searched by keyword. Note that creating index and updating search database is a never-ending process because of the constantly changing nature of the Web. As a result, the spiders are always “crawling.”

When users search the Web using a search engine, they are provided with the links of all the searched web pages. On clicking on the links provided in a search engine's search results, the current versions of the web pages are retrieved from the server.

18.6.1 Searching the Internet

With most search engines, you fill out a form with your search terms and then ask the search engine to find web pages relevant to the search terms. Some of the well-known search engines are *www.google.com*, *www.hotbot.com*, *www.lycos.com*, and *www.altavista.com*.

Let us assume that you want to search the Web to get information regarding computer generations. To do this, first open the search engine's website (such as *www.google.com*). Now type your keyword(s) in the search box and click the **Google Search** button or press **Enter** (Figure 18.37).

Figure 18.37 Google Search Engine

The engine searches its index and generates a page with links to those resources containing some or all of the search terms. Clicking on any link opens a website or web page in which the searched keyword appears (Figure 18.38).

Figure 18.38 Search Result Page

18.6.2 Refining the Search

The major search engines allow the user to choose whether to search for the exact typed phrase, all the words in the phrase, any of the words in a phrase, and so on. However, a few search engines do not provide such options. In this case, the user Figure 18.38 Search Result Page can refine the search by adding one or more words or symbols to the search topic.

- **AND**: It is used to search for two or more terms on the same page. Type the word AND between the terms (for example, poverty AND crime) or put a plus sign right before the second term (example: poverty+crime).
- **OR**: It is used to search for either of two (or more) terms on the same page. Type the word OR between the terms (for example, college OR university).
- **NOT**: It is used to search for pages that include the first term but not the second. Type the words NOT between the terms (for example, cats NOT dogs) or put a minus sign right before the second term (for example, cats-dogs).
- " ": Double quotes are used to search for an exact phrase. For example, "desktop wallpapers."
- ( ): Parentheses are used to group parts of the search phrase. For example, type desktop AND (themes OR wallpapers). This returns pages with the words desktop and themes or both the words themes and wallpapers.
- *: Asterisk is used to search for various forms of a word. For example, prior* returns pages with the words prior, prioritize, priority, and so on.

18.7 INSTANT MESSAGING

Instant messaging is a web-based service to exchange messages in real time between two or more people over the Internet. It is a combination of e-mail and chat room. For sending instant messages, users create a list of users with whom they wish to communicate. The messages sent through this service take less time and the delivery of messages is instantaneous. However, to send and receive instant messages, Internet connection along with instant messaging software (known as *instant messengers*), such as *Yahoo Messenger* or *MSN Messenger*, is required. Instant messenger is one of the most popular Internet tools, which allows the users to

communicate over the Internet through voice, live pictures, or text. It is free and faster than e-mail because instant messages can be sent to friends whether they are online or offline. It allows the users to carry multiple conversations at a time.

Usually, users send their messages in text format. However, while sending instant messages, a user can also communicate through webcam (video chatting) as well as through speakers and microphone (voice chatting). In addition, one can send audio or video files along with instant messages. The recipients can view the received files at that particular time or later as desired by them.

18.7.1 Features of Messenger

A messenger is installed on the user's computer and it facilitates sending and receiving instant messages, graphical symbols, audio, and visual information. The features of messenger that make it a favourite among Internet users include:

- **Chatting**: With a messenger, one can chat using a webcam or microphone and send textual messages. In addition, users can find and meet people instantly.
- **Audibles**: Audibles are faces used to say something or express the attitude of the user in the form of greetings, games, and fun. These audibles are used for sending verbal animations directly from the instant messaging application.
- **Emoticons/Smileys**: The emoticons or smileys are used to express the feelings in the form of little symbols and smiling faces.
- **Games**: When connected online, users can play multi-player games such as chess, checkers, and many more. Users can even see how they stack up against their friends by comparing their scores.
- **Address Book**: Messengers also provide address books so that users can easily find and manage their friend's contact details.

Note: *In this book, we have used Yahoo Messenger 6.0 for explaining instant messaging concepts. Yahoo Messenger 6.0 is freely available from www.google.com.*

18.7.2 Getting Started with Yahoo Messenger

To use Yahoo Messenger, a unique Yahoo ID (such as *itlesl@yahoo.com*) is required. When you start the messenger, it will display a screen asking you to enter your yahoo mail ID and password (Figure 18.39).

Figure 18.39 Signing in to Yahoo Messenger

If you are using the messenger for the first time, you may wish to add the e-mail addresses of all those known ones who also use Yahoo messenger. Once you and your friends are online, you can chat and send instant messages. To add a friend to the list, follow the steps given below:

1. Click on the Add button on the toolbar of Yahoo Messenger.
2. Enter the e-mail address and mobile number (optional) of the person to be added in the list (Figure 18.40).

Figure 18.40 Add to Messenger List Dialog Box

3. The messenger will then ask you to type a message that will be sent to the person asking him/her to approve your request to add him/her to your list.
4. The messenger will then display that your request has been sent to the other person.
5. Click Finish to add that person in your Friends list Similarly, you can add other persons to the list (Figure 18.41).

Figure 18.41 Friends List

18.7.3 Sending Instant Messages

A message can be sent to a friend whose e-mail address you have stored in your Friends List. To send a message, the following steps should be followed:

1. Double-click on the name of a friend that you have added in the Friends List.
2. Type the message that you want to send and press Enter or click Send button.
3. If your friend is online, he/she may reply to your message (see Figure 18.42).

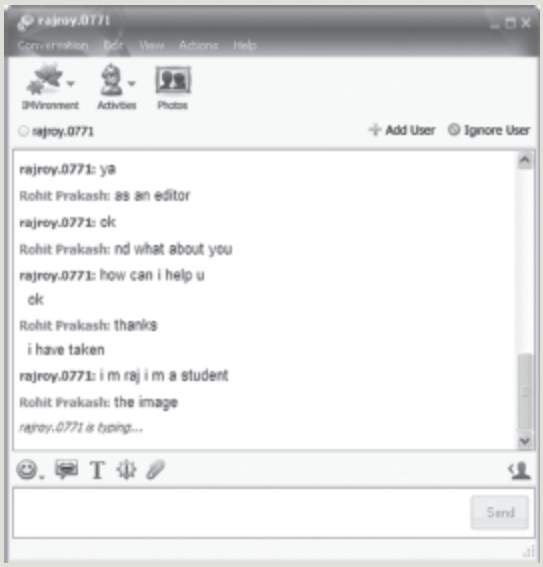


Figure 18.42 Instant Messaging

4. Along with your message, you can also send emoticons (smileys) or choose diffierent in-built backgrounds for the message. Several popular smileys are listed in Table 18.2.

Table 18.2 Smileys

Keyboard Short-cut	Smiley	Description	Keyboard Short-cut	Smiley	Description
:-) or :)		Happy	:-&		Sick
B-)		Cool	X(		Angry
:-O		Surprised	:O)		Clown
:-D		Laughing	<:-P		Party
8-}		Silly	:-?		Thinking
(:|		Yawn	:-SS		Nail Biting
=D>		Applause	:^O		Liar
:-(		Sad	:-}		Big Grin

18.7.4 Sending Files

In addition to the messages, you can also send and receive text, image, audio, or graphical files using the messenger. To send a file, follow the steps given below:

1. Choose Send a File option from the Actions menu or click Send a File button when the Message window is active (see Figure 18.42). This displays the Select the file to send dialog box (Figure 18.43).

Figure 18.43 Sending a File

2. Locate the file that is to be sent and click the Open button. This will return you to the Message window, which now displays the attached file also.
3. Type the description of the file in the box where you type the chat and click the **Send** button to complete the process.

LET US SUMMARIZE

1. A web browser (or simply browser) is a software application, which provides a GUI so that the user can navigate the Internet easily by clicking on menus, icons, or buttons instead of learning difficult keyboard commands.
2. There are many web browsers available having diffierent features. However, every browser shares some of the basic features like handling requests for HTML files, interpreting links, saving web pages, connecting to an e-mail program, and so on.
3. The content area is the portion of the window that holds the web page present in the current tab. Text, images, animation, and links available on the web page appear in this area.
4. Status bar displays the current state of activity on the web page. The information that appears on the Status bar includes the item currently being downloaded, the current web page's download progress, error (if any) in downloading, and zoom level of web page.
5. Browsing is an act of looking for information on the Internet by repeatedly going through several web pages usually with the help of hyperlinks. It is also known as navigating or surfing the Web.
6. Many web pages have links to files (like .ZIP or .EXE files) that are not natively displayed in the Internet Explorer content area. Such files have to be copied on the hard disk so that the user can use them.
7. Internet Explorer uses a feature called Favorites to put websites, one visits often, within easy reach. That way, a website can be opened simply by a click without the user having to remember or type anything.
8. Electronic mail can be defined as the process of exchanging messages electronically, via a communications network, using the computer. E-mails allow users to communicate with each other in less time and at nominal cost as compared to traditional telephone or mail services.
9. In order to use e-mail, one must have access to the Internet and an e-mail account. An e-mail account is a service that allows the user to send and receive e-mails through the Internet.
10. Netiquettes are rules that should be followed while communicating over the Internet. Disobeying netiquettes while sending e-mail can create trouble.
11. A newsgroup is a virtual place where people with common interests can ask questions and get answers on just about any imaginable topic ranging from sports to space. Newsgroups generate huge volumes of messages for which you need a newsreader to read.
12. To conduct a search on the Internet, a special search tool known as search engines is used. Search engine searches a database of Internet files collected by a computer program called a wanderer, crawler, or spider. It allows the user to enter

keywords relating to particular topics and retrieve information about the Internet sites containing those keywords.

13. Instant messaging is a web-based service to exchange messages in real time between two or more people over the Internet. It is a combination of e-mail and chat room. Usually, users send their messages in text format. However, while sending instant messages, a user can also communicate through webcam (video chatting) as well as through speakers and microphone (voice chatting).

EXERCISES

Fill in the Blanks

1. Internet Explorer is a ..... that allows user to communicate, locate, display,and use web documents.
2. A web browser uses the ..... protocol to request web pages from the web server.
3. The tool used for searching and getting refined information from the Internet is known as .....
4. To check e-mails, the user has to click ..... folder to view the received messages.
5. Internet Explorer uses a feature called ..... to put websites, one visits often,within easy reach.
6. The standard protocol used for sending e-mail is called .....
7. ..... allows user to send a file with e-mail.
8. ..... allows users to exchange messages in real time with other people over the Internet.
9. ..... shows the live images of a friend chatting with you on either side.
10. The little animated symbols used to express your feelings in e-mail are known as .....

Multiple-choice Questions

1. Which of the following symbols must an e-mail address have?
    1. _
    2. &
    3. @
    4. None of these
2. Usually the e-mail address is in
    1. Proper case
    2. Lower case
    3. Upper case
    4. All of these
3. In which of the following ways, a browser supports multimedia data?
    1. Native support
    2. Plug-ins
    3. Helper programs
    4. All of these
4. In Outlook Express, which of the following folders stores the deleted e-mails?
    1. Inbox

2. Remove items
3. Sent
4. Deleted items

5. ISP stands for

    1. Internet service promoter
    2. Internet service provider
    3. Intranet service provider
    4. None of these

6. Which one of the following is not present in the Compose Mail page of Yahoo?

    1. To
    2. Cc
    3. Content
    4. Subject

7. The verbal animations expressing feelings of a person in messengers are known as

    1. Smileys
    2. Audibles
    3. Emotion faces
    4. Audio faces

8. Which of the following buttons does not appear on the Command bar?

    1. Home
    2. Refresh
    3. Print
    4. Read Mail

9. The practice of attacking people on Internet is known as

    1. Spamming
    2. Flaming
    3. Hacking
    4. Fighting

10. Which of the following connectors is not used for searching?

    1. *
    2. {}
    3. " "
    4. ()

State True or False

1. A newsgroup is a virtual place where people with common interests can ask questions and get answers on just about any imaginable topic ranging from sports to space.
2. The e-mail sent by a user can only be web based.
3. E-mail address has no spaces in between.
4. Yahoo Messenger allows exchanging only text and not video or audio.
5. Search engine employs a special software called spiders to find information on the million of web pages on the Internet.
6. To view your friend on a webcam, it is not essential for you to have one.
7. Status bar displays the percentage of the page being downloaded, when the user is not using the Internet.
8. The e-mails sent to you are stored on the mail server of your ISP until you retrieve them.
9. Sending an e-mail is costlier than sending a message through post.

10. A user requires a microphone to chat with a friend on the other side.

Descriptive Questions

1. What is e-mail? Explain its working with the help of an example.
2. Describe some of the functions of Outlook Express and compare it with Internet Explorer?
3. What are search engines? How do they help users in using information on the Internet?
4. Describe some of the components of Internet Explorer.
5. Explain instant messaging with the help of a messenger.
6. Suppose you want to send an e-mail to one of your friends having e-mail address niceperson@indiatimes.com. Type a message, inviting him on your birthday (November 8 2004). Also attach a file named “Invitation Card” describing the venue and time of party.

ANSWERS

Fill in the Blanks

1. Browser
2. HTTP
3. Search engine
4. Inbox
5. Favorites
6. SMTP
7. Attachment
8. Instant messaging
9. Webcam
10. Smileys/Emoticons

Multiple-choice Questions

1. (c)
2. (b)
3. (d)
4. (d)
5. (b)
6. (c)
7. (b)
8. (b)
9. (b)
10. (b)

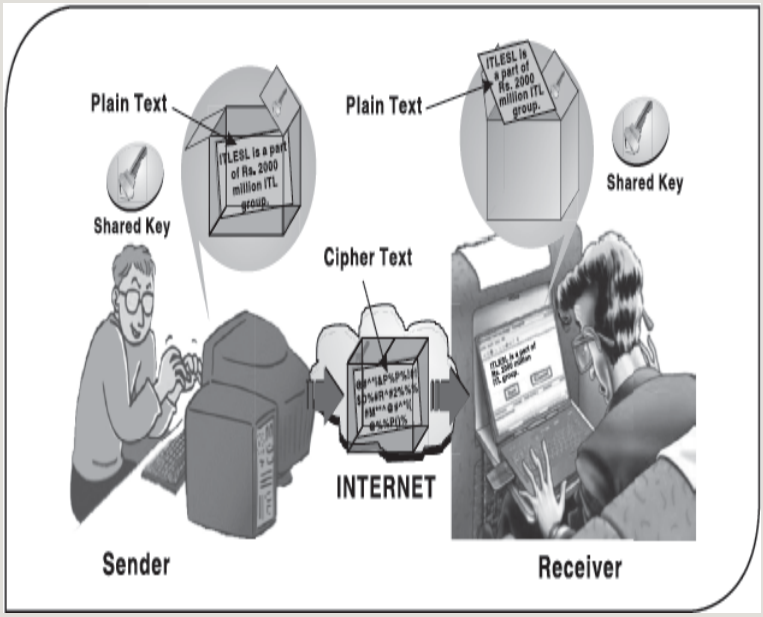
State True or False

1. True
2. False
3. True
4. False
5. True
6. True

7. False
8. True
9. False
10. False

Chapter 19

Computer Security



This chapter provides a foundation to understanding computer security. It is defined as technological and managerial procedures applied to the computer systems to ensure the availability, integrity and confidentiality of information against unwanted access, damage, and modification. The chapter begins with a brief overview of computer-related security aspects. It further gives an insight into the malicious programs that affect the computer system. The discussion continues on various techniques (cryptography, digital signature) that play an important role in transmitting data securely over the Internet. Firewall application along with its types and an overview on different ways of identification and authentication of the network is explained. Finally, the

chapter emphasises on training programs and policies for security.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- Computer security—a technological and managerial procedure applied to secure the computer system
- Malicious programs—unusual activity that penetrates along with the useful files
- Cryptography—the process of altering messages to hide their meaning from adversaries who might intercept them
- Digital signature—the digitized images of paper signature used to verify the authenticity of electronic document
- Firewall—an application that prevents certain outside connections from entering into the network
- Identification and authentication, which is another line of defence against the unauthorized people from entering into a computer system
- Data backup methods and recovery tools to ensure recovery of data in the event of data loss
- Security awareness and its related policies

19.1 INTRODUCTION

Security has always been an overriding concern of humankind. For any organization, information plays a fundamental role in running business. Therefore, it is required to safeguard information from reaching the illegal hands. When data are in the digital form, a different security procedure is required. Computer security refers to the technological and managerial procedures applied to computer systems to ensure the availability, integrity, and confidentiality of information managed by the computer system against unauthorized

access, modification, or destruction. It deals with the transmission of data in a secured environment to the people sitting thousand miles away from each other. The growth in the area of telecommunication and networking has enabled people to make use of electronic medium as the fastest way to stay in touch with each other. However, it has created unreliability and insecurity in the minds of people using it.

Information technology has also some loopholes associated with it like the possibility of stealing of vital information and intentionally implanting destructive programs onto the other's computer system. The motive behind such activities is to slow down the pace of an organization and harm it economically. Intruders penetrate into the computer using different ways; they make use of malicious programs to cause destruction and breach privacy. Therefore, it has become essential for an organization to take preventive measures to safeguard their data and privacy. Security experts make use of firewall and cryptography techniques to prevent suspicious data from reaching to the host computer and use algorithms to encrypt the data while sending it across the network.

19.2 COMPUTER SECURITY: DEFINITION

Computer security refers to the protection given to computers and the information contained in them from unauthorized access. The practice of computer security also includes policies, procedures, hardware, and software tools that are necessary to protect the computer systems and the information processed, stored, and transmitted by the systems. It involves the measures and controls that ensure *confidentiality, integrity,* and *availability* of the information, processed and stored by a computer. These three aspects are responsible for effective computer security (Figure 19.1).

Figure 19.1 Computer Security

Confidentiality: Confidentiality ensures that information is available only to those persons who are authorized to access it. Strict controls must be implemented to ensure that only those persons who need access to certain information have that access. The most common form of access control is the use of passwords. Requiring passwords, smart cards, or single-use-password devices is the first step to prevent unauthorized individuals from accessing sensitive information and is the first layer of defence in access control. Therefore, keeping password confidential is one of the most fundamental principles of computer security.

Integrity: Integrity ensures that information cannot be modified in unexpected ways, as loss of integrity could result from human error, intentional tampering, or even catastrophic events. The consequence of using inaccurate information can be disastrous; therefore, an effort must be made to ensure the accuracy and integrity of data at all times. When the validity of information is of utmost importance, it is often helpful to design controls and checks to ensure accuracy of information. For this,

encryption process is used, which transforms information into some secret form to prevent unauthorized individuals from accessing the data. Such a technique prevents an intruder from reading or modifying the information.

Availability: Availability prevents resources from being deleted or becoming inaccessible. This applies not only to information, but also to the machines on the network and other aspects of the technology infrastructure. This inability to access the required resources is called “denial of service.” Intentional attacks against computer systems often aim to disable access to data. Another aspect of availability ensures that needed resources are usable when and where needed, thus providing system redundancy, in the form of backup data and power source.

19.2.1 Security Threats

Computer systems are vulnerable to many kinds of threats that can cause various types of damages, which may result in significant data loss. These damages can range from errors that can cause harm to database integrity to natural calamity destroying entire computer centres. Precision in estimating computer security-related losses is not possible because many losses are never discovered. The effects of various threats vary considerably; some affect the confidentiality or integrity of data while others affect the availability of a system. A threat can come from any person, object, or event that, if realised, could potentially cause damage to the computer network. It can also arise from intentional modification of sensitive information or accidental error in a calculation or because of a natural disaster like flood, storm, or fire. Some of the commonly occurring threats to a network are discussed below.

THINGS TO REMEMBER

Errors and Omissions

User, system operators, and programmers frequently make unintentional errors, which contribute to security problems directly and indirectly resulting in system crashes. In other cases, errors create vulnerabilities. It can occur in all phases of the system life cycle. Programming and development errors, often called bugs, range in severity from gentle to catastrophic. In the past decade, software quality has improved measurably to reduce this threat, yet software "errors" still bound. Errors and omissions are important threats to data integrity.

Errors and Omissions: Errors and omissions are important threats to data and system integrity. These errors are caused not only by data entry operators, processing hundreds of transactions per day, but also by users who create and edit data. Many programs, especially those designed by users for personal computers, lack quality-control measures. However, even the most sophisticated programs cannot detect all types of input errors or omissions. A sound awareness and training program can help an organization reduce the number and severity of errors and omissions.

Fraud and Theft: Information technology is increasingly being used to commit fraudulent and theft activity. Computer systems are exploited in numerous ways, both by automating traditional methods of fraud and by using new methods. For example, individuals may use a computer to steal money from a large number of financial accounts, thus generating a significant sum

for their own use. Financial systems are not the only institutions facing fraudulent activity, but the systems such as time and attendance systems, inventory systems, school grading systems, or long-distance telephone systems, which control access to any resource, are also the targets.

The majority of fraud uncovered on computer systems is committed by insiders who are authorized users of a system. Since insiders have both access to and familiarity with the victim computer system, including what resources it controls and where the flaws are, they are in a better position to commit crimes. An organization's former employees may also pose threats, particularly if their access is not terminated promptly.

Loss of Physical and Infrastructure Support: Loss of physical and infrastructural support in an organization also contributes to the security threat. The infrastructural support includes power failures, loss of communications, water outages and leaks, lack of transportation services, natural calamity, and so forth. Recent study has shown that more loss is associated with fires and floods than with viruses and other more widely publicized threats. A loss of infrastructure often results in system downtime, sometimes in unexpected ways.

Hacker and Cracker: The term hacker refers to the person with the intention of finding some weak points in the security of websites and other computer systems in order to gain unauthorized access. The activities of hackers are not limited to only gaining the unauthorized access to systems, but also include stealing and destroying the confidential information. Hackers can also introduce viruses in the network, which can enter database or other applications and crash the whole server. In addition, they can also modify links in websites

to redirect the sensitive information to the database of their interests.

In hacking community, hackers have been classified into two categories, namely, *white-hat hackers and black-hat hackers* depending on their intent behind hacking. The hackers who break into the computer security with non-malicious reasons are known as white-hat hackers. Usually, such hackers are security experts working with manufacturers. On the other hand, the hackers who break into the computer security without authorization for ulterior purposes, such as property theft, credit card theft, terrorism, etc., are known as black-hat hackers or *crackers*.

Note: *In mass media, the terms hacker and cracker are often used interchangeably.*

Malicious Code: *Malicious codes* are the programs that generate threats to the computer system and precious data. This code can be in the form of viruses, worms, Trojan horses, logic bombs, and other "uninvited" software. *Virus* is a small segment of code, which replicates by attaching copies of itself to existing executables files. The new copy of the virus is executed when a user executes the new host program. *Trojan horse* is a program that performs a desired task, but also includes unexpected functions. *Worm* is a self-replicating program, which is self-contained and does not require a host program. The program creates a copy and causes it to execute; no user intervention is required. Worms commonly utilize network services to propagate other host systems.

The number of known viruses and worms is increasing, and the rate of virus attacks is growing alarmingly. Most organizations use Antivirus software and other protective measures to limit the risk of virus infection.

Foreign Government Espionage: In some instances, threats can be posed by foreign government intelligence services. In addition to possible economic espionage, foreign intelligence services may target unclassified systems to collect information about intelligence missions. Some unclassified information that may be of interest includes travel plans of senior officials, defence, and emergency preparedness, manufacturing technologies, satellite data, personnel and payroll data, investigative, and security files. Therefore, adequate guidance must be sought from the security office regarding such threats.

19.3 MALICIOUS PROGRAMS

It is often seen that users are made to open certain files (screen saver, game, utility, and so on) on the Internet. On opening such files, a user begins to face unusual activities from the computer such as malfunctioning of the applications or inefficient running of hardware resources and so on. Such unusual activity is the result of malicious programs, which penetrate along with the useful file. These malicious programs are often called virus, worms, Trojan horse, logic bomb, and so on.

19.3.1 Virus

Viruses are programs, which are designed to replicate, attach to other programs, and perform unsolicited and malicious actions. It executes when an infected program is executed. On MS-DOS systems, these files usually have the extensions *.exe, .com*, or *.bat*. Virus enters computer systems from an external software source and easily hides in software. Just as flowers are attractive to the bees that pollinate them, virus host programs are deliberately made attractive to victimize the user. They become destructive as soon as they enter a system or are programmed to lie dormant until activated by a trigger. The different types of virus are discussed below.

Boot Sector Virus *Boot sector virus* infects the master boot record of a computer system. This virus either moves the boot record to another sector on the disk or replaces it with the infected one. It then marks that sector as a bad spot on the disk. This type of virus is very difficult to detect since the boot sector is the first program that is loaded when a computer starts. In effect, the boot sector virus takes full control of the infected computer (Figure 19.2).

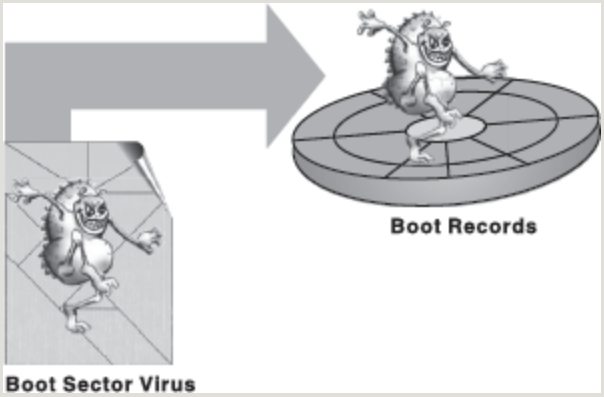


Figure 19.2 Boot Sector Virus

File-Infecting Virus: *File-infecting virus* infects files with extension *.com* and *.exe*. This type of virus usually resides inside the memory and infects most of the executable files on a system. The virus replicates by attaching a copy of itself to an uninfected executable program. It then modifies the host programs and subsequently, when the program is executed, it executes along with it. File-infecting virus can only gain control of the computer if the user or the operating system executes a file infected with the virus (Figure 19.3).

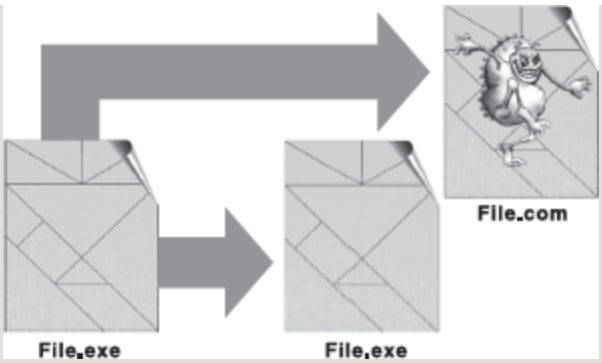


Figure 19.3 File Infecting Virus

Polymorphic Virus: *Polymorphic virus*, unlike other viruses, consists of static virus program that gets copied from file to file as it propagates. Such virus is difficult to detect because each copy it generates appears different from the other one. It uses encryption algorithm to multiply new copies of the program. For an encrypted virus to execute, it must decrypt the encrypted portion of itself. When an infected program launches, the virus decryption routine gains control of the computer and decrypts the rest of the virus body so that it can execute normally (Figure 19.4).

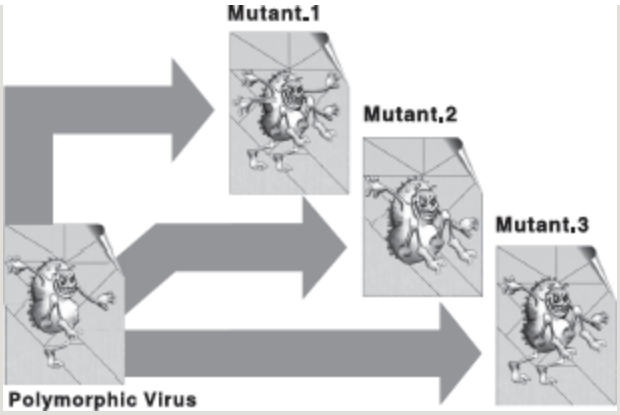


Figure 19.4 File Infecting Virus

Stealth Virus: *Stealth virus* attempts to conceal its presence from the user. It reads system files or system sectors and when some other program requests for information from portions of the disk, it changes back into the correct (unchanged) form. Use of stealth virus is the major reason why most antivirus programs operate best when the system is started (booted) from a known-clean floppy disk. When this happens, the virus does not gain control over the system and is immediately available to be seen and dealt with. The Stoned Monkey virus is an example of stealth virus. This virus uses “read stealth” capability and if a user executes a disk editing utility to examine the main boot record, the user would not find any evidence of infection (Figure 19.5).

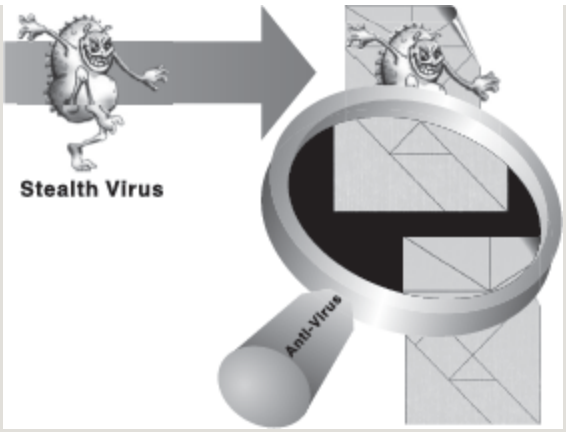


Figure 19.5 Stealth Virus:

Multipartite Virus: *Multipartite virus* infects both boot sectors and executable files and uses both mechanisms to spread. It is the worst virus of all because it can combine some or all of the stealth techniques along with polymorphism to prevent detection. For example, if a user runs an application infected with a multipartite virus, the virus activates and infects the hard disk's master boot record. Moreover, the next time the computer starts; the virus gets activated again and starts infecting every program that the user runs. One-half virus is an example of a multipartite virus, which exhibits both stealth and polymorphic behaviour (Figure 19.6).

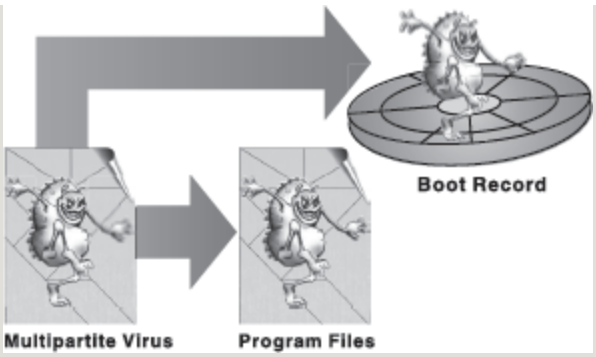


Figure 19.6 Multipartite Virus

19.3.2 Other Destructive Programs

Apart from viruses, other threats, which harm computer, are *worms, Trojan horses*, and *logic bombs*. Each of these programs can also be used as a medium to propagate any kind of virus.

Worms: *Worms* are the programs constructed to infiltrate on the legitimate data processing programs and alter or destroy the data. They often use network connections to spread from one computer system to another, thus, worms attack systems that are linked through communication lines. Once active within a system, worms behave like a virus and perform a number of disruptive actions. To reproduce themselves, worms make use of network medium such as:

THINGS TO REMEMBER

Worms

A worm with harmful objectives could perform a wide range of destructive activities such as deleting files on each affected computer. On October 16, 1989, a worm named **WANK** infected many VAX and VMS computers on the SPAN network. This

worm changed the system announcement message to “Worms Against Nuclear Killers!” The message was then graphically displayed as the first letters of each word and the last three letters of the last word.

- *Network mail facility*, in which a worm can mail a copy of itself to other systems.
- *Remote execution capability*, in which a worm can execute a copy of itself on another system.
- *Remote log in capability*, whereby a worm can log into a remote system as a user and then use commands to copy itself from one system to another.

The worm's replication mechanism can access the system by using any of the three methods given below:

- It employs password cracking, in which it attempts to log into systems using different passwords such as words from an online dictionary.
- It exploits a trap door mechanism in mail programs, which permits it to send commands to a remote system's command interpreter.
- It exploits a bug in a network information program, which permits it to access a remote system's command interpreter.

By using a combination of these methods, the network worm is able to copy itself to different computers, which are using similar versions of the operating system.

Trojan Horses: The term “Trojan Horse” is an ancient Greek mythology. In the war between Greeks and Troy, the Greek army barricaded the city of Troy but was unable to penetrate inside the city. Therefore, they decided to trick the enemy by building a large wooden horse with soldiers hidden secretly inside it and presenting it as a gift to the citizen of Troy. During night, the warriors came out from the horse and overran the city.

In computer terminology, there are programs, which perform similar destructive activities. These programs enter into a computer through an e-mail or free programs that have been downloaded from the Internet. Once safely get into the computer, it usually opens the

way for other malicious software (like viruses) to enter into the computer system. In addition, it may also allow unauthorized users to access the information stored in the computer.

Trojan horses spread when users are convinced to open or download a program because they think it has come from a legitimate source. They can also be included in software that are freely downloadable. They are usually subtler especially in the cases where they used for espionage. They can be programmed to self-destruct, without leaving any evidence other than the damage they have caused. The most famous Trojan horse is a program called *back orifice*, which is an unsubtle play of words on Microsoft's Back Office suite of programs for NT server. This program allows anybody to have the complete control over the computer or server it occupies.

Logic Bombs: *A logic bomb* is a program or portion of a program, which lies dormant until a specific part of program logic is activated. The most common activator for a logic bomb is a date. The logic bomb checks the computer system date and does nothing until a pre-programmed date and time is reached. It could also be programmed to wait for a certain message from the programmer. When logic bomb sees the message, it gets activated and executes the code. A logic bomb can also be programmed to activate on a wide variety of other variables such as when a database grows past a certain size or a user's home directory is deleted. For example, the well-known logic bomb is a *Michelangelo*, which has a trigger set for Michelangelo's birthday. On the given birth date, it causes system crash or data loss or other unexpected interactions with existing code.

19.3.3 Affecting Computer System

To safeguard computer system against viruses, it is important to understand how virus spreads and what

they do to infect the computer system.

How Virus Spreads? Virus is designed to proliferate and propagate in computer network. This means any contact between two or more computers is an opportunity for infection. Unauthorized users break into the system and easily cause destruction by planting virus in the most sensitive locations of the computer. Viruses come through many sources and being the software code, they can be transmitted with any other software, for example, on a disk, through network, or using e-mail.

With the advent of the Internet, people can now access and exchange information and resources. Most people do so using the World Wide Web, FTP, or e-mails. These viruses are predominantly in the form of macro viruses incorporated into Microsoft Word and Excel documents. A new type of virus infects the computer by seeing the "preview" in the e-mail client of Microsoft Outlook (or Outlook Express).

Virus infection is growing exponentially, and there have been several incidents involving e-mail-borne viruses such as *Melissa, Code Red, Happy99*, and many more. Viruses are never likely to go away; therefore, all users need to be vigilant and know how they should combat infection.

System Components Affected by Virus: These two conditions can cause a virus to attack on the computer system.

- Executable file on the storage disk.
- Write-enabled disk (floppy disk, hard disk)

The simple act of write-protecting floppy disks by covering the notch in 5¼ disks or opening the hole in 3½ disks can prevent many virus infections. Unfortunately, this procedure does not always work. Viruses can hide in a number of places on a disk. If a

virus simply attaches itself to an executable file stored on a hard disk or a floppy disk, change in the file size can help detect the presence of virus. Another technique used is the *checksum*. A checksum adds up various file characteristics to produce a hexadecimal number. Unfortunately, smart viruses record the original checksum and report it when program is tested, rendering this test null and void. Viruses, which locate in the logical drive of bootstrap program, are more difficult to trace. Another possible location for virus to reside is on the physical sector of the hard disk. Virus also modifies directory entries for DOS format and installs a copy in the directory that executes when other programs run. These linked viruses change directory listings so that they can hide themselves from simple directory display commands. Simply scanning for executable files whose size has been changed since the last scan is not sufficient. Hidden files, system files, and areas of the disk normally invisible to software must also be explored.

19.3.4 Protecting Computer System

Different organizations have different styles of operation. This fact extends to the way they set up their computer networks and operating procedures. If computer operations consist of one or two personal computers used by fewer people, then the need for an elaborate defence system is not of utmost importance. However, if system is large enough to include the worldwide networks used by large corporations, then a detailed and systematic defence system is required.

Using Antivirus Software: In the early days of computer networking, computers were not networked very well, and computer viruses spread extremely slowly. Files were transmitted by means of BBSs (Bulletin Board Systems) or on diskette. As a result, the transmission of infected files was not that fast and easy. However, as the connectivity improved, mostly by the use of computers in

the workplace, the scope of virus threat widened. First, there was LAN, and then there was WAN and now the Internet. The extensive use of e-mail has also contributed to the significant rise in the number of virus incidents. As a result, probability of getting infected by a virus today is more than it was a few years ago.

Until the preventive measures for virus were not invented, the only option to get rid of them was to get rid of that valuable file which was infected. This proved very costly to the companies whose valuable work was destroyed due to viruses. As the computer technology gained heights, these viruses also advanced in causing destruction. Then came a software utility called *Antivirus* to the rescue. Antivirus is a software utility, which (upon installing on a computer) scans the hard disk for viruses and tries to remove them if found. Most Antivirus programs include an auto-update feature that enables the program to download profiles of new viruses so that it can check for the new viruses as soon as they are discovered. The most popular Antivirus software available include *Norton Antivirus, McAfee Antivirus, and Avira AntiVir*.

THINGS TO REMEMBER

Antivirus

Antivirus software has normally a built-in scanner, which scans all files on the computer's hard disk. It looks for changes and activities in computers typical in case of a virus attack. They look for particular types of code within programs. The software generally relies on having prior knowledge of the virus. As a result, frequent update to the tools is necessary. The important thing is to be aware of the

possibility of an attack, to possess a good virus checking software, and to have data backups.

Figure 19.7 depicts a typical virus detection mechanism used by an Antivirus program. The image illustrates that if an Antivirus program is not installed on the computer, the virus in the e-mail gets into the computer. However, once an Antivirus program is installed in the computer, it checks all the incoming files (mails), detects the viruses, and removes them before storing the files onto the user's machine. However, in some cases, Antivirus is not able to remove the virus and thus, we have to delete that file.

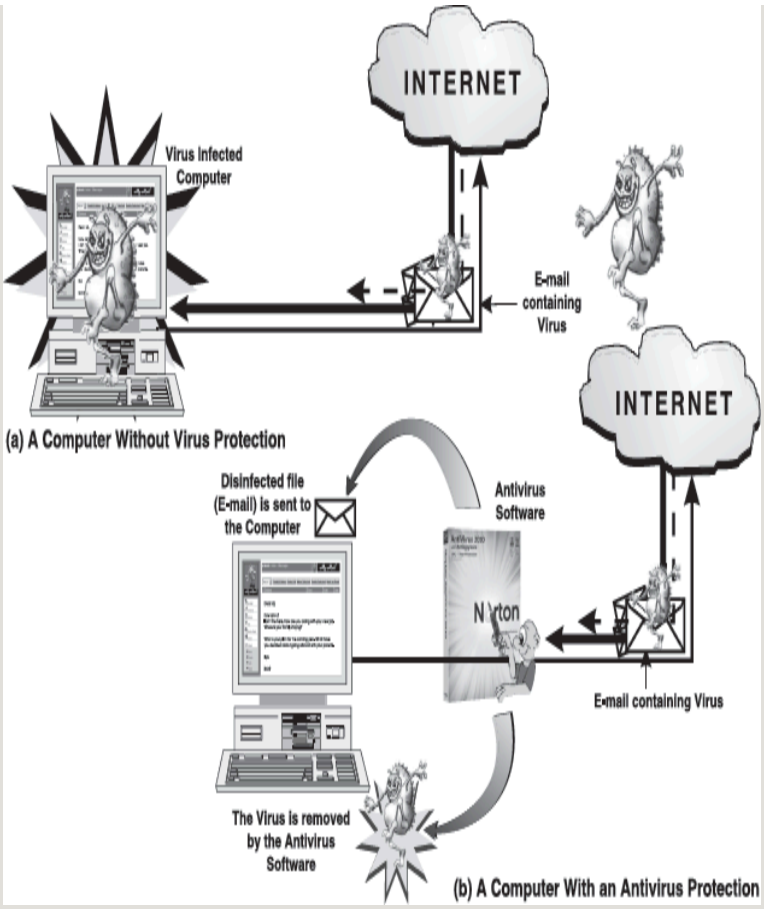


Figure 19.7 Virus Protection Using Antivirus

19.4 CRYPTOGRAPHY

When data are sent over the Internet, it passes through many different servers before it reaches the desired destination. This data remain on servers for months and at any stage of journey, it is vulnerable to interception. Most of the time, this may not concern the user, but there are times when sensitive data (business quote or personal details) are transmitted. Therefore, the best way is the use of an encryption program or the cryptography technique. In simple terms, *cryptography* is the process of altering messages to hide their meaning from adversaries who might intercept them. In data and telecommunications, cryptography is an essential

technique required for communicating over any untrusted medium, which includes any network, particularly the Internet. Cryptography provides an important tool for protecting information and is used in many aspects of computer security. Cryptography is traditionally associated only with keeping the secrecy of data. However, nowadays these are used to provide many security services such as electronic signatures and non-modification of data. Cryptography relies upon two basic components: an *algorithm* (or cryptographic methodology) and a *key*. Communication over the Internet, for example, e-mail, is not secure if no encryption is used. Hackers may be able to read messages, or even modify it if cryptography or any other medium is not present.

In modern cryptographic systems, algorithms are the complex mathematical formulae and keys are strings of bits. For two parties to communicate over a network (Internet), they must use the same algorithm (or algorithms that are designed to work together). In some cases, they must also use the same key. In all cases, the initial unencrypted data are referred to as *plain text*, which is encrypted into *cipher text*, and in turn decrypted into usable plain text. Cryptography techniques are broadly classified into three types:

- **Secret Key Cryptography (SKC)**: Uses a single key for both encryption and decryption.
- **Public Key Cryptography (PKC)**: Uses one key for encryption and another for decryption.
- **Hash Functions**: Uses a mathematical transformation to irreversibly "encrypt" information (Figure 19.8).

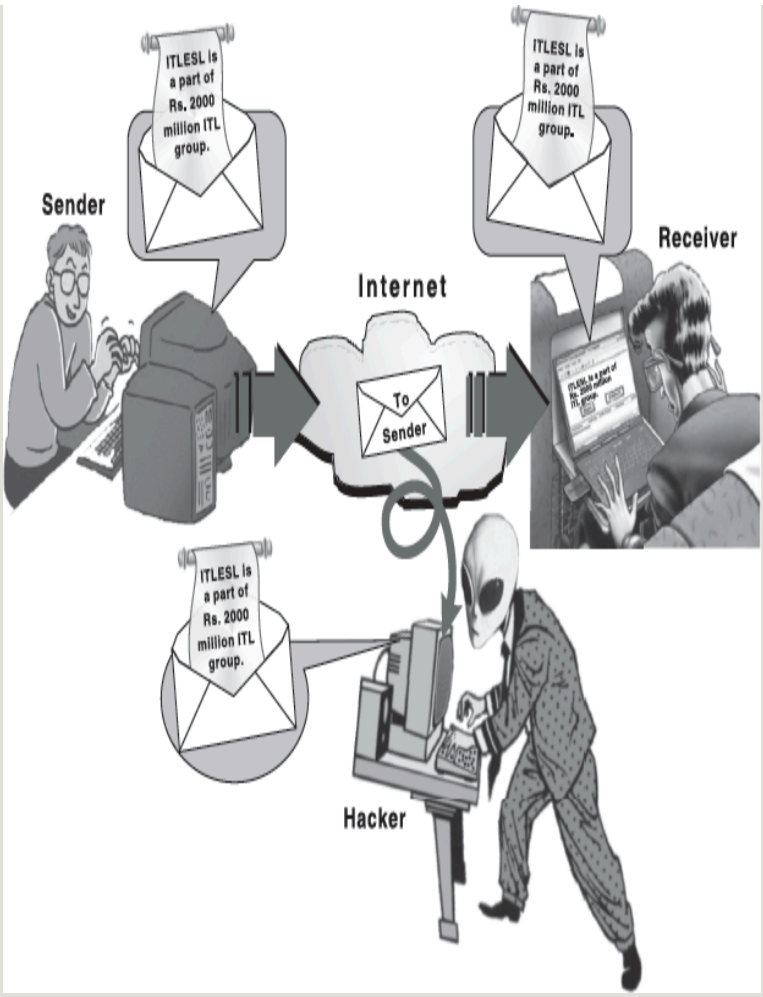


Figure 19.8 Message Exchange without Encryption

19.4.1 Secret Key Cryptography

With secret key cryptography (SKC), a single key is used for both encryption and decryption of data. With this form of cryptography, it is obvious that the key must be known to both the sender and the receiver. If the key is compromised, the security offered by SKC is severely reduced or eliminated. SKC assumes that the parties who share a key rely upon each other not to disclose the key and protect it against modification. SKC schemes are generally categorized as being either *stream ciphers or block ciphers*. Stream ciphers operate on a single bit at a

time and implement some form of feedback mechanism so that the key is constantly changing. On the other hand, a block cipher scheme encrypts one block of data at a time using the same key on each block. In general, the same plain text block will always encrypt to the same cipher text when using the same key in a block cipher whereas the same plain text will encrypt to different cipher text in a stream cipher.

As shown in Figure 19.9, the sender uses the key to encrypt the plain text and sends the cipher text to the receiver. The receiver applies the same key to decrypt the message and recover the plain text. As there is a usage of single key for both functions, SKC is also called *symmetric encryption*.

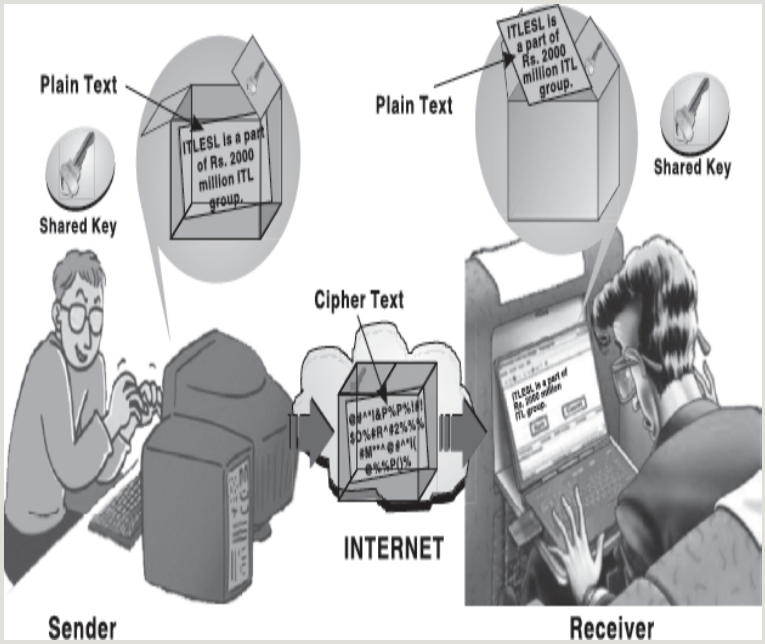


Figure 19.9 Message Exchange Using Secret Key

The main problem in SKC is getting the sender and receiver to agree on the secret key without anyone else finding out. If they are in separate physical locations, they must trust on a courier, or a telephone system, or any other transmission medium to prevent the disclosure of the secret key being communicated. Anyone who overhears or intercepts the key in transit can later read, modify, and forge all messages encrypted or authenticated using that key.

19.4.2 Public Key Cryptography

This concept was introduced in 1976 by Whitfield Diffie and Martin Hellman of Stanford University to solve the problem, found in secret key-based cryptography. In this technique, each person gets a pair of keys, one called the *public key* (used for encryption) and the other called the *private key* (used for decryption). Each person's public key is published while the private key is kept secret. The need for the sender and receiver to share secret information is eliminated and all communications involve only public keys, and no private key is ever transmitted or shared. No longer is it necessary to trust some communication channel to be secure against betrayal. The only requirement is that public keys are associated with their users in a trusted (authenticated) manner (for instance, in a trusted directory). Anyone can send a confidential message by using public information, but the message can only be decrypted with a private key, which is in the sole possession of the intended recipient.

When a sender wants to send a secret message to the receiver, the sender uses receiver's public key to encrypt the message. When the receiver receives the encrypted message through the Internet, he/she uses his/her private key to decrypt it. Figure 19.10 shows an overview of the whole process.

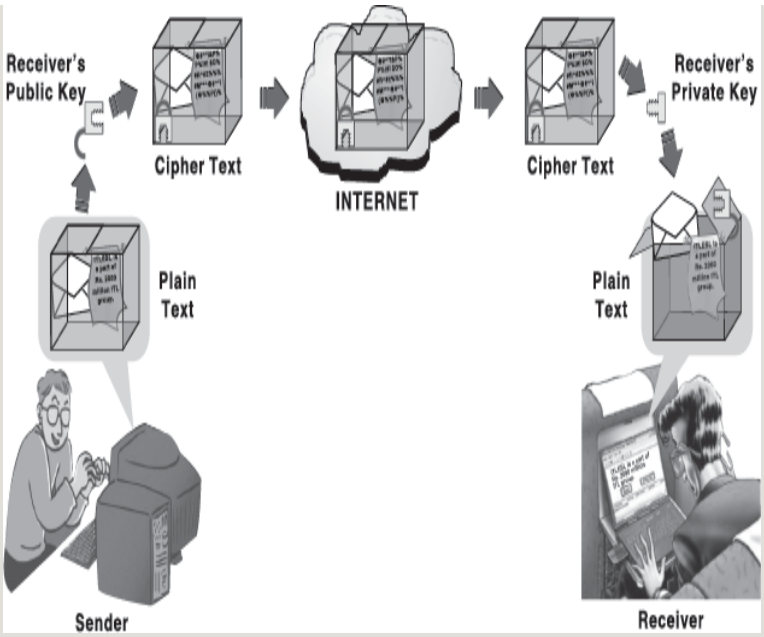


Figure 19.10 Message Exchange Using Public Key

19.4.3 Hash Function

A *hash function* is a one-way encryption algorithm that does not use any key to encrypt or decrypt the data. This technique generates a sequence of bit values of fixed length from the original message. This particular technique works on a fixed-length hash value, which is computed based upon the plain text that makes it impossible for either the contents or length of the plain text to be recovered. Hash algorithm typically uses digital fingerprint of a file's contents, which is often used to ensure that the file is not altered by an intruder or virus. Hash functions are also commonly employed by many operating systems to encrypt passwords and preserve the integrity of a file (Figure 19.11).

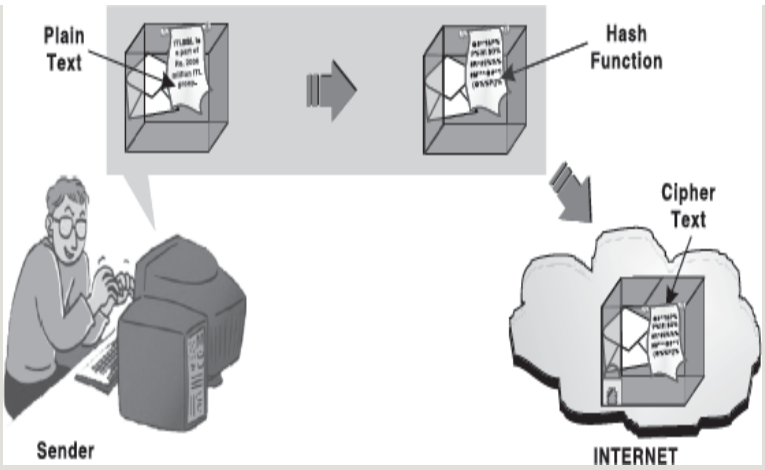


Figure 19.11 Message Exchange Using Hash Function

19.5 DIGITAL SIGNATURE

In today's commercial environment, establishing a framework for the authentication of computer-based information requires familiarity from both the legal advisor and computer security fields. Combining these two disciplines is not an easy task. The historical legal concept of “signature” is defined as any mark made with the intention of authenticating the marked document. Digital signature refers to the digitized images of paper signature used to verify the authenticity of electronic document. In other words, digital signatures play the role of physical signatures in verifying electronic documents.

To understand the concept of digital signature in a better way, we must first know the legal implications of digital signature. A signature is not part of the substance of a transaction, but is a representation. Signature serves the following general purposes:

- **Evidence**: A signature authenticates writing by identifying the signer with the signed document. When the signer makes a mark in a distinctive manner, the writing becomes attributable to the signer.
- **Ceremony**: The act of signing a document calls to the signer's attention towards the legal significance of the signer's act, and thereby

helps preventing inconsiderate engagements.

- **Approval**: In certain contexts defined by law or custom, a signature expresses the signer's approval or authorization of the writing, or the signer's intention that it has legal effect.
- **Efficiency and Logistics**: A signature on a written document often imparts a sense of clarity and finality to the transaction and may lessen the subsequent need to inquire beyond the face of a document.

Although the basic nature of transactions has not changed, the cyber law has only begun to adapt new technology. The legal and business communities must develop rules and practices, which use new technology to achieve and surpass the effects historically expected from paper forms. To achieve the basic purposes of signatures outlined above, a signature must have the following attributes:

- **Signer Authentication**: A signature should indicate who signed a document, message, or record and should be difficult for another person to produce without authorization.
- **Document Authentication**: A signature should identify what is signed, making it impracticable to falsify or alter either the signed matter or the signature without detection.

19.5.1 How Is Digital Signature Used?

Digital signatures use public key cryptography (PKC) technique, which employs an *algorithm* using two different but mathematically related *keys*; one for creating a digital signature or transforming data into a seemingly unintelligible form, and another key for verifying a digital signature or returning the message to its original form. Computer equipment and software utilizing two such keys are often collectively termed as an *asymmetric cryptosystem*.

The complementary keys of an asymmetric cryptosystem for digital signatures are arbitrarily termed the *private key*, which is known only to the signer and is used to create the digital signature, and the *public key*, which is ordinarily more widely known and is used by a relying party to verify the digital signature. If many people need to verify the signer's digital signatures, the public key

must be available or distributed to all of them. Although the keys of the pair are mathematically related, if the asymmetric cryptosystem has been designed and implemented securely, it is computationally infeasible to derive the private key from the knowledge of the public key. Thus, many people may know the public key of a given signer and use it to verify that signer's signatures; they cannot discover the signer's private key and use it to forge digital signatures.

Another fundamental process, termed as *hash function*, is used in both creating and verifying a digital signature. A hash function is an algorithm, which creates a digital representation in the form of a hash value of a standard length, which is usually much smaller than the message but nevertheless, substantially unique to it. Thus, use of digital signatures usually involves two processes, one performed by the signer and the other by the receiver of the digital signature. These are:

- **Digital Signature Creation**: It uses a hash result derived from and unique to both the signed message and a given private key. For the hash result to be secure, there must be only a negligible possibility that the same digital signature could be created by the combination of any other message or private key.
- **Digital Signature Verification**: It is the process of checking the digital signature by reference to the original message and a given public key, thereby determining whether the digital signature was created for that same message using the private key that corresponds to the referenced public key.

To have a clear understanding of how digital signature is applied, let us consider an example. Suppose Mr. A wants to send his signed message to Mr. B through Internet, he can use the public key cryptosystem to provide digital signatures. Mr. A creates the digital signature using his private key. This process makes message in encrypted form and results as signed message. The signed message is then sent through the Internet to Mr. B (Figure 19.12).

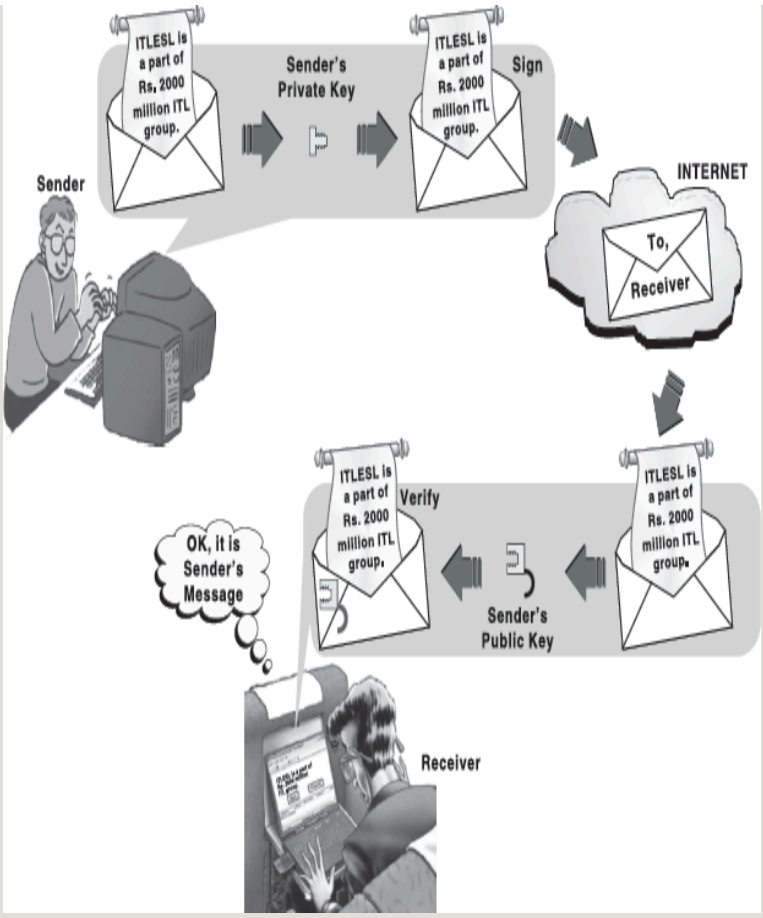


Figure 19.12 Message Exchange Using Digital Signature

Mr. B receives the message along with the digital signature from the Internet. Mr. B gets a copy of Mr. A's public key and using this key, he verifies the signature by A. This process results in the decryption of the message with A's public key. If the decrypted result is the same as that of transmitted message, Mr. B can believe that the message has really come from Mr. A. This is because only Mr. A holds his private key, which is needed to generate the digital signature.

Note that if the message or the digital signature is modified during the transmission, the decrypted form of

digital signature will not match with the message. From here, Mr. B can conclude that either the message transmission has tampered or the message has not been generated by Mr. A.

19.6 FIREWALL

The ongoing occurrences of incidents pertaining to network security caused a great concern to the people, using computers as their medium to exchange data across the country. A need was felt for a method of controlling the traffic, which allows access of information to computers. Organizations required an application that could protect and isolate their internal systems from the Internet. This application is called *Firewall*. Simply put, a firewall prevents certain outside connections from entering into the network. It traps inbound or outbound packets, analyses them, and then permits access or discards them (Figure 19.13).

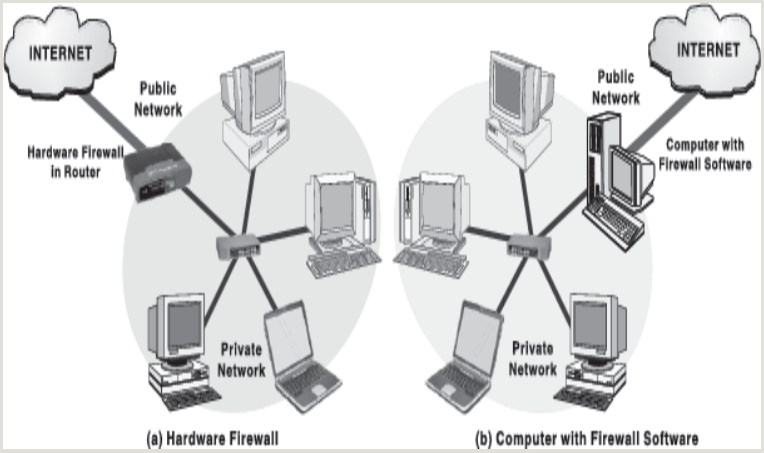


Figure 19.13 Firewall Software and Hardware

THINGS TO REMEMBER

Firewall

Firewall makes decisions on whether data should be allowed to pass based upon the security policy. For each packet of data, the firewall compares known components of the packet to a security rule set and decides if the packet should be allowed to pass. In addition, firewall has security rules that involve altering the packet in some basic way before passing the data. With a sensible security policy and a security rule set designed to implement that policy, a firewall is used in protecting local area networks from attacks.

Generally, firewall system comprises software (embedded in a router), computer, host, or a collection of hosts set up specifically to shield a site or subnet from protocols and services that can be a threat from hosts outside the subnet. It serves as the gatekeeper between an untrusted network (Internet) and the more trusted internal networks. If a remote user tries to access the internal networks without going through the firewall, its effectiveness is diluted. For example, if a travelling manager has an office computer that he or she can dial into while travelling, and his or her computer is on the protected internal network, then an attacker who can dial into that computer has circumvented the firewall. Similarly, if a user has a dial-up Internet account, and sometimes connects to the Internet from his or her office computer, he or she opens an unsecured connection to the Internet that circumvents the firewall.

19.6.1 How Firewall Works?

To understand the working of a firewall, consider an example where an organization is having hundreds of computers on the network. In addition, the organization will have one or more connections to the Internet lines. Now, without a firewall in place, all the computers are directly accessible to anyone on the Internet. A person who knows what other people are doing can probe those computers, try to make FTP (file transfer protocol) connections to them, or telnet connections, and so on. If one employee makes a mistake and leaves a security hole, hackers can get to the machine and exploit that hole.

With a firewall in place, the network landscape becomes much different. An organization will place a firewall at every connection to the Internet (for example, at every T1 line coming into the company). The firewall can implement security rules. For example, one of the security rules may be: out of the 300 computers inside an organization, only one is permitted to receive public FTP traffic. A company can set up rules like this for FTP servers, web servers, telnet servers, and so on. In addition, an organization can have control on how employees connect to websites, whether files can be sent as attachments outside the company over the network, and so on. Firewall provides incredible control over how people use the network. It provides protection against the following:

- Block unwanted traffic.
- Direct incoming traffic to more trustworthy internal systems.
- Hide vulnerable systems, which cannot be secured from the Internet.
- Log traffic to and from the private network.
- Hide information like system names, network topology, network device types, and internal user ID's from the Internet.
- Provide more robust authentication than standard applications.

One can also customize firewalls according to the specific needs. This means that one can add or remove filters based on several conditions:

- **IP Addresses:** IP addresses are 32-bit numbers, normally expressed as four "octets" between the periods. It is a unique address assigned on the Internet. A typical IP address looks like this: 135.27.84.129. If a certain IP address outside the company is reading too many files from a server, the firewall can block all traffic to or from that IP address.
- **Domain Names:** As numeric strings are hard to remember, which make up an IP address, all servers on the Internet also have human-readable names, called domain names. For example, it is easier for us to remember www.itlesl.com than 134.26.56.161. An organization might block all access to certain domain names or allow access only to specific domain names.
- **Ports:** Server machine, which makes services available to the Internet, uses numbered ports. For each service that is available on the server, a corresponding port number is assigned. For example, if a server machine is running a web (HTTP) server, the web server would typically be available on Port 80. A firewall can be configured to block Port 80 on all machines.
- **Specific Words and Phrases:** Firewall can be made to configure by sniffing (search through) each packet of information for an exact match of the text listed in the filter. For example, a user can instruct the firewall to block any packet with the word "ADWARE" in it.

19.6.2 Types of Firewall

A firewall intercepts the data between the Internet and the computer. All data traffic passes through it, and it allows only authorized data to pass into the corporate network. Firewalls are typically implemented using one of the three primary architectures: *packet filtering, application-level gateway, and circuit-level gateway*.

Packet Filtering: Packet filtering is the most basic firewall protection technique used in an organization. It operates at the network layer to examine incoming and outgoing packets and apply a fixed set of rules to the packets to determine whether they will be allowed to pass. The packet filter firewall is typically very fast because it does not examine any of the data in the packet. It simply examines the IP packet header, the source and destination IP addresses, and the port combinations and then it applies filtering rules. For example, it is easy to

filter out all packets destined for Port 80, which might be the port for a web server. The administrator may decide that Port 80 is off limits except for specific IP subnets, and a packet filter would suffice for this. Packet filtering is fast, flexible, transparent (no changes are required at the client), and cheap. This type of filter is commonly used in small-to-medium businesses that require control over users to use the Internet (Figure 19.14).

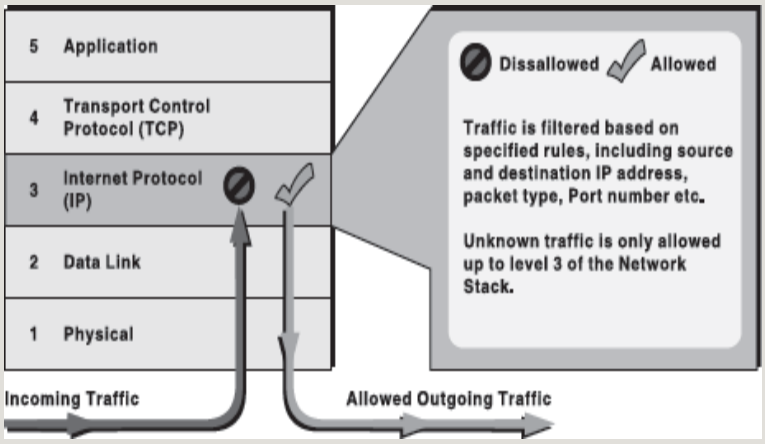


Figure 19.14 Packet Filtering Firewall

Application-Level Gateway: An application-level gateway firewall uses server programs (*called proxies*), which run on the firewall. These proxies take requests from the external network, examine them, and forward the legitimate requests to the internal host computer. This type of firewall supports functions such as user authentication and logging. As this type of firewall is considered the most secure type, it provides a number of advantages to the medium-to-high risk websites:

- The firewall can be configured as the only host address that is visible to the outside network, requiring all connections to and from the internal network to go through the firewall.

- The use of proxies for different services prevents direct access to services on the internal network, protecting the enterprise against insecure or misconfigured internal hosts.
- Strong user authentication can be enforced with application gateways.
- Proxies can provide detailed logging at the application level.

This type of firewall requires every client program to be set up as a proxy. In addition, the firewall must have a proxy in it for each type of protocol that can be used. This can cause a delay in implementing new protocols if the firewall does not support it. The penalties paid for this additional level of security are performance and flexibility. Proxy server firewalls have large processor and memory requirements in order to support many simultaneous users, and introduction of new Internet applications and protocols can often involve significant delays while new proxies are developed to support them. True proxy servers are undoubtedly the safest, but impose an overhead in heavily loaded networks.

Firewall working on application-level gateways technique requires a proxy for each service (such as FTP and HTTP), which is compatible to it. When a service is required that is not supported by a proxy, an organization has three possible choices to perform:

- Deny the service until the firewall vendor has developed a secure proxy.
- Develop a custom proxy.
- Pass the service through the firewall (Figure 19.15).

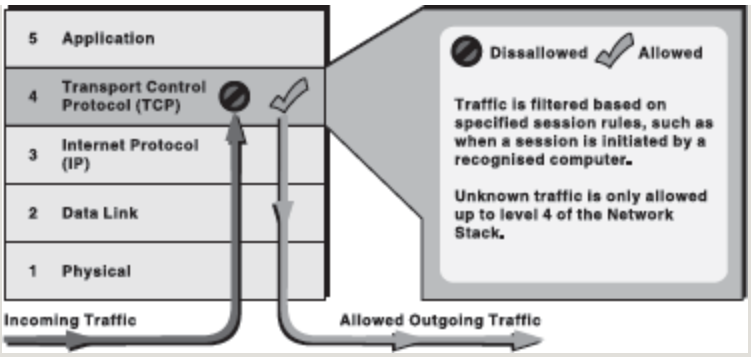


Figure 19.15 Application-level Gateway Firewall

Circuit-Level Gateway: In the circuit-level firewall, all connections are monitored, and only those connections that are found to be valid are allowed to pass through the firewall. This generally means that a client behind the firewall can initiate any type of session, but clients outside the firewall cannot see or connect to a machine protected by the firewall. Stateful inspections usually occur at the network layer, thus making it fast and preventing suspicious packets from travelling up the protocol stack. Unlike static packet filtering technique, stateful inspection makes its decisions based on all the data in the packet (corresponding to all the levels of the OSI model) (Figure 19.16).

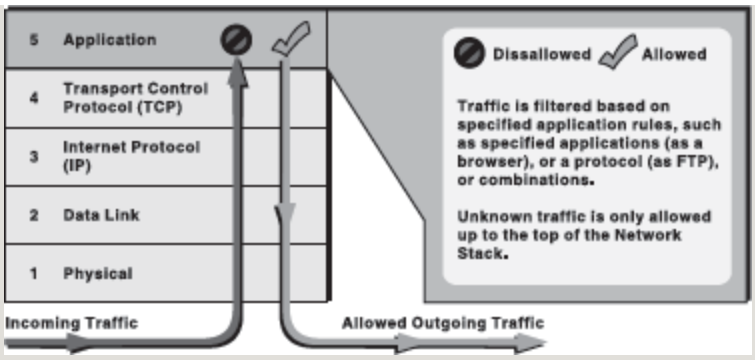


Figure 19.16 Circuit-level Gateway Firewall

Using this information, the firewall builds dynamic state tables and then uses these tables to keep track of the connections, which go through it. Rather than allowing all packets that meet the rule set requirements to pass, it allows only those packets, which are part of a valid, established connection. A typical use of circuit-level gateways is a situation in which the system administrator trusts the internal users.

19.7 USER IDENTIFICATION AND AUTHENTICATION

Identification and authentication (*I&A*) is another line of defence against the unauthorized people from entering into a computer system. I&A is a critical building block of computer security as it forms the basis for most types of access control and for establishing user's accountability. Such access control often requires a system to identify and differentiate among different users. For example, access control is often based on least privilege, which refers to granting of accesses to only those users who are required to perform their duties. User accountability requires the linking of activities on a computer system to specific individuals and, therefore, requires system to identify users.

Often people confuse *identification* from authentication as both have similar aspects. Identification is the means through which a user provides a claimed identity to the system. On the other hand, *authentication* refers to establishing validity of the claim. Computer systems make use of data authentication for recognizing people, which the systems receive. Authentication presents several challenges such as collecting authentication data, transmitting the data securely, and identifying the same person who was earlier authenticated and is still using the computer system.

There are three ways of authenticating the users, each of which can be used either individually or in combination with others:

- **Users Requirement**: (password, PIN, cryptographic key).
- **Users Possessions**: (ATM card or smart card).
- **Users Biometric**: (voice pattern, handwriting dynamics, fingerprint).

These ways provide strong authentication; however, there are certain problems associated with them. If people want to pretend to be someone else on a computer system, they can guess or learn an individual's password; they can also steal or fabricate tokens.

19.7.1 User Requirements

The most common form of authentication is the combination of a user ID and password. This technique is based solely on user requirement. In general, password systems work by requiring the users to enter a user ID and password. The system compares the password to a previously stored password for that user ID. If there is a match, the user is authenticated and granted access. This type of security access has been successfully providing security to computer systems for a long time. They are integrated into many operating systems and users and system administrators are familiar with them. When properly managed in a controlled environment, they can

provide effective security. However, this technique is dependent upon keeping passwords secret. Unfortunately, there are many ways that the secret key may be divulged.

- **Finding Passwords**: If users create their own passwords, they may tend to make it easy to remember. On the other hand, assigned passwords may be difficult to remember, so users are more likely to write them down. Many computer systems are equipped with administrative accounts that have preset passwords. Because these passwords are standard, they are easily "guessed."
- **Giving Passwords:** Users may share their passwords with a co-worker in order to share files. In addition, people can be tricked into divulging their passwords.
- **Electronic Monitoring:** When passwords are transmitted to a computer system, they can be electronically monitored. This can happen on the network used to transmit the password or on the computer system itself.
- **Accessing Password File:** If the password file is not protected by strong access controls, the file can be downloaded. Password files are often protected with one-way encryption so that plain text passwords are not available to system administrators or hackers (if they successfully bypass access controls) (Figure 19.17).

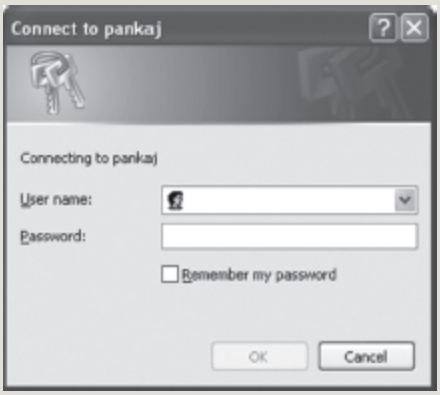


Figure 19.17 User ID and Passward

19.7.2 User Possessions

Although some techniques are based solely on users' requirements, most of the techniques are based on what the users possess. This technique uses *token* system.

Such tokens are divided into two categories: *memory tokens and smart tokens*.

Memory Tokens: Memory tokens are meant for storing information (Figure 19.18). It requires special reader/writer devices for writing and reading of data to and from the tokens. The most common type of memory token is a magnetic strip card, in which a thin stripe of magnetic material is affixed to the surface of a card (for example, as on the back of credit cards). A common application of memory tokens for authentication to computer systems is the automatic teller machine (ATM) card.

Figure 19.18 Momory Token

Memory tokens when used with PINs provide significantly more security than passwords. A hacker must have both a valid token and the corresponding PIN to pretend to be someone else. This is much more difficult than obtaining a valid password and user ID combination. Tokens can be used in support of log generation without the need for the employee to key in a user ID for each transaction or other logged event since

the token can be scanned repeatedly. If the token is required for physical entry and exit, then people will be forced to remove the token when they leave the computer. This can help maintain authentication. However, this method also has certain limitations. Although sophisticated technical attacks are possible against memory token systems, most of the problems associated with them relate to their cost, administration, token loss, user dissatisfaction, and the compromise of PINs. Most of the techniques for increasing the security of memory token systems relate to the protection of PINs.

Smart Tokens: A smart token is the functionality expansion of memory token, incorporating one or more integrated circuits into the token itself. When used for authentication, a smart token is another example of authentication based on users' possession category. A smart token requires a user to provide something the user knows (PIN or password) in order to "unlock" the smart token for use. Smart tokens offer great flexibility and are used to solve different authentication problems. The benefits of smart tokens vary, for the type it is used. In general, they provide greater security than memory cards. It can solve the problem of electronic monitoring even if the authentication is done across an open network by using one-time passwords.

However like memory tokens, most of the problems associated with smart tokens relate to their cost, the administration of the system, and user dissatisfaction. Smart tokens are generally less vulnerable to the compromise of PINs because authentication usually takes place on the card. However, smart tokens cost more than memory tokens and are more complex (Figure 19.19).

Figure 19.19 Smart Token

19.7.3 Biometric Technique

Biometric authentication technologies use the unique characteristics (or attributes) of an individual to authenticate the person's identity. These include physiological attributes (such as fingerprints, hand geometry, or retina patterns) or behavioural attributes (such as voice patterns and hand-written signatures). Biometric authentication technologies based upon these attributes have been developed for computer log in applications. Biometric authentication is technically complex and expensive, and user acceptance can be difficult.

Biometric systems provide an increased level of security for computer systems, but the technology is still new as compared to memory tokens or smart tokens. Biometric authentication devices provide imperfection, resulting from technical difficulties in measuring and profiling physical attributes as well as from the somewhat variable nature of physical attributes. These may change, depending on various conditions. For example, a person's speech pattern may change under stressful conditions or when suffering from a sore throat or cold.

Due to their relatively high cost, biometric systems are typically used with other authentication means in environments requiring high security (Figure 19.20).

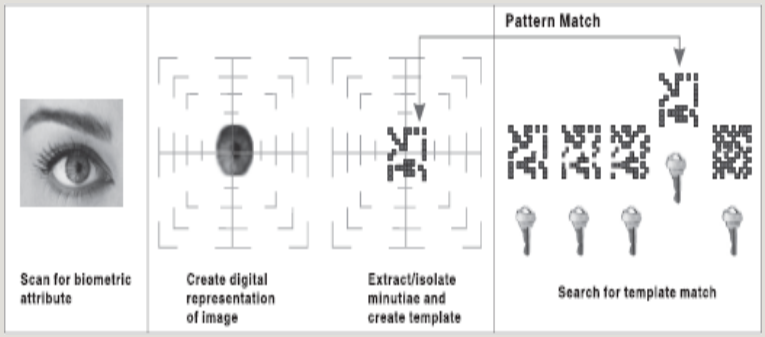


Figure 19.20 Biometric Techniques

19.8 DATA BACKUP AND RECOVERY

In spite of all the precautions you take to protect your data, some failure such as hardware failure, software failure, or natural disaster may occur resulting in data loss. To protect against such failures, you must take a backup of your data as data loss is very common to all.

Data backup (also known as *dump*) refers to making duplicate copies of your data and storing them onto some permanent storage device such as magnetic tape, magnetic disk, CD, DVD, pen drive, etc. The data backup is essential for recovering the data as quickly, and with as little damaging impact on users, as possible. *Data recovery* is a process of restoring data to the most recent consistent state that existed before the failure.

19.8.1 Backup Methods

There are four backup methods, which are *full backup, differential backup, incremental backup,* and *mirror backup*.

Full Backup Method: This is the basic backup method that backs up all the files on the computer system. Generally, full backups are performed on a weekly or monthly basis. This method is advantageous because only a single backup is enough to restore all the backed up files. However, it consumes a lot of time as well as secondary memory space to back up all the files.

Incremental Backup Method: This method backs up only those files that have been modified since the most recent backup (either full, differential, or incremental), thereby making each backup an increment to the last backup. Hence, restoring the files after a crash requires all the backups. For example, suppose a full backup has been taken on the tenth of month and incremental backup on the twelfth and the fourteenth. Then, restoring the files after a crash on the fifteenth requires the full backup as well as the two incremental backups. An important advantage of this method is that it consumes less time and memory space to take a backup. However, a single backup is not sufficient to restore all the files, which makes the restoring process slower.

Differential Backup Method: This method backs up all the files that have any difference from the last full backup. Hence, restoring the files requires the full backup and only the last differential backup. This makes restoring little faster than incremental backup, which requires all the previous backups. However, backing up is slower and storage requirement is higher than incremental backup.

Mirror Backup Method: All the methods discussed above compress the files being backed up into a single file. However, the mirror backup method copies the files or folders being backed up without any compression. Thus, it keeps each file separate in the destination,

thereby making the destination a mirror of the source. This feature makes this method fastest of all the methods. However, it requires more storage space as compared to other methods.

19.8.2 Online Backups

Online backup means taking backup of data on a remote computer over the Internet. There are several advantages of this method over other backup methods. One is that this method not only enables you to recover data after file corruption or computer failure, but you can also recover data even if a natural disaster (such as fire, flood, theft, etc.) has destroyed everything at your place. However, it is not affordable for individual users or small businesses to subscribe for online backup services.

19.8.3 Recovery Tools

Sometimes, you lose important data because of disk failure, accidental deletion of file, etc. The situation even becomes worse if you are not having backup of data. In such situations, recovery tools come to our rescue.

A recovery tool enables you to recover your lost data. A wide variety of recovery tools are available in the market, some of which are PC Inspector File Recovery, EASEUS Data Recovery Wizard, VirtualLab Data Recovery, and Stellar Phoenix Windows Data Recovery.

19.9 SECURITY AWARENESS AND POLICIES

In today's computing environment, everyone in an organization has access to system resources and, therefore, has the potential to cause harm. Each computer on networks is potentially a door in which access to other computers on the network is possible. Therefore, a need for security awareness and training is required to implement computer security in an organization. Computer security awareness and training

is an issue that affects all computer users, whether they use a personal computer or terminals connected to the mainframe computer.

19.9.1 Awareness

The main purpose behind security awareness is to enhance security by improving awareness of the need to protect system resources, developing skills and knowledge so that computer users can perform their jobs more securely and build knowledge needed to design, implement, or operate security programs for organizations and systems. Awareness is used to reinforce the fact that security supports the mission of the organization by protecting valuable resources. In addition, it is also used to remind people of basic security practices such as logging off a computer system or locking doors.

To spread awareness of computer security, various teaching methods are deployed, such as video tapes, newsletters, posters, bulletin boards, briefings, short reminder notices at logon, discussions, and lectures. Awareness is often incorporated into security training and is used to change employee attitudes. However, employees often regard computer security as an obstacle to productivity. To help motivate employees, it must be emphasized how security, from a broader perspective, contributes to productivity. The consequences of poor security must be explained, while avoiding the fear and intimidation that employees often associate with security.

Providing training is also an awareness activity, which teaches skills to the people that enable them to perform their jobs in a more secure manner. This includes teaching people what they should do and how they should (or can) do it. Training addresses many levels, from basic security practices to more advanced or

specialized skills. It can be specific to one computer system or generic enough to address all systems. Many personnel's need advanced or specialized training rather than just basic security practices. For example, managers may need to understand security consequences and costs so that they can take the security factor into their decisions. There are different ways to identify individuals or groups who need specialized or advanced training. One method is to look at job categories, such as executives, functional managers, or technology providers.

A security-training program normally includes training classes, either strictly devoted to security or as added special sections or modules within existing training classes. Training is either computer or lecture based and may include hands-on practice and case studies.

19.9.2 Security Policy

A *security policy* is a formal statement of the rules for people who are given access to an organization's technology and information assets. When developing a security policy, care must be taken to identify and understand relevant and valid issues. Mostly, resources are wasted on reacting to a high-profile hoax call while a serious issue goes unnoticed. When evaluating the effectiveness of a particular security policy, the resources being protected must be analysed, the information stored in today's computer ranges from public domain material such as telephone numbers to highly sensitive data, for example, an individual's genome. It is not practical nor is it possible to firmly secure all this information. The goal is to protect information in line with its relative value and importance to the business process. A good security policy should focus on allowing employees to access only the resources he or she needs to perform their job function.

The main purpose of security policy is to inform users, staff, and managers of their obligatory requirements for protecting technology and information assets. The policy should specify the mechanisms through which these requirements can be met.

Components of Security Policy: Once a security policy has been established, it should clearly be communicated to users, staff, and management. Having all personnel signing on the statement indicating that they have read, understood, and agreed to abide by the policy is an important part of the process. To retain the value and genuineness of the policy, it must include the following components:

THINGS TO REMEMBER

Security Policy

The important characteristics of security policy are as follows:

- It must be implementable through system administration procedure, publishing of acceptable user guidelines, or other appropriate methods.
- It must be enforceable with security tools, where appropriate, and with sanctions, where actual prevention is not technically feasible.
- It must clearly define the areas of responsibility for the users, administrators, and management.

- **Computer Technology Purchasing Guidelines:** Specify required (or preferred) security features. These should supplement existing purchasing policies and guidelines.
- **Privacy Policy:** Defines reasonable expectations of privacy regarding such issues as monitoring of electronic mail, logging of keystrokes, and access to users' files.
- **Access Policy** Defines access rights and privileges to protect assets from loss or disclosure by specifying acceptable guidelines for users, operations staff, and management.

- **Accountability Policy:** Defines the responsibilities of users, operations staff, and management.
- **Authentication Policy:** Establishes trust through an effective password policy and by setting guidelines for remote location authentication and the use of authentication devices.
- **Information Technology System and Network Maintenance Policy:** Describes how both internal and external maintenance people are allowed to handle and access technology. One important topic to be addressed here is whether remote maintenance is allowed and how such access is controlled.
- **Violations Reporting Policies:** Indicate types of violations to be reported and to whom the reports are to made. A non-threatening atmosphere and the possibility of anonymous reporting results in a greater probability that a violation will be reported if it is detected.

LET US SUMMARIZE

1. *Computer security* refers to the protection given to computers and the information contained in them from unauthorized access. It involves the measures and controls that ensure *confidentiality, integrity, and availability* of the information, processed, and stored by a computer.
2. Computer systems are vulnerable to many kinds of threats that can cause various types of damages, which may result in significant data loss. These damages can range from errors that can cause harm to database integrity to natural calamity destroying entire computer centres.
3. *Malicious codes* are the programs that generate threats to the computer system and precious data. This code can be in the form of viruses, worms, Trojan horses, logic bombs, and other "uninvited" software.
4. *Viruses* are programs, which are designed to replicate, attach to other programs, and perform unsolicited and malicious actions. It executes when an infected program is executed. On MS-DOS systems, these files usually have the extensions .exe, .com, or .bat. The different types of virus include *boot sector virus, file-infecting virus, polymorphic virus, stealth virus, and multipartite virus*.
5. *Worms* are the programs that are constructed to infiltrate on the legitimate data processing programs and alter or destroy the data. They often use network connections to spread from one computer system to another. Once active within a system, worms behave like a virus and perform a number of disruptive actions.
6. A *Trojan horse* is a destructive program, which is concealed in a piece of software. It enters into a computer through an e-mail or free programs that have been downloaded from the Internet. Once safely get into the computer, it usually opens the way for other malicious software (like viruses) to enter into the computer system. In addition, it may also allow unauthorized users to access the information stored in the computer.
7. A *logic bomb* is a program or portion of a program, which lies dormant until a specific part of program logic is activated. The most common activator for a logic bomb is a date. The logic bomb checks the computer system date and does nothing until a preprogrammed date and time is reached.

8. *Antivirus* is a software utility, which (upon installing on a computer) scans the hard disk for viruses and tries to remove them if found. Most Antivirus programs include an auto-update feature that enables the program to download profiles of new viruses so that it can check for the new viruses as soon as they are discovered.
9. *Cryptography* is the process of altering messages to hide their meaning from adversaries who might intercept them. It provides an important tool for protecting information and is used in many aspects of computer security. Cryptography techniques are broadly classified into three types: *SKC, PKC*, and *hash functions*.
10. *Digital signature* is used to verify the authenticity of electronic document. In other words, digital signatures play the role of physical signatures in verifying electronic documents. It uses PKC technique, which employs an *algorithm* using two different but mathematically related keys; one for creating a digital signature or transforming data into a seemingly unintelligible form and another key for verifying a digital signature or returning the message to its original form.
11. *Firewall* is an application, which prevents certain outside connections from entering into the network. It traps inbound or outbound packets, analyses them, and then permits access or discards them. It is typically implemented using one of the three primary architectures: packet filtering, application-level gateway, and circuit-level gateway.
12. *Identification and authentication (I&A)* is another line of defence against the unauthorized people from entering into a computer system. I&A is a critical building block of computer security as it forms the basis for most types of access control and for establishing user's accountability.
13. *Data backup* (also known as *dump*) refers to making duplicate copies of your data and storing them onto some permanent storage device such as magnetic tape, magnetic disk, CD, DVD, pen drive, etc. There are four backup methods, which are full backup, differential backup, incremental backup, and mirror backup.
14. *Data recovery* is a process of restoring data to the most recent consistent state that existed before the failure. A recovery tool enables you to recover your lost data. A wide variety of recovery tools are available in the market, some of which are PC Inspector File Recovery, EASEUS Data Recovery Wizard, VirtualLab Data Recovery, and so on.
15. *Computer security awareness and training* is an issue that affects all computer users, whether they use a personal computer or terminals connected to the mainframe computer. Awareness is used to reinforce the fact that security supports the mission of the organization by protecting valuable resources. Training is an activity, which teaches skills to the people that enable them to perform their jobs in a more secure manner.
16. A *security policy* is a formal statement of the rules for people who are given access to an organization's technology. A good security policy should focus on allowing employees to access only the resources he or she needs to perform their job function.

EXERCISES

Fill in the Blanks

1. Computer security involves the measures and controls that ensure ....., ..., and ..... of the information processed and stored by a computer.
2. Computer systems are vulnerable to many kinds of ..... that can cause various types of damages, which may result in significant data loss.
3. A person with the intention of finding some weak points in the security of websites and other computer systems in order to gain unauthorized access is termed as .....
4. ..... is a small segment of code, which replicates by attaching copies of itself to existing executables files.
5. ..... are the programs constructed to infiltrate on the legitimate data processing programs and alter or destroy the data.
6. A destructive program, which enters into a computer through an e-mail or free programs that have been downloaded from the Internet, is called .....
7. .....resides inside the memory and infects most of the executable files on a system.
8. A ..... is a program or portion of a program, which lies dormant until a specific piece of program logic is activated.
9. ..... is a software utility, which scans the hard disk for viruses and tries to remove them if found.
10. The process of altering messages to hide their meaning from adversaries who might intercept them is called .....

Multiple-choice Questions

1. Some of the commonly occurring threats to a network are .....
    1. Errors and omissions
    2. Fraud and theft
    3. Malicious code
    4. All of these
2. .....is a program, named after the Greek mythology about the Trojan war.
    1. Worm
    2. Logic bomb
    3. Trapdoor
    4. Trojan horse
3. Virus, which infects the master boot record of a computer system, is called
    1. File-infecting virus
    2. Stealth virus
    3. Boot sector virus
    4. Multipartite virus
4. Cryptography technic relies on .....
    1. Algorithm
    2. Key
    3. Both (a) and (b)
    4. None of these
5. The shared key concept of cryptography is used in the .....

1. Public-key cryptography
2. Secret-key cryptography
3. Hash function
4. None of these

6. Firewall application is used for .....

    1. Trapping inbound or outbound packets
    2. Scanning viruses
    3. Encrypting messages
    4. None of these

7. ..... is a 32-bit number, normally expressed as four “octets” between the periods.

    1. Port
    2. IP address
    3. Domain name
    4. None of these

8. Firewall technique, which requires a proxy for each service to be supported through it, is called .....

    1. Packet filtering
    2. Circuit-level gateway
    3. Application-level gateway
    4. None of these

9. ..... comes into the category of Identification and authentication, based on the things a user possesses.

    1. Smart token
    2. Biometric
    3. Username and password
    4. Both (a) and (c)

10. Biometric technique makes use of ..... for authentication.

    1. Fingerprint
    2. Eye's retina
    3. Credit card
    4. Both (a) and (b)

State True or False

1. A security policy is a formal statement of the rules for people who are given access to an organization's technology and information assets.
2. Errors and omissions are not an important threat to data and system integrity.
3. Fraud can be committed by both insiders and outsiders of an organization.
4. A person who breaks into the computer security without authorization for ulterior purposes is called white-hat hacker.
5. Worm is a self-replicating program, which is self-contained and does not require a host program.
6. A Trojan horse is a useful program, which is concealed in a piece of software.
7. Stealth virus attempts to conceal its presence from the user.
8. Cryptography comprises conversion of plain text into cipher text and vice versa.
9. Hash function is a two-way encryption algorithm.
10. Digital signature is created and verified by means of cryptography.

Descriptive Questions

1. Define computer security. Why is the need for security policy felt in an organization?
2. What are the various components of security policy?
3. Discuss various types of security threats arising from computer network.
4. Write short notes on:
    1. Worm
    2. Trojan horse
    3. Logic bomb
    4. Recovery tools
5. What is a computer virus? Discuss its types.
6. Explain how virus spreads in the computer system and the preventive measures to counter it.
7. What is cryptography? Discuss its types.
8. What do you mean by data backup? Explain different backup methods along with their advantages and disadvantages.
9. Explain how digital signature is used while sending a document across a network.
10. Discuss how firewall prevents unauthorized access to the computer system.
11. How Identification and authentication help in securing the user's security?
12. What is the advantage of online backup over other backup methods?

ANSWERS

Fill in the Blanks

1. Confidentiality, integrity, availability
2. Threats
3. Hacker
4. Virus
5. Worms
6. Trojan horse
7. File-infecting virus
8. Logic bomb
9. Antivirus
10. Cryptography

Multiple-choice Questions

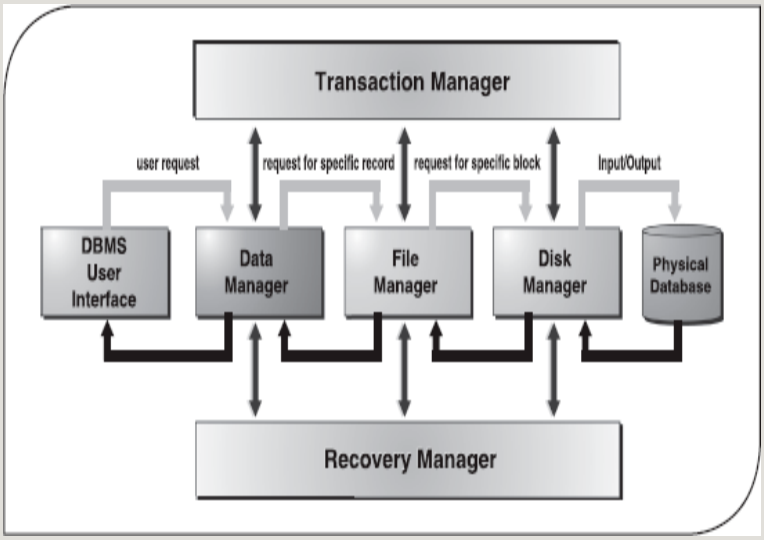
1. (d)
2. (d)
3. (c)
4. (c)
5. (b)
6. (a)
7. (b)
8. (c)
9. (a)
10. (d)

State True or False

1. True
2. False
3. True
4. False
5. True
6. False
7. True
8. True
9. False
10. True

Chapter 20

Database Fundamentals



This chapter provides a basic understanding of Database Management System (DBMS). DBMS is a software, designed to control the information in a database (the warehouse of data and information). The chapter deals with the basic concepts of data and its organization, database fundamentals, database management system along with its components and architecture. It provides a clear understanding of various database models and how databases are normalized. The chapter concludes with an overview of types of databases, data warehousing and data mining.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The basic concepts of database
- The definition of database
- Logical data concepts, which reflect the way data are viewed by the programmer or the end-user
- The manner in which data are physically stored in the hardware
- The database management system, which is a collection of interrelated data and a set of programs to access that data
- DBMS architecture, which provides an abstract view of the database
- Database models, which represent how different elements of database are associated
- Normalization, which is the process of reducing data redundancy in a relational database
- The various types of database
- Database warehousing and data mining, which are used for large-scale processing and analysis of data

20.1 INTRODUCTION

In today's changing technological environment, information is power. Owing to the rapid growth of IT in the last few decades, organizations and institutions have realized the value of information as a resource and the importance of the speed and ease with which this resource can be managed. The eternal quest for data management has led to the development of database technology. A *database* can be defined as a collection of related data from which users can efficiently retrieve the desired information. In addition to the storage and retrieval of data, certain other operations can also be performed on a database. These operations include

adding, updating, and deleting data. All these operations on a database are performed using a database management system (DBMS). Essentially, a DBMS is a computerized record-keeping system.

About half a century ago, computer was developed. It was a revolutionary technology, and almost instantly, computers were employed to resolve the problem of information storage and retrieval. Computers replaced tons of paper, file folders, and file cabinets as the principal media for storing important information. With the advent of small, powerful personal computers, databases have become more readily usable by the average computer user.

20.1.1 Data versus Information

Before defining database, we should know about the two terms, *data* and *information*, which are used frequently with databases. Data can be anything such as a number, a person's name, images, sounds, and so on. Hence, *data* can be defined as a set of isolated and unrelated raw facts, represented by values, which have little or no meaning, simply because they lack a context for evaluation. Usually, the values are represented in the forms of characters, numbers, or any symbol such as “Monica,” “35,” and “chef.” Note that although these words and numbers have certain meaning, it is difficult to figure out exactly what these values signify. However, when the data are processed and converted into a meaningful and useful form, it is known as *information*. Hence, information can be defined as a set of organized and validated collection of data. For example, “Monica is 35 years old and she is a chef.”

Strictly speaking, data refer to the values physically recorded in the database whereas information refers to the conclusion or meaning drawn out from it. With respect to database, these terms are synonymous. In

reality, information in one sense may be data in another. Sometimes, it is hard to tell the difference between the two. For example, the bill at the grocery store can be used as data as well as information. As information, the bill represents the amount you owe to the store. However, when the store manager uses his computer to calculate total sales for the day based on all purchases made, it is treated as data. In summary, whether something is data or information depends on how it is used in relation to the given context.

Other than data and information, one more term, knowledge, is frequently used with database technology.*Knowledge* is the act of understanding the context in which the information is used. It can be based on learning through information, experience, and/or intuition. Based on the knowledge, the information can be used in a particular context, for example, if an hotelier uses the information about Monica (*she is a chef*) to hire her, he is using his knowledge. Hence, knowledge can also be referred to as a person's capability and wisdom and how much that person knows about a particular subject. Consequently, it can be said that data constitute information, and information constitutes knowledge (Figure 20.1).

Figure 20.1 Data, Information, and Knowledge

20.2 DATABASE: DEFINITION

A *database* can be defined as a collection of related data in an organized way. The organized information (that is, database) serves as a base from which the desired

information can be retrieved, conclusions can be drawn, and decisions can be made, by further reorganizing or processing this data. In essence, a database not only stores information, but the information is integrated so that when the database is interrogated, users can get something useful from it. A database can be anything from a simple collection of roll numbers, names, addresses, and telephone numbers of students to a complex collection of sound, images, and even video or film clippings. Though databases are generally computerized, instances of non-computerized databases from everyday life can be cited in abundance. A dictionary, a telephone book, a collection of recipes, and a TV guide are all common examples of non-computerized databases. The examples of computerized databases include customer files, employee rosters, books catalogue, equipment inventories, and sales transactions.

20.2.1 Basic Database Terms

Before discussing the concepts related to the database, it is required to be familiar with the basic database terms which are as follows.

Field/Column: A *field* represents one related part of a table and is the smallest logical structure of storage in a database. It holds one piece of information about an item or subject. For example, in a database maintaining information about employee, the fields can be *Code, Dept, Name, Address, City*, and *Phone* (see Figure 20.2).

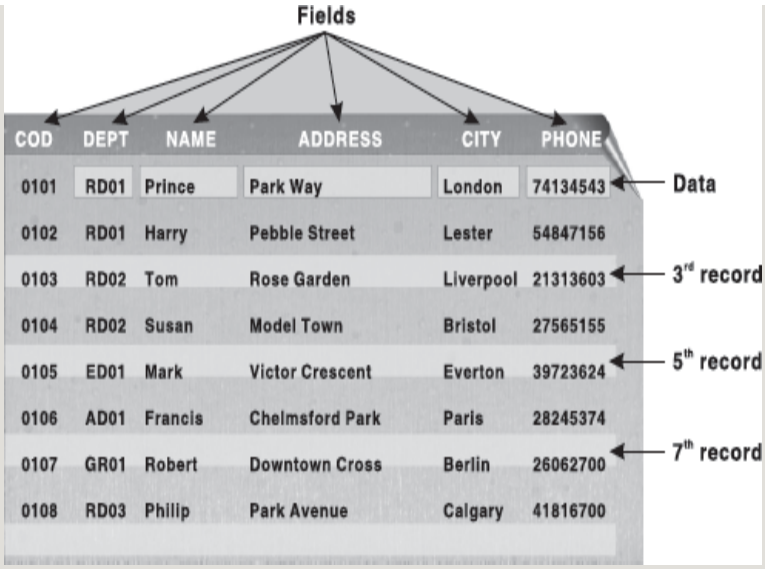


Figure 20.2 Table

Record/Row: A *record* is a collection of multiple related fields that can be treated as a unit. For example, fields *Code, Dept, Name, Address, City*, and *Phone* for a particular employee form a record. Figure 20.2 contains eight records (0101 to 0108), and each record has six fields. In database terminology, each row is a record. However, database experts prefer to use the word row, since record sometimes has other connotations outside the databases.

Table: A *table* is a named collection of logically related multiple records. For example, a collection of all the employee records of a company form employee table. Note that every record in a table has the same set of fields. Depending on the database software, a table can also be referred to as a *file*. The collection of multiple related files (tables) forms the database (Figure 20.2).

Note: *The database software can save each collection of records (table) in a separate file or saved in a single database, which is logically separated*

in tables.

Data Type: A data type determines the type of data that can be stored in a column. Although many data types are available, the most commonly used data types are Character, Numeric, and Date and Time.

- **Character Data Type:** It is used to store characters, numbers, special characters, or combinations of any of these. Note that if a numeric value is stored in an alphanumeric field, the value is treated as a character, not a number.
- **Numeric Data Type:** It is used to store only numeric values.
- **Date and Time Data Type:** It is used to store date and time values. The values for this data type vary widely depending on the database software being used.

20.3 LOGICAL DATA CONCEPTS

The overall database design and implementation process starts with requirement gathering and specifications. Once the requirements of the user have been specified, the next step is to construct an abstract or conceptual model of a database based on the requirements of the user. The conceptual model represents various pieces of data and their relationships at a very high level of abstraction. It mainly focuses on what data are required and how it should be organized rather than what operations are to be performed on the data. The conceptual model can be represented using *Entity–Relationship model* (*E–R model*). The E–R model views the real world as a set of basic objects (known as *entities*), their characteristics (known as *attributes*), and associations among these objects (known as *relationships*). The *entities, attributes*, and *relationships* are the basic constructs of an E–R model.

20.3.1 Entity

An *entity* is an object that has an independent existence in the real world. It includes all those “things” about which data are collected. An entity may be a tangible object such as a student, a place, or a part. It may also be non-tangible such as an event, a job title, or a customer

account. Diagrammatically, entities are represented in rectangles. For example, the entities, *CUSTOMER* and *ITEMS* can be represented as shown in Figure 20.3.

Figure 20.3 Entities

20.3.2 Attribute

Attributes are the properties of an entity that characterize and describe it. In a database, entities are represented by tables and attributes by columns. For example, a *CUSTOMER* entity might have numerous attributes such as *Code, Name*, and *Address*. Similarly, the *ITEMS* entity may have attributes like *Item_Id* and *Price*. They are drawn in elliptical shapes along with the entity rectangles (see Figure 20.4).

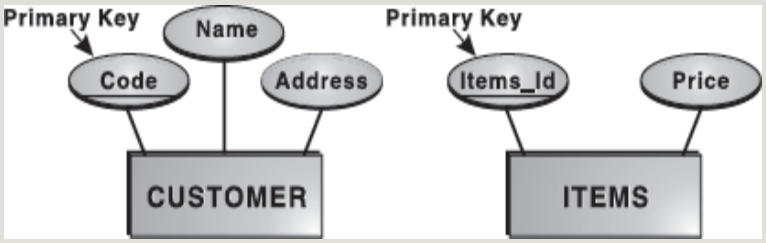


Figure 20.4 Attributes

20.3.3 Relationship

Relationship is an association, dependency, or link between two or more entities and is represented by a

diamond symbol. A relationship describes how two or more entities are related to each other. For example, the relationship *Buys* (shown in Figure 20.5) associates a *CUSTOMER* with *ITEMS*.

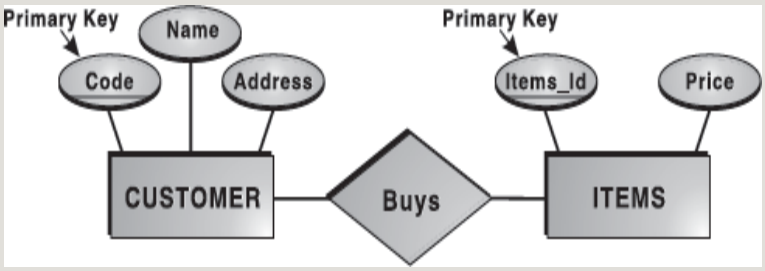


Figure 20.5 Entities, Attributes, and Relationship

Even though a relationship may involve more than two entities, the most commonly encountered relationships are binary, involving exactly two entities. Generally, such binary relationships are of three types: *one-to-one, one-to-many*, and *many-to-many*.

One-to-One Relationship (1:1): In one-to-one relationship, one record in a table is related to only one record in another table. For example, a department cannot be headed by more than one department head and a department head can be the head of only one department (Figure 20.6).

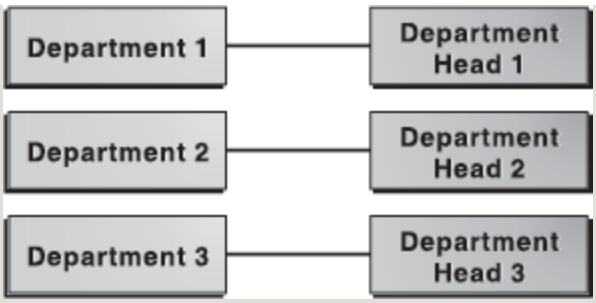


Figure 20.6 One-to-One Relationship

One-to-Many Relationship (1:M): In one-to-many relationship, one record in a table (parent table) can be related to many records in another table (child table). For example, a father may have more than one child but the child has only one father (Figure 20.7).

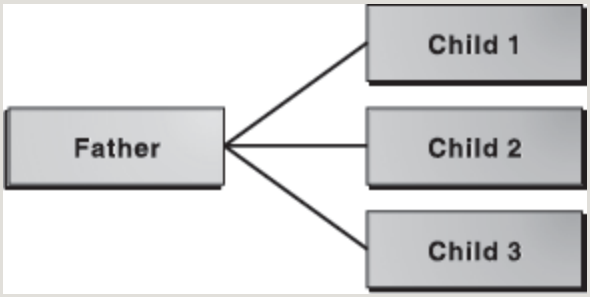


Figure 20.7 One-to-Many Relationship

Many-to-Many Relationship (M:M): In many-to-many relationship, one record in a table can be related to one or more records in second table, and one or more records in the second table can be related to one or more records in the first table. For example, a customer can

buy many items and the same item can be bought by many customers (Figure 20.8).

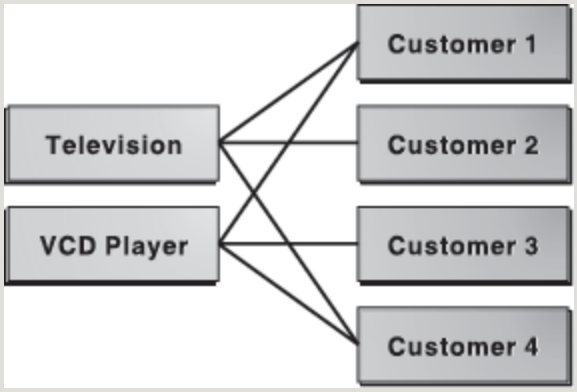


Figure 20.8 Many-to-Many Relationship

20.4 PHYSICAL DATA CONCEPTS

Physical concepts of data refer to the manner in which data are physically stored on the hardware (like hard disk). Fundamentally, it involves the physical organization of the records of a file for the convenience of storage and retrieval of data. The physical organization is much more related with optimizing the use of the storage medium when a particular logical structure is stored on, or in it. Usually, the files are organized in three fashions—*sequential, direct*, and *indexed sequential*.

20.4.1 Sequential Files

In sequential files, data are stored and/or retrieved in a logical order, that is, in a sequence. The records are stored one after another in an ascending or descending order, based on the key field (which is unique for each record) of the records. Generally, these files are stored on sequential storage devices like magnetic tapes. In such files, to retrieve a record, all the records must be

traversed sequentially before reaching to the desired record. An analogy to sequential files may be taken as an audiocassette. If the listener is at the first song and wants to listen to the fourth song, the second and third songs need to be traversed through, respectively.

Sequential files are easy to organize and maintain. The hardware, associated with sequential organization, is also relatively cheaper, compared to the other file organizations. However, sequential search takes a very long time. If there are *N* records in the file, on an average, the user will have to go through *N/2* records before finding the right one. Similarly, making changes (such as adding, deleting, and updating records) to a sequential file is also a difficult operation. For example, to add a record in the middle of a sequential file, the entire file has to be sorted, rewritten, and stored on the storage device again (Figure 20.9).

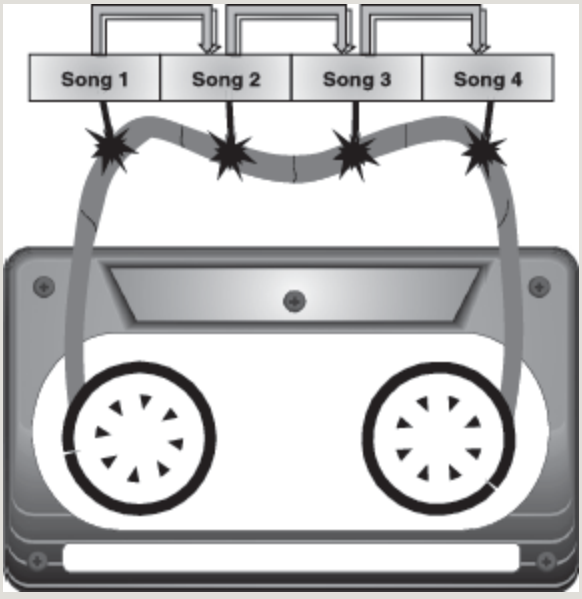


Figure 20.9 Sequential Access

20.4.2 Direct Files

Direct files facilitate accessing any record directly or randomly without having to traverse the sequence of the records. These files are also known as *random* or *relative files*. Even though only one item can be accessed at a time, that item may be stored anywhere in the file. For example, in case of CDs (compact discs), any song can be played randomly. Generally, these files are stored on direct access storage devices (DASDs) such as hard disks and CDs (Figure 20.10).

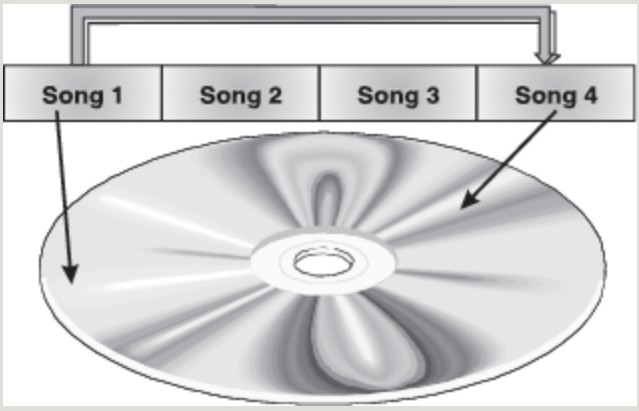


Figure 20.10 Direct Access

When the records are stored, the program itself generates the address for the record by applying certain techniques to its key field, which is unique for every record. These addresses are stored in a *master index*, which is loaded into main memory before any database manipulation takes place. It defines the relationship between the primary key and the corresponding disk addresses (a piece of information that tells the disk R/W head where to look for data). Since direct search is done in the computer's main memory, it takes less amount of

time. Based on the address found in master index, R/W head immediately goes to the correct disk position and retrieves the desired record. As a result,updation, addition, and deletion of record becomes much faster. However, this method is expensive than sequential, because usually the direct files are stored on direct access devices, which are expensive as compared to sequential storage devices. Moreover, direct access is less storage efficient because they tend to generate wide gaps in between any two records. In addition, address generation overhead is also involved for accessing records due to addressing functions.

20.4.3 Indexed Sequential Files

Before discussing the indexed sequential file organization, let us first know about indexes. You might have used indexing many times while going through a book. At the back pages of most of the books, certain key words (topics), along with the page numbers on which these topics are discussed, are given. In similar fashion, an indexed file includes an index table (also known as *reference table*) that relates key field values to storage locations of the corresponding records. Hence, indexing can be defined as a technique of ordering the records in a table without making any modification in the table (Figure 20.11).

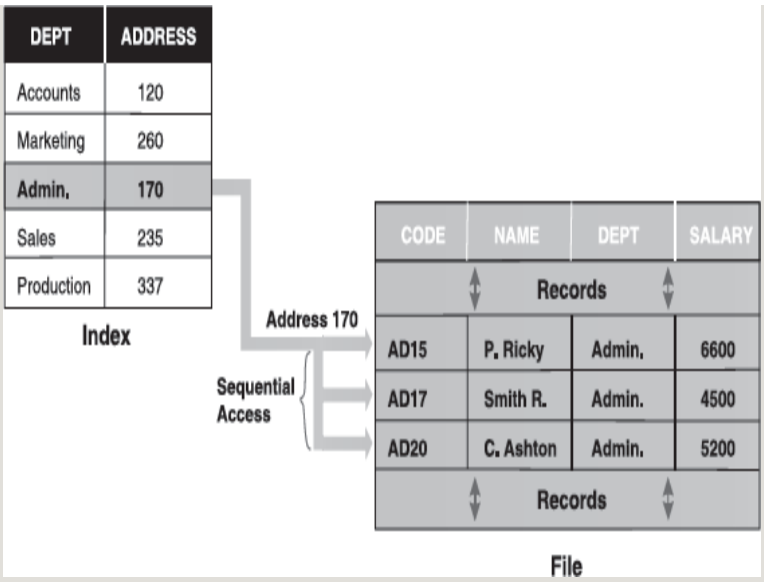


Figure 20.11 Indexed Sequential Access

Essentially, indexed sequential technique is a hybrid of sequential and direct file organization. It provides a combination of access types that are supported by a sequential and a direct file. The indexed file organization uses a separate index file, which contains the key values and the location of the corresponding record. The records are organized in an orderly sequence and the index table is used to access the records without searching through the entire file. The records may be in random sequence, but the index table is stored in sorted sequence on the key values. Since the index table is in sorted sequence, the file management system simply accesses the records in the order of the index values. To access a particular record, a search is made in the index table to determine the address of the first record of the segment in which the data are placed. Then a direct access is made to that address. After that, a sequential search is invoked until the desired record is located.

Index file requires considerably less space than the data file as it requires only two fields—the key field and the address of record's location, even if the record may have numerous fields. However, it is still an overhead, as other file organizations do not use any extra file. Indexes do not affect the file organization of records physically, that is, various indexes can use the same file without rearranging its records. Indexing speeds up the data retrieval, but it may slow down the update process because indexes also need to be maintained along with the data file while adding and deleting. Since the index sequential files require DASDs, they are an expensive means of file organization.

20.5 DATABASE MANAGEMENT SYSTEM

As discussed earlier, a database is a collection of well-organized data. In order to carry out operations like insertion, deletion, and retrieval, the database needs to be managed by a software package. This software is called a database management system (*DBMS*). Hence, DBMS can be defined as a collection of interrelated data and a set of programs to access that data. The primary goal of a DBMS is to provide an environment that is congenial and efficient to retrieve and store information. It allows a user to store, update, and retrieve data in abstract terms. DBMS eases the maintenance and retrieval of information from a database. A DBMS also relieves the user from having to know about the exact physical representations of data and from specifying detailed algorithms for storing, updating, and retrieving the data.

Principally, a database system is a computerized record-keeping system whose overall purpose is to maintain information and to make it available on demand. The management of data involves defining the structures for information storage and provision of mechanisms for its manipulation. DBMS also provides mechanisms for the

security of the stored information from unauthorized access and recovery of the data during system failures. In addition, the database system avoids possible anomalous results in case the data are to be shared among several users.

20.5.1 Need for DBMS

Historically, information systems employed stand-alone systems for separate applications each with their own set of files (file-based systems). This implied that data, such as addresses, may be duplicated in many separate systems, such as an employee's address may be held once in the payroll system and again in the personnel system. This causes waste of space and inconsistency. There may be a possibility for the address to be updated in one system but not in the other, hence, making the data inconsistent.

The problems described above led to the development of the database approach. In the database approach,an organization attempts to build an integrated corporate database to support all the data requirements of the organization (in this context database means all the data held by the organization, whether held in computerized or non-computerized database applications). DBMSs are very good at organizing and managing large collections of persistent data. A DBMS is like a suitcase where all the stuff (data) are put so that it is in one place and easy to reach. This protects data from unauthorized access and accidental corruption or loss due to hardware (like power outages) and software failures (like operating system crashes). DBMS allows concurrent access, meaning that a single data set can be accessed by more than one user at a time. Although concurrent access may introduce unwanted problems, DBMSs provide mechanisms to prevent concurrent access problems.

File-based Approach vs. DBMS Approach:
Traditionally, data were stored and processed on multiple files with the help of a program (or a group of programs) for each application. This is known as *file-based approach* of database management. A file may be defined as systematized self-containing collection of records. It may consist of data (data file) or it may contain a sequence of basic statements (program). Each user works with a different program that handles its own independent data.

THINGS TO REMEMBER

File-oriented Approach

The file structure is defined in the program code and the data used for file are dependent on it. Usually programs are written in different languages that lead to incompatible file formats and one program file may not easilyaccess other files. Since programs are written to comply with particular functions, only a fixed number of queries can beperformed and any new requirement will need a new program.

Each program maintains its own set of data. Users of one program may not be aware of potentially useful data held by other programs. In Figure 20.12, assume that the *EMPLOYEE* file has fields like *Code, Name, Dept*, and *Salary*, and the *ALLOWANCE* file contain fields like *Code, HRA, DA*, and *CCA*. Note that the same data for *Code* field is held by different programs. As a result, space is wasted and potentially different values and/or different formats for the same item may be stored in different files.

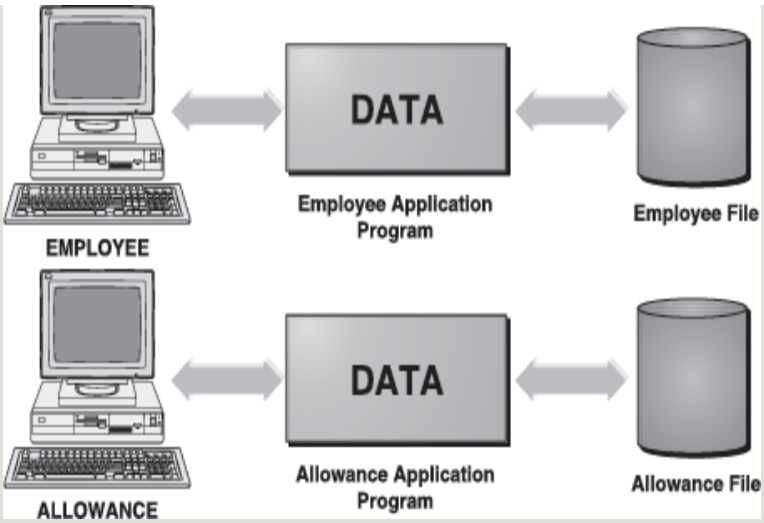


Figure 20.12 File-Oriented Approach

DBMS is a software system that enables users to define, create, and maintain the database and provides controlled access to this database. This approach has a *database engine* placed between the applications and the data. The engine is the central component of a DBMS and it provides access to the database. It also coordinates all the other functional elements (manipulate, add, delete, search, select, and store data) of the DBMS. Only the engine knows how the data are physically stored and applications pass requests to the engine to read and write data. The definition of data is embedded in application programs, rather than on separate and independent files. The data are logically related which comprise entities, attributes, and relationships of the information (Figure 20.13).

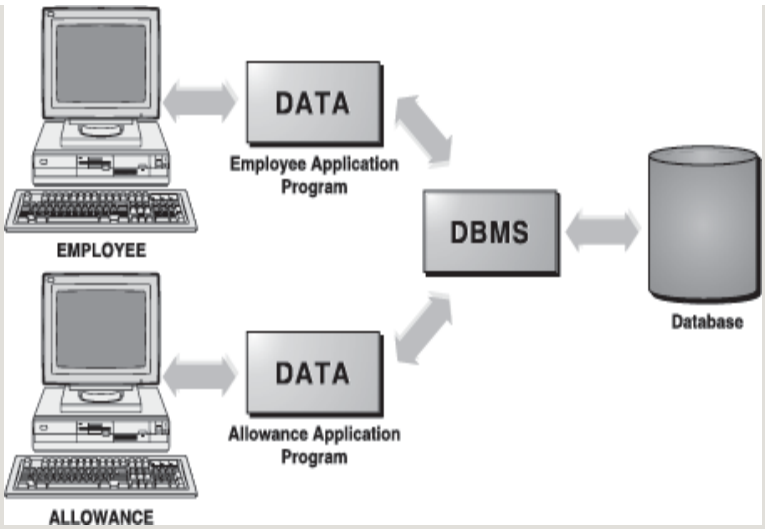


Figure 20.13 DBMS Approach

In general, the data in a DBMS are integrated as well as shared.

- **Integrated:** A database can be considered as a unification of several data files (tables), with any redundancy among those files eliminated (either wholly or partly). For example, a database might contain both the *EMPLOYEE* and the *ALLOWANCE* files. To process the *ALLOWANCE* file, one need not include the redundant field in this database as it can easily be extracted by referring to the *EMPLOYEE* table.
- **Shared:** Shared data means that individual pieces of data in the database can be shared among numerous users. Each user can have access to the same piece of information, and every user can use it for different purpose. In the above example, department information in the *EMPLOYEE* file can be shared by users in the personnel department and the accounts department.

20.5.2 Benefits of DBMS

Let us discuss some of the major advantages provided by the DBMS.

- **Reduction in Data Redundancy:** Data redundancy refers to duplication of data. In non-database systems, each application has its own separate files. This can often lead to redundancy in stored data, which results in wastage of space. For example, in case of file-based system of a magazine publishing house, the customer department keeps a separate customer details file. A similar file is also kept with

the postage-handling department. Hence, the same data are duplicated in two separate places. The other problem with this system is that the data may not be updated consistently. Suppose a customer requests for change in his/her posting address to the customer department and the changed address is reflected in his/her record. However, it is not guaranteed that the same change is also done in the postage-handling department's file. If the database system forgets or delays to update this file, this may lead to postage of the subscribed magazine to the old address.

A DBMS does not maintain separate copies of the same data. All the data are kept at one place and various applications refer to the data from this centrally controlled system. As a result, when the customer demands the change in address, it is done at one place from where the concerned department's application uses the data to process.

- **Reduction in Inconsistency:** This is a corollary of the previous point. Suppose that the customer's xyz address is represented by two distinct entries in the stored database and the database system is not aware of this duplication (that is, redundancy is not controlled), then there will be some occasions on which the two entries will not agree, that is, when only one of the two entries has been updated. At such times, the database is said to be inconsistent. As a result, inconsistent or contradictory information will be supplied to the users.

  It is obvious that if the given fact is represented by a single entry, then such inconsistencies may not occur. Alternatively, even if the redundancy is not removed but is controlled, then the DBMS ensures that the database is always consistent by ensuring that any change made to either of the two entries is automatically applied to the other one also. This process is known as *propagating update*. Note that recent DBMS packages like Oracle and Sybase are built with capabilities of propagating updates automatically.

- **Sharing of Data:** Sharing of data allows the existing applications to use the data in the database. Simultaneously, it also helps in developing new applications, which will use the same stored data. Due to shared data, it is possible to satisfy the data requirement of new applications without having to create any additional stored data or with marginal modification. For example, in the banking corporation, the account holder's name and address can be shared by all the applications. As a result, any changes in the current account policies can be implemented with virtually no changes in the table's structure.
- **Enforcement of Standards:** Since access to the database must be through the DBMS, standards are easier to enforce. Standards may relate to the naming of the data, the format of the data, the structure of the data, and so on. Standards may include the yardsticks laid by the company or organization using the database. There may be certain industry, national, or international standards that must be satisfied by the data. Standardization of data particularly helps in migrating or interchanging data between different systems.
- **Improvement in Data Security:** In conventional systems, applications were developed in an impromptu manner. Usually, different systems of an organization would access different components of the operational data. In such an environment, enforcing security can be quite difficult. Setting up of a DBMS makes it easier to enforce security restrictions since the data are stored centrally. DBMS

can ensure that the only means of access to the database is through the authorized channel. Hence, data security checks can be carried out whenever access is attempted to sensitive data. To ensure security, DBMS provides security tools such as user codes and passwords. Different checks can be established for each type of access (addition, modification, deletion, etc.) to each piece of information in the database.

- **Maintenance of Data Integrity:** Data integrity refers to ensuring that the data in the database are accurate. Since, in DBMS, the data are centralized and is used by a number of users at a time, it is essential to enforce integrity controls. Integrity may be compromised in many ways. For example, someone might input the salary data of a full-time employee as 50000 rather than 5000.

  If many users are allowed to update the same data item at the same time, there is a possibility of incorrect and inconsistent updates. For example, in an airline reservation DBMS, there could be a situation where the number of bookings made is larger than the capacity of the aircraft. Therefore, appropriate controls must be introduced to prevent occurrence of such errors because of concurrent updating activities. However, since all data are stored only once, it is often easier to maintain integrity than in conventional systems.

- **Better Interaction with Users:** As compared to traditional database systems, a DBMS often provides better service to the users. In conventional systems, usually the information is poorly arranged and, as a result, the availability of up-to-date information becomes poor. However, in the case of DBMS, the availability of up-to-date information improves since the data can now be shared and the DBMS makes it easy to respond to unforeseen information requests. Centralizing the data in a database also means that users can obtain new and combined information that would have been impossible to obtain otherwise. In addition, use of a DBMS allows the novice users (who do not know programming) to interact with the data more easily.
- **Efficient System:** It is very common to change the contents of the stored data. These changes can easily be made in a DBMS than in a conventional system as these changes do not need to have any impact on application programs. The cost of developing and maintaining systems is also lower. It is much easier to respond to unforeseen requests when the data are centralized in a database than when it is stored in file-based systems. Although the initial cost of setting up of a DBMS can be high, its overall cost is less than that for a similar service using conventional systems.

20.5.3 Components of a DBMS

Usually a DBMS is a large software package that carries out diverse tasks including the provision of facilities to enable the user to access and modify information in the database. On one hand, it is an intermediate link between the physical database, the computer, and the operating system, and, on the other hand, the users.

Essentially, this system involves data, the hardware that physically stores that data, the software that utilizes the hardware's file system in order to store the data and provide a standardized method for retrieving or changing the data, and finally, the users who turn the data into information.

- **Data:** Data stored in a database include numerical data including whole numbers and floating-point numbers, and non-numerical data such as characters, date, or logical values (true or false). More advanced systems may include more complicated data entities such as pictures and images as data types.
- **Hardware:** Hardware of the system can range from a PC to a network of computers. It also includes various storage devices (like hard disks) and input and output devices (like monitor, printer, etc.).
- **Software:** Software of a DBMS includes the DBMS, operating system, network software (if necessary), and the application programs.
- **Users:** In DBMS, generally three broad classes of users are considered. These are application programmers, end-users, and DBA (database administrator). The application programmers develop the application programs. These programs can manipulate the database in all the possible ways. The end-users access the database from a terminal using a query language provided by the database system or through application programs developed by application programmers. The database administrator (DBA) is the person who is responsible for the design, construction, and maintenance of a database.

Figure 20.14 illustrates a typical interaction among various DBMS components. The user requests for specific information with the help of user interface, that is, application programs. This request is processed by DBMS's data manager and after processing, the data manager requests for the specific records to the file manager. The file manager then requests for the specific block (where the requested records are physically stored on the hard disk) to the disk manager. The disk manager then retrieves the block and sends it to file manager, which sends the required record to data manager. The transaction manager supervises the data transaction that is carried out between the data manager, file manager, and the disk manager. The recovery manager keeps a check on the transacted data so that in case of system failure,the data can be protected (Figure 20.14).

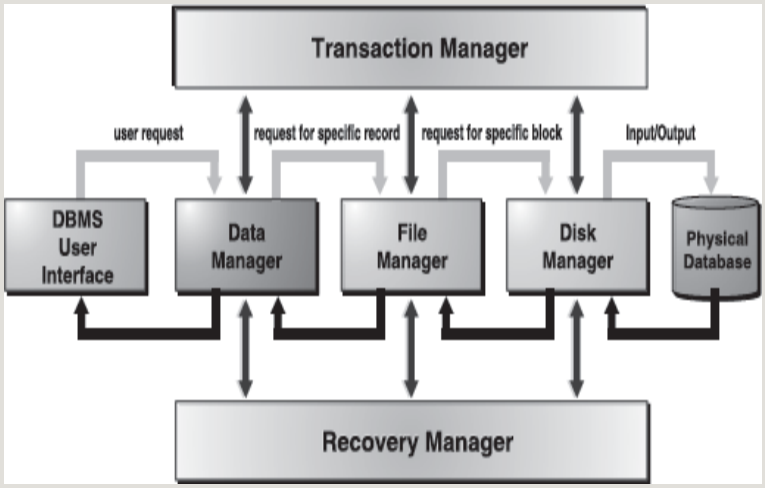


Figure 20.14 DBMS Components

Note: *A database manager is a program, which provides the interface between the physical database and the application programs and queries submitted to the system.*

20.5.4 Database Administrator

One of the main reasons for having DBMSs is to have central control of both data and programs accessing that data. The person having the central control over the system is called database administrator (*DBA*). The DBA is an information technology expert or a well-trained, computer literate person who is responsible for the technical operations of a database or all databases in an organization. In reality, the DBA may be a department or group rather than one person. They are the custodian of database accuracy, efficiency, maintenance, and development. Largely, they control access to information, ensuring that it is updated accurately and regularly. As part of a database team, they collaborate in the design and development of databases to meet ongoing user needs.

A DBA needs to stay abreast of emerging technologies and new design approaches. Typically, a DBA either has a degree in computer science and some on-the-job training with any database product or more extensive experience with a range of database products. Usually he/she should be experienced in major database management products, such as structured query language (SQL), SAP, and oracle-based database management software. DBAs should have an excellent knowledge of computers and organizational structures.

The DBA has many different responsibilities, but the overall goal of a DBA is to maintain the database system and to provide users with access to the required information when they need it. The DBA makes sure that the database is protected and that any chance of data loss is minimized. Typical responsibilities of the DBA are listed below:

- Granting different types of authorization to regulate which parts of the database various users can access.
- Ensuring regular backups of a database, and in case of failure (or disaster like fire or flood), suitable recovery procedures are used to restore the database services with as little downtime as possible.
- Ensuring regular and accurate update of database.
- Collaborating in the design and development of databases to meet new user needs and respond to anticipate technological innovations.
- Communicating regularly with internal technical applications and operational staff to ensure the database integrity and security.
- Identifying and resolving user's problems.
- Facilitating sharing of common data by overseeing proper key management.
- Procuring and maintaining database software and related documents and tools.

20.5.5 Database Languages

A DBMS mainly provides two database languages, namely, *data definition language* and *data manipulation language* to implement the databases.

- **Data Definition Language (DDL):** DDL is used to create and delete database and its objects. These commands are primarily used by

the DBA during the building and removal phases of a database project.

- **Data Manipulation Language (DML):** DML is used to retrieve, insert, modify, and delete database information. These commands will be used by all database users during the routine operation of the database.

20.5.6 Data Dictionary

Apart from the data, the database also stores *metadata*, which describes the tables, columns, indexes, constraints, and other items that make up the database. In simple words, metadata is data about data. This metadata are stored in an area called the *data dictionary*. Hence, a data dictionary defines the basic organization of a database. It contains the list of all files in the database, the number of records in each file, and the names and types of each field. Data dictionaries do not contain any actual data from the database, but only the information for managing it.

Most database systems keep the data dictionary hidden from users to prevent them from accidentally destroying its contents. Different users use the dictionary in different ways. The DBA needs a dictionary to ensure consistency among the data items, to educate users about the database content, and to help ensure that different department defines the same data in the same way. The programmers may use it to ensure that they have the name and coding of the data items or segments correct in their programs. Managers may use it as a guide to decide what data could be made available to them.

20.6 DBMS ARCHITECTURE

A major purpose of a DBMS is to provide users with an abstract view of the database. This means that the system does not provide all the details of data, rather it hides the details of how the data are stored and maintained. However, in order for the system to be usable, data must be retrieved efficiently. This concern leads to the design of complex data structures for the representation of data

in the database. Since many database system end-users are not computer trained, the complexities are hidden from them through several levels of abstraction in order to simplify their interaction with the system.

Before getting into the details of the database system's architecture, let us discuss two important terms associated with it.

- **View:** A database can be accessed by many users, and each of them may have a different perspective or view of the data. A database system provides a facility to define different views of the data for different users. A *view* is a subset of the database that contains virtual data derived from the database files but it does not exist in physical form. That is, no physical file is created for storing the data values of the view, rather only the definition of the view is stored.
- **Schema:** The *database schema* refers to the overall structure of the database, that is, all the information that is going to be represented in the database like name of the tables in the database, data types of the each data item of each table, constraints, relationship between the tables, etc. Note that once the schema of the database is created, usually it is not changed. If in case it needs to be modified, only the DBA, who has access to manipulate the structure of any object in the schema, can modify it.

  According to the ANSI/SPARC DBMS Report (1977), a DBMS can be envisioned as a three-layered system: *internal, conceptual*, and *external*.

- **Internal Level:** It is the lowest level of data abstraction that deals with the physical representation of the database on the computer and thus is also known as *physical level*. It describes *how* the data are physically stored and organized on the storage medium. At this level, various aspects are considered to achieve optimal run-time performance and storage space utilization. These aspects include storage space allocation techniques for data and indexes, access paths such as indexes, data compression and encryption techniques, and record placement.
- **Conceptual Level:** This level of abstraction deals with the logical structure of the entire database and thus is also known as *logical level*. It describes *what* data are stored in the database, the relationships among the data, and complete view of the user's requirements without any concern for the physical implementation. That is, it hides the complexity of physical storage structures. The conceptual view is the overall view of the database and it includes all the information that is going to be represented in the database.
- **External Level:** It is the highest level of abstraction that deals with the user's view of the database and thus is also known as *view level*. In general, most of the users and application programs do not require the entire data stored in the database. The external level describes a part of the database for a particular group of users. It permits users to access data in a way that is customized according to their needs, so that the same data can be seen by different users in different ways, at the same

time. In this way, it provides a powerful and flexible security mechanism by hiding the parts of the database from certain users, as the user is not aware of existence of any attributes that are missing from the view.

Figure 20.15(a) shows the three-level architecture for *EMPLOYEE_INFO* database containing details like employee code, employee name, salary, etc.

Figure 20.15(b) can be interpreted as follows:

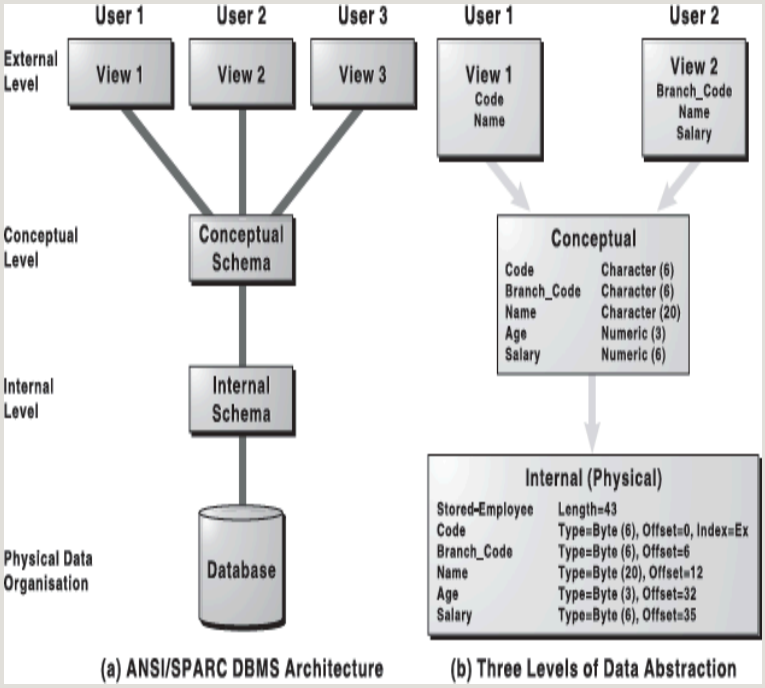


Figure 20.15 DBMS Architecture

- At the conceptual level, the database contains information regarding *EMPLOYEE* having fields like *Code* (6 characters), *Branch_Code* (6 characters), *Name* (20 characters), *Age* (3 digits), and *Salary* (6 digits). This is the conceptual schema of the database.
- At the physical level, STORED-EMPLOYEE represents the total record length (43 bytes). This structure contains all the five fields along with their byte representation and the pointer to the next record. Note that the records are indexed on *Code* field by an index called "*Ex.*" This is the internal schema of the database.

- External view is for two different users. The first user's external view incorporates only two fields, *Code* and *Name*. The second user's external view incorporates three fields, namely, *Branch_Code*, *Name*, and *Salary*. This is the external schema of the database.

The main advantage of the three-level database architecture is that it allows a clear separation of the conceptual view from the external data representation and from the physical data structure layout. This clear separation between different views results in data independence.

20.6.1 Data Independence

Data independence is the ability to change the schema at one level of the database system without having to change the schema at the other levels. Data independence is of two types, namely, *logical data independence* and *physical data independence*.

- **Logical Data Independence:** The separation of the external views from the conceptual view, which enables the users to change the conceptual view without affecting the external views or application programs, is called *logical data independence*. Examples of logical schema change can be addition/removal of entities in the database.
- **Physical Data Independence:** The separation of the conceptual view from the internal view enables us to provide a logical description of the database without the need to specify physical structures. This is often called *physical data independence*. Modification at the physical level is occasionally necessary in order to improve performance. Examples of physical schema change can be reorganization of files, adding a new access path, and so on.

Note: *Explicitly, the way in which the data are organized in storage devices and the techniques for accessing it are dictated by the requirements of the application under development. Moreover, knowledge of that organization and that access technique is built into the application logic and code.*

20.7 DATABASE MODELS

As discussed earlier, the conceptual level describes what types of data are actually stored in the database and the relationships among the data. However, how and what data structures are required to represent the data and the relationships among them depend on the data model being used for the DBMS. A *database model* or simply a

data model is an abstract model that describes how the data are organized and represented. A data model consists of two parts:

- A mathematical notation for describing the data and relationships
- A set of operations used to manipulate that data

Every database and DBMS is based on a particular database model. There are four basic types of database models: *hierarchical, network, relational*, and *object oriented*. These models provide different conceptualizations of the database and they have different outlooks and different perspectives.

20.7.1 Hierarchical Database Model

The hierarchical data model is the oldest type of data model, developed by IBM in 1968. This data model organizes the data in a tree-like structure in which each *child node* (also known as *dependents*) can have only one *parent node*. In other words, a hierarchal database is a collection of records connected to one another through links. The top of the tree structure consists of a single node that does not have any parent and is called the *root node*. The root may have any number of dependents; each of these dependents may have any number of lower-level dependents.

For example, consider an *EMPLOYEE_INFO* database that includes the record types (collection of similar type of records) *EMPLOYEE, PROJECT*, and *DURATION*. The *EMPLOYEE* record type includes the fields *Employee_Code, Employee_Name*, and *Salary* of an employee. The *PROJECT* record type includes *Project_Code* and *Project_Description*. The *DURATION* record type includes *No_of_hrs* spent by the employee on a project. Every employee is assigned some projects.

The hierarchical model for the *EMPLOYEE_INFO* database is shown in Figure 20.16. One complete record

of each record type represents a node. The node for an *EMPLOYEE* record is linked to all those nodes of the *PROJECT* record type that represent the projects on which that employee is working. In addition, each node of the *PROJECT* record type is linked to a specific node of *DURATION* record type.

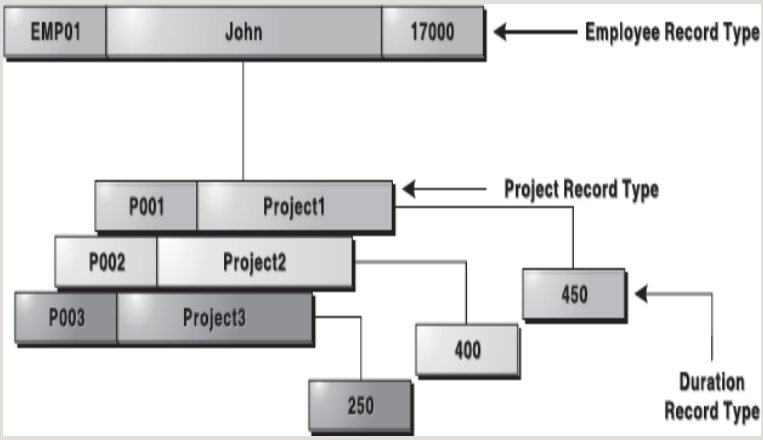


Figure 20.16 Hierarchical Database Model

The main advantage of the hierarchical data model is that the data access is quite predictable in structure, and therefore, both retrieval and updates can be highly optimized by a DBMS. However, the main drawback of this model is that the links are “hard coded” into the data structure, that is, the links are permanently established and cannot be modified. The hard coding makes the hierarchical model rigid. In addition, the physical links make it difficult to expand or modify the database, and the changes require substantial redesigning efforts.

20.7.2 Network Database Model

The first specification of network data model was presented by Conference on Data Systems Languages (CODASYL) in 1969, followed by the second specification in 1971. It is powerful but complicated. In a network model, the data are represented by a collection of records and the relationships among data are represented by links. A link is an association between precisely two records.

The network model of the *EMPLOYEE_INFO* database is shown in Figure 20.17. In this figure, each record of all the record types represents a node. All the nodes are linked to each other without any hierarchy. For instance, each *No_of_hrs* spent on a particular project by an employee is linked to that node of the *PROJECT* record type. All the nodes of the *DURATION* record type containing the *No_of_hrs* spent by the employee are linked to the *EMPLOYEE* record type (Figure 20.17).

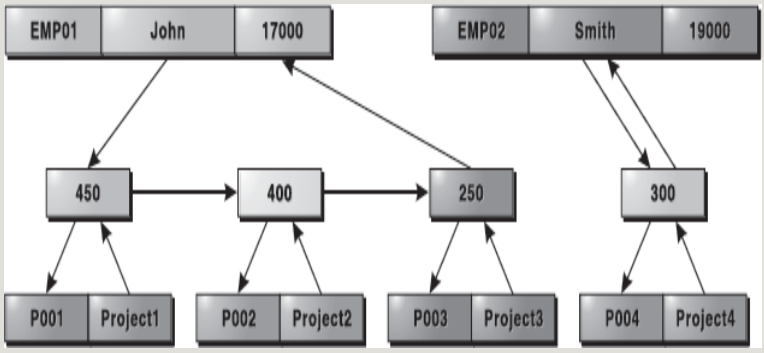


Figure 20.17 Network Database Model

The main limitation of the network data model is that it can be quite complicating to maintain all the links, and a

single broken link can lead to problems in the database. In addition, since there are no restrictions on the number of links, the database design can become overwhelmingly complex.

20.7.3 Relational Database Model

The relational data model was developed by E.F. Codd of IBM in 1970. The hierarchical and network model, though more flexible than the traditional file systems, were still not flexible enough. The limitations of these database models led to the development of the relational data model. The key difference between the network and hierarchical database models, and the relational model is in terms of flexibility. It is relatively easy and quick to create a new database structure and amend the existing structures in relational systems as compared to a network and hierarchical database.

The relational data model represents the database as a collection of simple two-dimensional tables called *relations*. The rows of a relation are referred to as *tuples* and columns of a relation are referred to as *attributes*. The relationship between the two relations is implemented through a common attribute in the relations and not by physical links or pointers.

The relational model of the *EMPLOYEE_INFO* database is shown in Figure 20.18. In this figure, the Employee database includes three relations, *EMPLOYEE, PROJECT*, and *DURATION*. The *EMPLOYEE* relation includes the attributes *Employee_Code, Employee_ Name*, and *Salary*. The *PROJECT* relation includes the attributes *Project_Code* and *Project_Description*. The *DURATION* relation includes the attributes *Employee_Code, Project_Code* and *No_of_hrs*. Since an employee is assigned a project, the relations *EMPLOYEE* and *PROJECT* are linked with each other with the help of the relation *DURATION*.

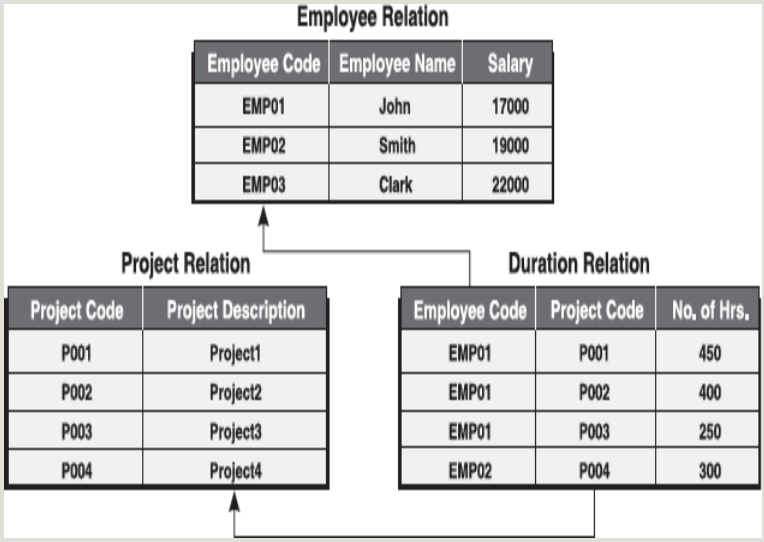


Figure 20.18 Relational Database Model

In a relational database, three basic operations are used to develop useful sets of data: *selection, projection*, and *join*. The selection operation retrieves certain records from a relation based on the user-specified criteria. The projection operation extracts fields from a relation, permitting the user to create new relations that contain only the required information. The join operation combines the data from the two relations based on a common column, providing the user with more information than is available in individual relations. Together, these three operations are part of *relational algebra*. Relational database systems use a query language called *SQL* to implement relational algebra operations.

Keys: Key is one of the important concepts of relational database. It can be *primary key, candidate key*, and *foreign key*.

Primary Key: An attribute or a set of attributes that uniquely identify each record in a relation is known as a *primary key*. This implies no two records in the relation can have same value for the primary key. In addition, the attribute chosen as primary key cannot accept null value. For example, in *EMPLOYEE* relation, the attribute *Employee_Code* is the primary key because all employee codes are unique. Note that there can be only one primary key in a relation.

Note: *The key composed of more than one attribute is known as composite key. Sometimes, it is also known as concatenated key or structured key.*

Candidate Keys: In a relation, there can be more than one attribute that can uniquely identify each record. All such attributes are known as *candidate keys*. One of these candidate keys is chosen as a primary key; the candidate keys that are not chosen as primary key are known as *alternate keys*. For example, in *EMPLOYEE* relation, there are two candidate keys, namely, *Employee_Code* and *Employee_Name* (considering no two employees can have the same name). If the attribute *Employee_Code* is chosen as a primary key, the attribute *Employee_Name* becomes the alternate key.

Foreign Key: An attribute of a relation that references the primary key of another relation is referred to as *foreign key*. Figure 20.19 illustrates how a foreign key constraint is related to a primary key constraint. Here, the attribute *Item_Code* in the *PURCHASE* relation references the attribute *Item_Code* in the *ITEM* relation. Thus, the attribute *Item_Code* in the *Purchase* relation is the foreign key (Figure 20.19).

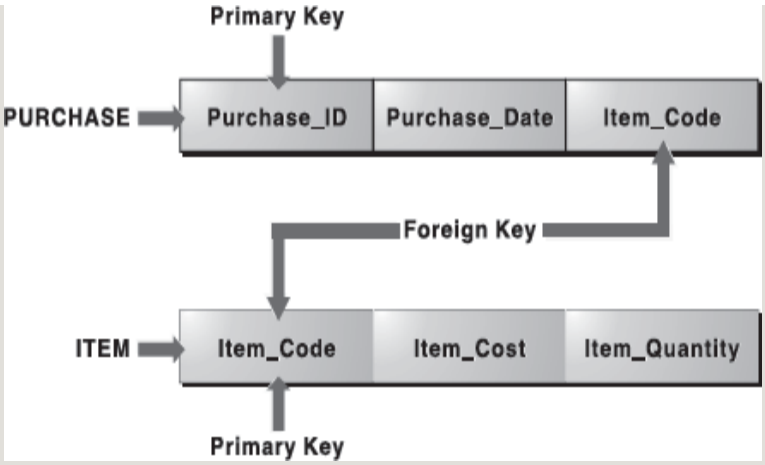


Figure 20.19 Foreign Key

Codd's Twelve Rules: The original concept for the relational model was proposed by Dr Edgar Codd in 1970. Later Dr Codd clarified his model by defining twelve rules (*Codd's Rules*) that a DBMS must meet in order to be considered a relational database. These rules provide a very useful yardstick for evaluating a relational system. The Codd's twelve rules are discussed below preceded by what is known as *rule zero*.

- **Rule Zero:** This rule states that all subsequent rules are based on the notion that in order for a database to be considered relational, it must use its relational facilities exclusively to manage the database.
- **Rule 1—The Information Rule:** This rule simply requires all information in the database to be represented in one and only one way, namely, by values in column positions within rows of tables. There is no hierarchical ranking of tables and the relationships among tables are logical only.
- **Rule 2—Guaranteed Access Rule:** This rule says that each item of data in an RDBMS is guaranteed to be logically accessible by specifying the name of the containing table, the name of the containing column, and the primary key value of the containing row.
- **Rule 3—Systematic Treatment of Null Value Rule:** According to this rule, a field should be allowed to remain empty (except for the primary key). This involves the support of a null value, which is distinct from an empty string or a number with a value of zero. Note that null means no value has been entered, that is, the value is unknown. For

example, if an item in a store is not marked, the price is unknown (null) but it is not free.

- **Rule 4—Dynamic Online Catalogue Based on the Relational Model Rule:** A relational database must provide access to its structure through the same tools that are used to access the data. This is usually accomplished by storing the structure definition within a special system catalogue. In addition to user data, a relational database contains data about itself. There are two types of tables in a RDBMS—user tables that contain the "working" data and system tables that contain metadata. The collection of system tables is also referred to as the system catalogue or data dictionary.
- **Rule 5—Comprehensive Data Sub-Language Rule:** In a relational system, there must be at least one language whose statements are expressible, per some well-defined syntax, as character strings and whose ability to support all the following is comprehensible:
  - Data definition
  - View definition
  - Data manipulation
  - Integrity constraints
  - Transaction boundaries

  Nowadays, all commercial relational databases use forms of the standard SQL as their supported comprehensive language.

- **Rule 6—View Updating Rule:** This rule specifies that all theoretically possible view updates should be possible. Each view should support the same full range of data manipulation that direct access to a table has available. In practice, providing update and delete access to logical views is difficult and is not fully supported by any current database.
- **Rule 7—High-Level Insert, Update, and Delete Rule:** This rule states that insert,update, and delete operations should be supported for any retrievable set rather than just for a single row in a single table.
- **Rule 8—Physical Data Independence Rule:** This implies that the user is isolated from the physical method of storing and retrieving information from the database. An application that accesses data in a relational database contains only a basic definition of the data (data type and length); it does not need to know how the data are physically stored or accessed. The changes can be made to the underlying architecture (hardware, disk storage methods, etc.) without affecting the access of the user.
- **Rule 9—Logical Data Independence Rule:** This rule specifies that the relationships among tables can change without impairing the function of applications and ad hoc queries. The database schema or structure of tables and relationships (logical) can change without having to re-create the database or the applications that use it.
- **Rule 10—Integrity Independence Rule:** In order to be considered relational, data integrity must be an internal function of the DBMS, not the application program. Data integrity means the consistency and accuracy of the data in the database. A minimum of the following two integrity constraints must be supported:
  - **Entity Integrity:** It implies that no component of a primary key is allowed to have a NULL value.

- **Referential Integrity:** It implies that if a foreign key is defined in one table, any value in it must exist as a primary key in another table.

- **Rule 11—Distribution Independence Rule:** Distributional independence specifies that the user should be unaware of whether or not the database is distributed (whether parts of the database exist in multiple locations).
- **Rule 12—Non-Subversion Rule:** This rule implies that there should be no way to modify the database structure other than through the multiple row database language (like SQL).

20.7.4 Object-oriented Database Model

The relational model has been successful in a wide variety of application areas. However, with the widespread networking of computers and the growth of Internet, the role of computers has expanded from its traditional base of transaction processing into new areas, many of which depend on the computer's multimedia capability. Whilst the relational model is well suited to transaction processing, it cannot manage the complex data structures such as image and sound, which are typical of multimedia applications. In addition, the relational model does not easily support the distribution of one database across a number of servers. Furthermore, there are several disciplines, such as medicine and multimedia, where relational database is not practical, due to the types of data involved. These fields of work needed more flexibility in data representation and accessibility. Due to these reasons, *object-oriented database management system* (*OODBMS*) was developed. In these databases, the users can define their own data access methods, the representation of data, and the methods of manipulating it. Similarly, new applications are emerging that are beyond the relational model's capability, such as those dealing with BLOBs (**B**inary **L**arge **OB**jects), like graphics, audio, and video. A relational database stores a BLOB outside the database that is referenced by pointers. These pointers allow the relational database to search for the BLOB, but the BLOB itself must be manipulated by conventional file I/O methods. However,

object-oriented databases provide native support to BLOBs.

The object-oriented model is a relatively new data model and provides an outlook for the future database models. An object-oriented database stores and maintains objects. An object is an item that can contain both data and the procedures that manipulate the data. For example, a student object might contain not only data about a student's name, roll number, and address, but also procedures on some tasks such as printing the student record or calculating the student's tuition fees. Like the other models, the object model assumes that objects can conceptually be collected together into meaningful groups known as classes. An object grouping is meaningful because objects of the same class have common attributes, behaviours, and relationships with other objects. Like the network model, the relationships among objects are specified via a “physical” link (pointer) between objects.

Figure 20.20 shows a simple example of object-oriented database structure. The class *Population* is the root of a class hierarchy, which includes the *Nation* class. The *Population* class is also the root of two subclasses, *Men* and *Women*. The Nation class is the root of other subclasses namely, *Country1*, *Country2*, and *Country3*. Notice that each class has its own set of attributes apart from the root class's attributes.

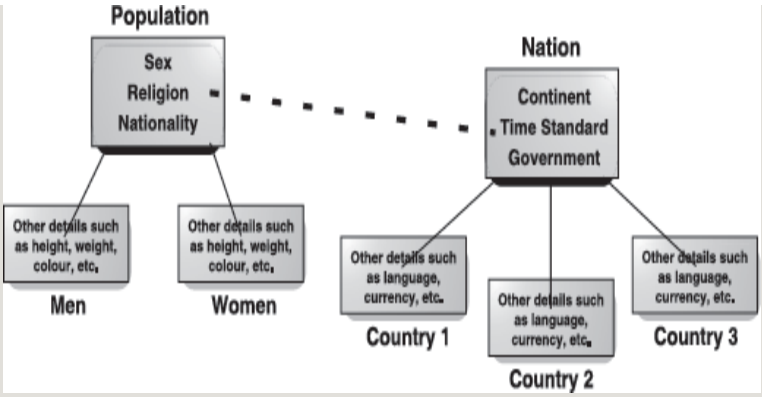


Figure 20.20 Object-oriented Database Model

Compared to relational database systems, OODBMSs can store more types of data and access these data much faster. With an object-oriented database, a system can store and access unstructured data such as pictures, video clips, sounds, and documents more efficiently. ObjectStore and Versant are some of the most common OODBMSs. Examples of applications appropriate for an object-oriented database include:

- Multimedia databases that store images, audio clips, and video clips.
- Computer-aided design (CAD) databases that store data about engineering, architectural, and scientific designs.
- Hypermedia databases that contain graphics, video, sound, and text links to other documents.

Object-oriented databases often use a query language called object query language (OQL) to manipulate and retrieve data. OQL is similar to SQL and they both use many of the same rules, grammar, and keywords. The main drawback of OODBMSs is its poor performance. Unlike RDBMSs, query optimization for OODBMSs is highly complex. Moreover, there is a lack of accepted standards for implementing OODBMSs. OODBMSs may not replace relational databases in conventional database markets such as inventory management, airline

reservation, finance, and investment management. The use of OODBMSs is likely to be restricted to complex applications such as design engineering and network management.

20.8 NORMALIZATION

While designing a relational database, the goal is to generate relations that allow storing information without redundancy. The process of reducing data redundancy in a relational database is called *normalization*. Essentially, normalization is a refinement process that the database designer undertakes after identifying the objects of the proposed database, their relationships, and defines the relations required and the attributes within each relation. This saves storage space and reduces the time queries taken to execute.

The fundamental principle of normalization is: *the same data should not be stored in multiple places*. No information is lost in the process; however, the number of relations generally increases as the normalization rules are applied. The main advantages of normalization are:

- It frees the database from certain insert, update, and delete anomalies.
- It reduces the need to restructure the database as new kinds of data are introduced.
- It avoids biasing the database design in favour of certain queries at the expense of others.
- It allows the operations needed for data access to be simple.

The normalization is the application of a set of simple rules called normal forms to the relation. There are six normal forms, namely, *first normal form (1NF), second normal form (2NF), third normal form (3NF), Boyce-Codd normal form (BCNF), fourth normal form (4NF), and fifth normal form (5NF)*. These forms are simply stages of database design, with each stage applying more strict rules to the types of information, which can be stored in a relation.

20.8.1 First Normal Form (1NF)

A relation is said to be in *1NF* if every attribute stores indivisible values. It implies any attribute in relation should not contain more than one value. For example, consider the relation *STUDENT* given below.

STUDENT Relation

Student_Name	Course_Name	Course_Fee	Course_Duration	Test_Marks	Discount (%)
Raja Khanna	.NET	10000	Six Months	80	5
Sameeksha Chopra	Java	8000	Three Months	80	5
Uma Handa	.NET	10000	Six Months	87	8
Nikhil Thakur	Networking	12000	One Year	90	10
Raja Khanna	Java	8000	Three Months	87	8

Clearly, this relation is in 1NF as every attribute contains indivisible values. However, this relation has certain anomalies. For example, if Nikhil Thakur (the only student of Networking) left the course, we must remove the corresponding row. This results in loss of information of Course Networking. Hence, this is not a good design of database and it must be normalized to next higher normal form.

20.8.2 Second Normal Form (2NF)

2NF depends on the concepts of the primary key and functional dependency (FD). FD is a certain constraint on the database that a relation must satisfy. FD is represented as A → B where *A* and *B* are attributes of a relation *R*. This FD implies that any two rows in *R* that have the same value in attribute *A* must also have same value in attribute *B*. In other words, the value of attribute

B in a row is determined by the value of attribute *A* in the row.

A relation is said to be in 2NF if it is in 1NF and every attribute is fully functionally dependent on the primary key. By fully FD, we mean that if primary key is composed of multiple attributes, then no attribute of relation should depend on a part of the primary key.

Now, consider our example relation *STUDENT*, which has the following functional dependencies.

Course_Name → Course_Fee, Course_Duration

Student_Name, Course_Name → Test_Marks

Test_Marks → Discount

If we designate Student_Name and *Course_Name* together as the primary key, then *Course_ Fee* and *Course_Duration* are not fully functionally dependent on the primary key, rather are partially dependent. By partial dependency, we mean *Course_Fee* and *Course_Duration* are dependent on *Course_Name* that is a part of the primary key. Note that Discount is functionally dependent on *Test_Marks*, which is fully functionally dependent on primary key *Student_Name*, *Course_Name*. Hence, we say *Discount* is transitively dependent on the primary key.

To normalize *STUDENT* relation to 2NF, partial dependency of *Course_Fee and Course_ Duration* must be eliminated by following the given steps.

1. A separate relation, say *COURSEDETAILS*, should be formed with attributes Course_ Name (as the primary key), *Course_Fee*, and *Course_Duration*.
2. Primary key of *COURSEDETAILS* relation should be included in the STUDENT relation as the foreign key.

COURSEDETAILS Relation

Course_Name	Course_Fee	Course_Duration
.NET	10000	6 months
Java	8000	3 months
Networking	12000	1 year

STUDENT Relation

Student_Name	Course_Name	Test_Marks (out of 10)	Discount (%)
Raja Khanna	.NET	8	8
Sameeksha Chopra	Java	8	8
Uma Handa	.NET	7	5
Nikhil Thakur	Networking	9	10
Raja Khanna	Java	7	5

Now, both the relations, *COURSEDETAILS* and *STUDENT*, are in 2NF. However, the STUDENT relation still has certain anomalies. For example, we cannot insert information of discount for score 6 in *Test_Marks* until a student scores 6 in the test. Hence, it must be normalized to the next normal form.

20.8.3 Third Normal Form (3NF)

A relation is said to be in 3NF if all the non-key attributes are independent, that is, no two non-key attributes of the relation depend on each other. The 3NF goes one step further in removing redundant data by removing attributes that are not fully dependent on the primary key. In essence, 3NF seeks to eliminate all attributes from a relation that are not directly dependent upon the primary key. Hence, a relation is said to be in 3NF if and only if it is in 2NF and every non-key attribute is non-transitively dependent on the primary key.

Consider the relations *COURSEDETAILS* and *STUDENT*. In relation *COURSEDETAILS*, non-key attributes are directly dependent on the primary key. Hence, it is in 3NF. However, in the *STUDENT* relation, the attribute discount is transitively dependent on the primary key. This dependency should be eliminated by creating a separate relation, say *DISCOUNTDETAILS*, consisting of *Test_Marks* (as the primary key) and Discount. Further, Test_Marks should be included as foreign key in the *STUDENT* relation.

STUDENT Relation

Student_Name	Course_Name	Test_Marks (out of 10)
Raja Khanna	.NET	8
Saneeksha Chopra	Jave	8
Uma Hnada	.NET	7
Nikhil Thakur	Networking	9
Raja Khanna	Jave	7

DISCOUNDETAILS Relation

Test_Marks (out of 10)	Discount (%)
7	5
8	8
9	10

Note: *There are additional normal forms (BCNF, 4NF, and 5NF) that are beyond the scope of this book.*

20.9 TYPES OF DATABASES

The modern business environment revolves around the accuracy and integrity of information. It is of utmost significance to every organization (big or small) today.

We can say that information is a key corporate asset and that it has to be managed, protected, and used like any other major asset. The corporate database that holds an organization's data is the underlying foundation for corporate information. Organizations are faced with questions regarding how and where to hold the corporate data. Advances in computer technology and the rapid development of graphical user interface (GUI)-based applications, networking, and communications have resulted in new dimensions in database management. Primarily, organizations adopt two approaches, *centralized* or *distributed* databases. If the entire database is kept in one centralized location, this type of database is a centralized database. On the other hand, if fragments of the database are physically placed at various locations, the database is a distributed database.

20.9.1 Centralized Database

In centralized databases, application processing is shared between numerous clients and a server. The client workstations are responsible for managing the user interface, and the database server is responsible for database storage and access. In this system, all database recovery, security, and concurrent access management are centralized at the database server. The server executes all data qualification statements and only the data that satisfy the criteria are sent to client stations. Examples of servers for centralized databases are DB/2 from IBM and Oracle Server from Oracle (Figure 20.21).

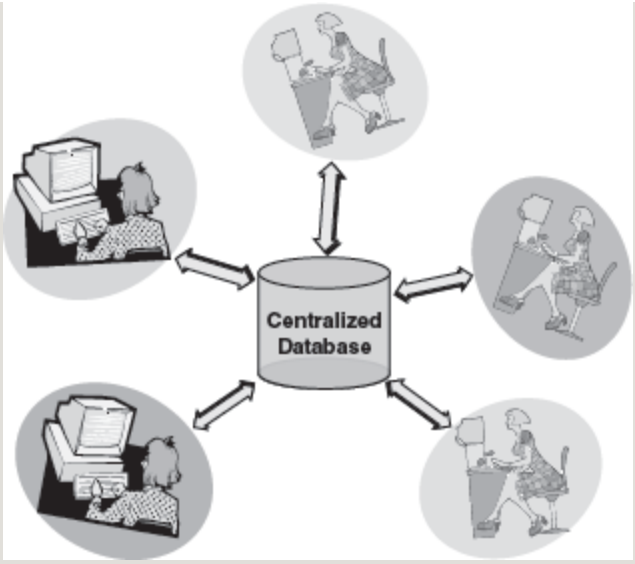


Figure 20.21 Centralized Database

Note: The central DBMS functions are called back-end functions; application programs on the client PCs are called front-end programs.

Advantages of Centralized Databases: The advantages of centralized databases are:

- Since multiple processors are being applied to the overall task, the back-end (database) and front-end (application) processing are being done in parallel. Thus, response time and throughput are improved.
- The server might be a custom-built machine that is tailored to the DBMS function and might thus provide better DBMS performance. Similarly, the front-ends may also be tailored to the needs of the end-users and thus be able to provide better interfaces, high availability, faster responses, and overall ease of use to the user.
- Several different client machines might be able to access the same server machine. Thus, a single database is shared across several different (and possibly widely different) front-end systems.

Disadvantages of Centralized Databases: The disadvantages of centralized databases are:

- Since all the data are stored at one place, any discrepancy may result in database corruption.
- The server must be able to grow in power and capacity to accommodate more clients, otherwise it will become a bottleneck.

- Centralized computing is more complex because proper processing requires close communication between clients and server; hence, specialized and expensive tools are necessary.

20.9.2 Distributed Database

A distributed database system consists of a collection of sites, connected together by the use of some kind of communication network. Each site is a database system site in its own right, and the sites have agreed to work together, so that a user at any site can access anywhere in the network as if the data were all stored at the user's local site. In other words, each site has its own local database, its own local users, its own local DBMS and transaction management software, and its own local data communication manager. In particular, a given user can perform operations on data at the user's own local site as if that site did not participate in the distributed system at all. The distributed database system can thus be regarded as a kind of partnership among the individual local DBMSs at the individual local sites.

Big corporations with widespread operations can benefit from distributed databases. In such organizations, computer processing is also distributed, with processing done locally at each location. A distributed database is fragmented into smaller data sets. Normally, the database is divided into data sets based on usage. If a fragment contains data that is most relevant to one location, that data set is kept at that location. Each fragment of data at every location may be managed with the same type of DBMS. For example, at every location Oracle DBMS may be used. In such case, the distributed database is referred to as a homogenous database. On the other hand, if the data fragments are managed at different locations with different DBMSs, then the distributed database is referred to as a collection of heterogeneous database systems (Figure 20.22).

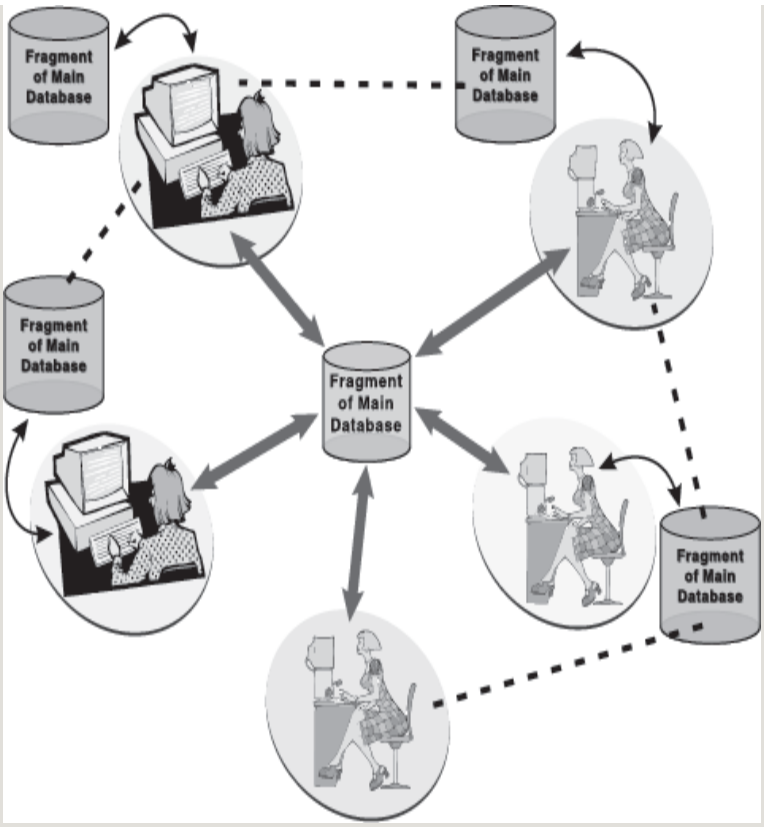


Figure 20.22 Distributed Database

Note: A distributed database is not the same as a decentralized database in which the computers are not connected by a network, so users cannot share data.

Advantages of Distributed Databases: The advantages of distributed databases are:

- When a centralized database fails, the database is unavailable to all users. A distributed system will continue to function at some reduced level even when a component fails.
- Certain operations can be processed at one site in the network, thus resulting in speed of processing and no communications overhead. The local organization is then fully responsible for the accuracy and safekeeping of the database.
- Usually, requests for data by users at a particular site can be satisfied by data stored at that site. This speeds up query processing.
- With a distributed system, data can be located closer to their point of use. This, like a centralized system, can reduce communication costs.

- Separate systems may be used to store data and/or process certain applications. These systems should have the capability to pass transactions among themselves. Data not available on one system may be obtained on another system.
- It is easier and more economical to add a local computer and its associated data to the distributed network than to expand a large central computer.

Disadvantages of Distributed Databases: The disadvantages of distributed databases are:

- Complex software (the DBMS) is required for a distributed database environment.
- The various sites must exchange messages and perform additional calculations to ensure proper coordination among the sites.
- A by-product of the increased complexity and need for coordination is the additional exposure to improper updating and other problems of data integrity.
- If the data are not distributed properly according to their usage, or if queries are not formulated correctly, response to requests for data can be extremely slow.

20.10 DATA WAREHOUSING AND DATA MINING

Data warehousing and data mining represent some of the latest trends in computing environment and information technology applications for large-scale processing and analysis of data. A *data warehouse* is a repository of subjectively selected and suitable operational data, which can successfully answer any ad-hoc, complex, statistical, or analytical queries. It is a collection of databases, data tables, and mechanisms to access the data on a single subject. Simply put, data warehouses store an aggregation of an organization's data. This vast database stores information like a data repository but goes a step further, allowing users to access data to perform research-oriented analysis (data mining). In simple terms, a data warehouse is a collection of data gathered and organized so that it can easily be analysed, extracted, synthesized, and otherwise used for the purposes of further understanding the data. It may be contrasted with data that are gathered to meet immediate business objectives such as order and payment transactions,

although this data would also usually become part of a data warehouse.

THINGS TO REMEMBER

Data Mining

Data Mining combine work from areas such as statistics, machine learning, pattern recognition, databases, and, more recently, high-performance computing. The goal of data mining is to discover interesting and previously unknown information in data sets. Tools for data mining have the ability to analyse enormous amount of data and discover significant pattern and relationships that might otherwise have taken a person thousands of hours to find. It is the exploration and analysis of large quantities of data in order to discover valid, novel, potentially useful, and ultimately understandable patterns in data.

According to William Inmon, the father of the modern data warehouse, a data warehouse is a subject-oriented, integrated, time-variant, non-volatile collection of data in support of decision-making. Based on this definition, we can envision four basic characteristics of data warehousing:

- **Subject Oriented**: A data warehouse is organized around a major subject such as customer, products, and sales. Data are organized according to subject instead of application. For example, an insurance company using a data warehouse would organize their data by customer, premium, and claim instead of by different products (auto sweep policy, joint life policy, etc.).
- **Non-Volatile**: A data warehouse is always a physically separate store of data. Due to this separation, data warehouses do not require transaction processing, recovery, concurrency control, and so on. The data are not updated if changed in any way once it enters the data warehouse, but is only loaded, refreshed, and accessed for queries.
- **Time Varying**: Data are stored in a data warehouse to provide a historical perspective. Every key structure in the data warehouse

contains, implicitly or explicitly, an element of time. The data warehouse contains a place of sorting data that is 5–10 years old, or older, to be used for comparison, trends, and forecasting.

- **Integrated**: A data warehouse is usually constructed by integrating multiple, heterogeneous sources such as relational databases and flat files. The database contains data from most or all of an organization's operational applications, and these data are made consistent.

Data warehousing involves the entire information delivery process—from access and transformation of data that resides in different operational stores, through the organization process that makes the data available for decision-making, to surfacing the data for exploitation using a range of decision support tools. The data warehousing process is given below:

- Data are extracted from each of the operational systems, which can include databases and external data (web feeds, purchased data, etc.).
- The data are transformed based on the business rules that apply to data usage.
- The data are stored in its transformed format in the data warehouse.

Data warehousing technology is becoming essential for effective business intelligence, business strategy formulation, and implementation in a globally competitive environment, wherein larger and larger amount of data is required to be timely processed for comprehension of its real meaning and impact. Rapid advances in both data storage and data capture technologies have resulted in a marked increase of data being stored in both business and scientific sectors. Due to these advances, millions of records are being generated, each of them containing tens or hundreds of fields. Many of these data sets are expanding on a daily basis. Traditional analysis of these types of data sets, involving human experts who manually analyse the data, are clearly no longer adequate. The field of data mining has developed for machine-oriented, automated methods for analysing large data sets. Using a combination of machine learning, statistical analysis, modelling techniques, and database technology, data mining finds patterns and subtle relationships in data and infers rules

that allow the prediction of future results. Some of the most common applications of data mining include market segmentation, customer profiling, evaluation of retail promotions, and credit risk analysis.

Data mining is practised by using certain software tools, which are geared for the user who typically does not know exactly what to search, but is looking for particular patterns or trends. These applications use pattern-recognition technologies as well as statistical and mathematical techniques and can have a key impact on the return on investment (ROI) for a technology expenditure upon discovering marketing or customer service data about one's clients. For example, searching and analysing customer sales detail to determine buying habits by ZIP code or other demographic criteria.

Data mining and data warehousing go hand in hand. Before one can effectively mine information, a data warehouse is constructed. Data warehousing involves integrating information from different systems, functions, and locations across an enterprise into a central database to allow more accurate analysis of customer needs, buying patterns, and profitability and improved decision-making and marketing. To use a simple analogy, it is finding the proverbial needle in the haystack. In this case, the needle is that single piece of intelligence the business needs and the haystack is the large data warehouse built up over a long period. With automated statistical analysis (data mining) techniques, businesses are discovering new trends and patterns of behaviour that previously went unnoticed. Once they have uncovered this vital intelligence, it can be used in a predictive manner for a variety of applications. Figure 20.23 illustrates a simplified process of mining data from a data warehouse.

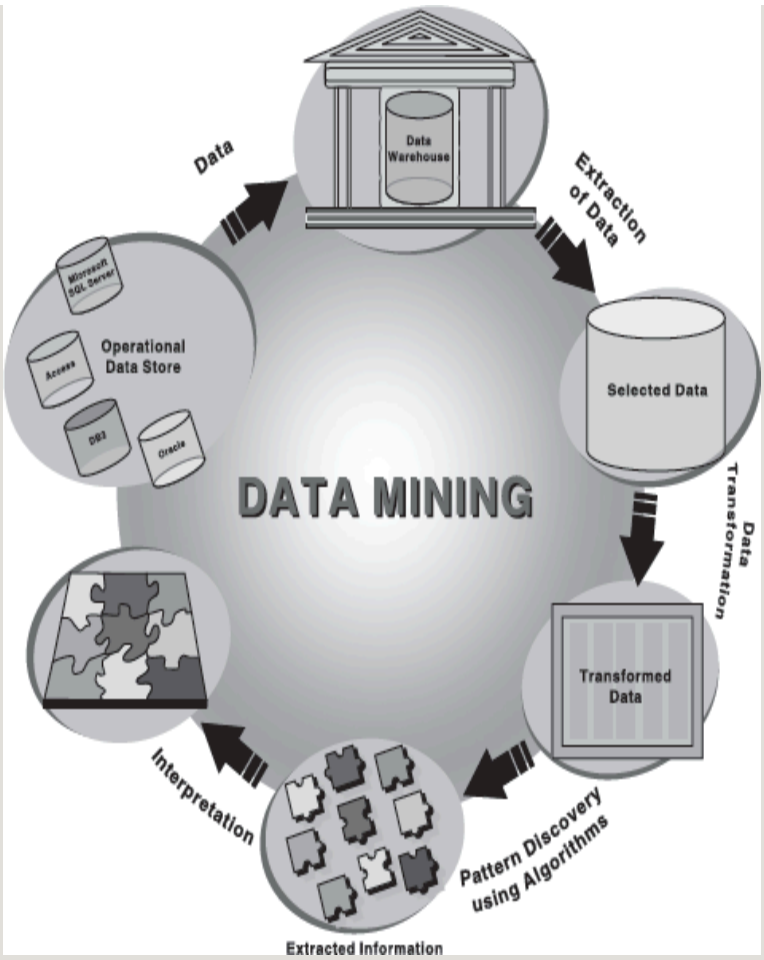


Figure 20.23 Data Mining Process

LET US SUMMARIZE

1. Data can be defined as a set of isolated and unrelated raw facts, represented by values, which have little or no meaning, simply because they lack a context for evaluation.
2. When data are processed and converted into a meaningful and useful form, it is known as *information*. Hence, information is a set of organized and validated collection of data. To be considered as valuable, information should be accurate, timely, complete, precise, and relevant.
3. Knowledge is the act of understanding the context in which the information is used. It can be based on learning through information, experience, guessing, and/or intuition.
4. A database can be defined as a collection of related data in an organized way. The organized information (that is, database) serves as a base from which the desired information can be retrieved, conclusions can be drawn, and decisions can be made, by further reorganizing or processing this data.

5. A field represents one related part of a table and is the smallest logical structure of storage in a database. It holds one piece of information about an item or subject.
6. A record is a collection of multiple related fields that can be treated as a unit and a table is a named collection of logically related multiple records.
7. An entity is a person, place, thing, activity, or event for which data are collected, stored, and maintained. Each characteristic describing a particular entity is called an attribute. In a database, entities are represented by tables and attributes are columns or fields.
8. Relationship is an association, dependency, or link between two or more entities. Generally, relationships are of three types: one-to-one (1:1), one-to-many (1:M), and many-to-many (M:M).
9. The relationship in which one record in a table is related to only one record in another table is known as one-to-one relationship. In one-to-many relationship, one record in a table can be related to many records in another table. When one record in a table can be related to one or more records in another table, and one or more records in the second table can be related to one or more records in the first table, the relationship is said to be many-to-many.
10. In a sequential file, the data are stored and/or retrieved in a sequence. The records are stored one after another in an ascending or descending order, based on the key field of the records. Usually, these files are stored on sequential storage devices like magnetic tapes.
11. A direct file allows accessing any record directly or randomly without having to traverse the sequence of the records. These files are stored on DASDs like hard disks.
12. Indexing is a technique of ordering the records in a table without making any modification in the original table. This technique is used by indexed sequential files, which uses a separate index file containing the key values and the location of the corresponding record. The records are organized in an orderly sequence and the index table is used to access the records without searching through the entire file.
13. DBMS can be defined as a collection of interrelated data and a set of programs to access that data. The primary goal of a DBMS is to provide an environment that is congenial and efficient to retrieve and store information.
14. The DBA is an IT expert or a well-trained, computer literate person who is responsible for the technical operations of a database or all databases in an organization.
15. A DBMS mainly provides two database languages, namely, DDL and DML to implement the databases. DDL is used to create and destroy database and its objects whereas DML is used to retrieve, insert, modify, and delete database information.
16. A data dictionary contains the logical structure of data in a database. It contains a list of all files in the database, the number of records in each file, and the names and types of each field.
17. A view is a subset of the database that contains virtual data derived from the database files but it does not exist in physical form.
18. Database schema refers to the overall structure of the database, that is, all the information that is going to be represented in the database like name of the tables in the database, data types of

the each data item of each table, constraints, relationship between the tables, etc.

19. According to the ANSI/SPARC, a DBMS can be envisioned as a three-layered system: internal, conceptual, and external. The internal level describes how the data are actually stored and organized on the storage medium. The conceptual level describes the type of data that is actually stored in the database, and the relationships among the data. The external level describes a part of the database for a particular group of users.
20. *Data independence* is the ability to change the schema at one level of the database system without having to change the schema at the other levels. The separation of the external views from the conceptual view is called as logical data independence and the separation of the conceptual view from the internal view is called *physical data independence*.
21. A database model or simply a data model is an abstract model that describes how the data are organized and represented. A data model consists of two parts: a mathematical notation for describing the data and relationships and a set of operations used to manipulate that data.
22. A hierarchal database model organizes the data in a tree-like structure in which each child node (also known as dependents) can have only one parent node. The top of the tree structure consists of a single node that does not have any parent and is called the *root node*.
23. In a network model, the data are represented by a collection of records and the relationships among data are represented by links. A link is an association between precisely two records.
24. The relational database model represents the database as a collection of simple two-dimensional tables called relations. The rows of a relation are referred to as tuples and columns of a relation are referred to as attributes. The relationship between the two relations is implemented through a common attribute in the relations and not by physical links or pointers.
25. An attribute or a set of attributes that uniquely identify each record in a relation is known as a primary key. In a relation, there can be more than one attribute that can uniquely identify each record. All such attributes are known as candidate keys.
26. An attribute of a relation that references the primary key of another relation is referred to as foreign key.
27. An object-oriented database model stores and maintains objects. An object is an item that can contain both data and the procedures that manipulate the data.
28. A refinement process that the database designer undertakes after identifying the objects of the proposed database, their relationships, and defines the relations required and the attributes within each table is known as normalization.
29. The normalization is the application of a set of simple rules called normal forms to the relation. There are six normal forms, namely, 1NF, 2NF, 3NF, BCNF, 4NF, and 5NF.
30. In *centralized* databases, application processing is shared between numerous clients and a server. The client is responsible for managing the user interface, and the database server is responsible for database storage and access.
31. A distributed database system consists of a collection of interconnected sites, where each site is a database system in its own right and the sites have agreed to work together, so that a user at any site can access anywhere in the network as if the data were all stored at the user's local site.

32. A data warehouse is a repository of subjectively selected and suitable operational data, which can successfully answer any ad-hoc, complex, statistical, or analytical queries.
33. *Data mining combines* work from areas such as statistics, machine learning, pattern-recognition, databases, and high-performance computing for discovering interesting and previously unknown information in data sets.

EXERCISES

Fill in the Blanks

1. In a table, columns represent ..... and rows represent ..... of data.
2. The ..... is an attribute or a set of attributes that uniquely identify each record in a relation.
3. The data, which describes the tables, columns, indexes, constraints, and other items that make up the database, are known as .....
4. In a database, entities are represented by ..... and attributes are represented by .....
5. .... refers to the duplication of data in a database whereas .....refers to ensuring that the data in the database is accurate.
6. The ability to modify a schema at one level without affecting the schema at other levels is known as .....
7. ..... is a program, which provides the interface between the physical database and the application programs and queries submitted to the database system.
8. ..... form of normalization removes transitive dependency.
9. ..... is a process where the data in a data warehouse can be analysed to reveal hidden patterns and trends in historical business activity.
10. In the ..... database model, data are organized in a tree-like structure.

Multiple-choice Questions

1. A DBMS provides ..... to retrieve, insert, modify, and delete database information.
    1. DDL.....
    2. DML
    3. XML.....
    4. Both (a) and (b)
2. Data integrity refers to .....
    1. Non-duplication of data
    2. Accuracy of data
    3. Security of data
    4. Centralized data
3. Which is NOT a component of a relational database?
    1. Entity
    2. Attribute
    3. Table
    4. Hierarchy

4. The of abstraction deals with the logical structure of the entire database.
    1. Physical level
    2. External level
    3. Conceptual level
    4. None of these
5. The ..... normal form is based on the concept of FD.
    1. First
    2. Second
    3. Third
    4. Fourth
6. A(n) is an attribute in a relation that serves as a primary key of another relation in the same database.
    1. Composite key
    2. Foreign key
    3. Identifer
    4. Primary key
7. Data are the ..... that describe people, events, things, and ideas.
    1. Words
    2. Graphics
    3. Numbers
    4. All of these
8. Father and son share ..... type of relationship.
    1. 1:M
    2. M:M
    3. 1:1
    4. None of these
9. The database model that uses a series of two-dimensional tables to store information is .....
    1. Object-oriented database model
    2. Hierarchical database model
    3. Relational database model
    4. Network database model
10. A logical collection of information that supports business analysis activities and decisionmaking tasks is
    1. An object-oriented database
    2. A data warehouse
    3. A database
    4. A data mining tool

State True or False

1. Knowledge is an ordered collection of facts, rules, understanding, and principles that can be written down.
2. Data and information are indistinguishable terms.
3. E.F. Codd developed the hierarchical DBMS model in the 1960s.
4. A relationship is an association between attributes.
5. A view is a subset of the database that contains virtual data derived from the database files; a physical file is created for storing the data values of the view.
6. Normalization is the process of decomposing relations to produce smaller, well-structured tables.
7. A distributed database requires multiple DBMSs, running at each remote site.

8. The hierarchical data model is the most commonly used model in contemporary database applications.
9. In a decentralized database, the computers are not interconnected by a network and users cannot share data.
10. A data warehouse is a repository of subjectively selected and suitable operational data, which can successfully answer any ad-hoc, complex, statistical, or analytical queries.

Descriptive Questions

1. What are some advantages and disadvantages of using indexed file organization?
2. List and describe any three benefits of database systems in comparison with file-oriented systems.
3. What is normalization and why is it done?
4. Is the relational data model better than the earlier data models? Give reasons.
5. Compare the features of the hierarchical, network, and relational data models. Explain how related data elements are linked in each model.
6. Enumerate in detail the fundamentals of database.
7. What do you mean by relationship? Explain various types of relationships.
8. What is DBMS? Explain its various components.
9. Explain in detail the architecture of DBMS.
10. Explain various Codd's rules and their importance.
11. What are the advantages and disadvantages of using a distributed database?
12. Give a descriptive note on the following topics:
    1. Hierarchical database model
    2. Network database model
    3. Relational database model
    4. Object-oriented database model

ANSWERS

Fill in the Blanks

1. Fields, Records
2. Primary key
3. Metadata
4. Tables, Fields
5. Data redundancy, Data integrity
6. Data independence
7. Database manager
8. Third
9. Data mining
10. Hierarchical

Multiple-choice Questions

1. (b)
2. (b)
3. (d)
4. (c)

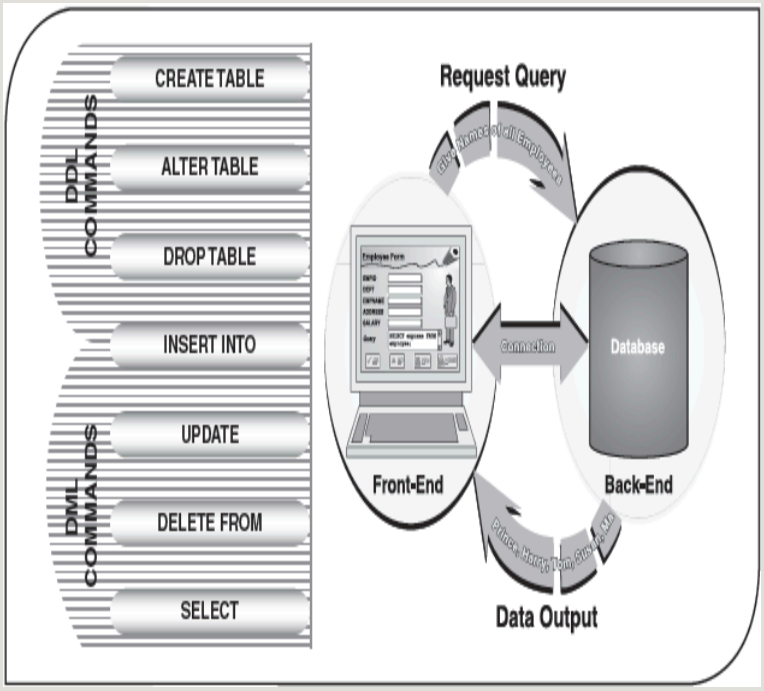
5. (b)
6. (b)
7. (d)
8. (a)
9. (c)
10. (b)

State True or False

1. True
2. False
3. False
4. False
5. False
6. True
7. True
8. False
9. False
10. True

Chapter 21

Structured Query Language (SQL)



This chapter introduces the reader to Structured QueryLanguage (SQL), which is a standard computer language foraccessing and manipulating databases. It is used by the useras a means of communication to the database managementsystem (DBMS). SQL commands consist of English-likestatements, which are used to query, insert, update, anddelete data. The chapter starts by describing the role andcharacteristics of SQL in a DBMS environment and how itevolved as a database language. The rest of the discussionis focussed on some of the most important and commonlyused SQL commands and queries.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- Structured query language, which is used to define and manipulate data stored in a database
- Concepts such as data definition language and data manipulation language
- DDL commands that are used to create or delete database objects such as tables and indexes
- DML commands, which operate on the data in a database
- Querying multiple tables that are used to maintain large databases
- SELECT statements, which can be used to break down a complex query into a series of logical steps

21.1 INTRODUCTION

In the previous chapter, we investigated the database, tables and records, and many such database-related issues. Database is a collection of interrelated information in a structured way, so that information can be easily retrieved, managed, and updated. Typically, a database is organized into tables, where each table is similar to a spreadsheet. A table has a set of columns (fields), each of which identifies some particular kind of information, and a set of rows (records) in which data are stored. There are lots of database management software available in the market such as Oracle and Sybase. Each database management system (DBMS) has its own programming language, using which a user can "talk" to the database and extract useful information. However, it would be difficult for a normal user to learn all these vendor-specific languages. For a user to write programs that interact with the databases easily, there has to be a way such that the user could use the same

method to obtain information from all these databases. Hence, a new language was developed to overcome this barrier and enable data to be retrieved from any database software. This query language is called *structured query language* (abbreviated SQL, pronounced as “see kwell” or “ess queue ell”).

THINGS TO REMEMBER

SQL

SQL eases the burden of programmers and other users, as they need not know how or where data are stored. They just have to specify what is required and how the specified data sets should be presented (for example, conditions or additional computations such as sum and average). It is just like querying in simple English, which allows the user to specify how the human conceptualizes the query in mind. Translating certain complex queries into SQL statements may be difficult and the constructs may not be powerful enough to define certain requirements.

SQL is an international standard (both **A**merican **N**ational **S**tandards **I**nstitute, ANSI, and **I**nternational **S**tandards **O**rganization, ISO, have standardized it) for defining and manipulating data sets from a database. Nowadays, almost all industrial-strength database vendors support SQL. Each vendor may have its own version of SQL implementing features, unique to that software, but all SQL-enabled databases support a common standardized subset of SQL. Using SQL, one can retrieve data from a database, create database and database objects, add data, modify existing data, and perform various other functions.

Note: *SQL provides a convenient method to create and manipulate databases. Moreover, SQL makes developing applications relatively easier.*

21.1.1 Communicating with Database Using SQL

SQL is simply a database language and by itself does not constitute a DBMS. It is just a medium of communicating to the DBMS what the user wants it to do. SQL is sometimes referred to as a *non-procedural* database language. It implies that the SQL statements describe *what* data are to be retrieved, rather than specifying *how* to find the data. The DBMS will take care of locating the requested data from the database. Generally, corporate databases are stored on several different computers. In such situations, the non-procedural nature of SQL makes flexible, ad-hoc queries for data retrieval. Users can construct and execute an SQL query, analyse the retrieved data, and change the query (if needed), all in a spontaneous manner (Figure 21.1).

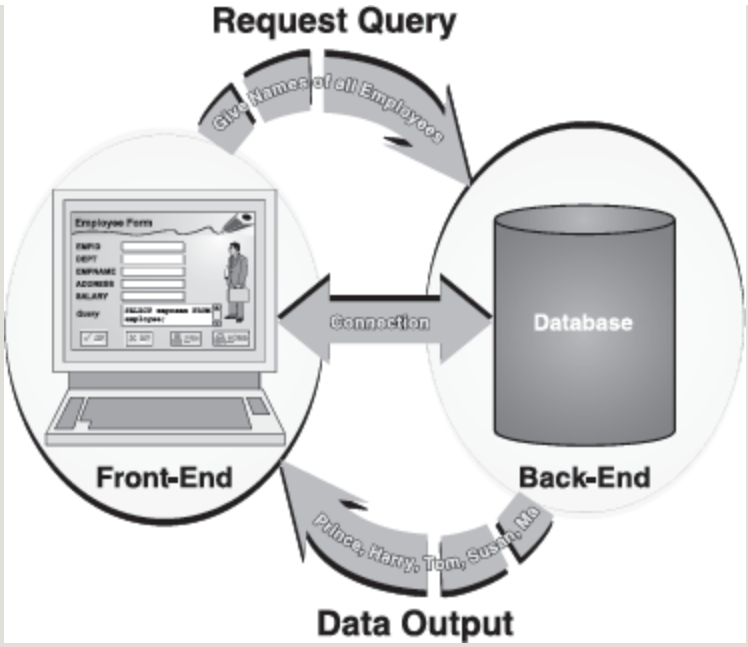


Figure 21.1 Communicating with Datavbase Using SQL

As discussed previously, SQL is used for communication with databases. The communicating parties are typically a *front-end* and a *back-end*. The front-end sends an SQL statement (containing instructions to create, read, change, or delete data) across a connection to a back-end that holds the data. Usually, the front-end includes an interface to users, such as Visual Basic or C++, with a "form" to gather data and buttons to carry out tasks. It may also contain fixed SQL statements or use code to create SQL statements on the fly using the data provided by the user. Back-ends are DBMSs such as Oracle and Sybase. They can accept an SQL statement and return a result. Usually, the back-end has two parts, the actual store of data and a database engine. The engine carries out the SQL statement and typically has a code that can search, read, write, index, and execute the actual interaction with the information. The front-end and back-end are linked via a "connection," which provides a

way to get the SQL statement to the back-end and then to return results to the front-end. A common connection used nowadays is ODBC (**O**pen **D**atabase **C**onnectivity).

By analogy, if the front- and back-end are like telephone handsets, then the connection is like a telephone connection over which a spoken language travels, although the language itself is free. To start a call, we first establish a connection by dialling the other party number and waits for the exchange to connect the correct wires. Then we can begin the conversation using a language like English across the connection. In the same way, a connection between front- and back-end software is created, and SQL statements are used to communicate across that connection.

21.1.2 Characteristics of SQL

SQL, tested and true over many years, is the standard way to communicate with a database. It is almost universally understood by both front-ends and back-ends. The amalgamation of SQL language and relational database management systems (RDBMSs) is one of the most cherished achievements of the computer industry. Over the last few decades, DBMS has grown tremendously and this alone is worth tens of billions of dollars, and SQL stands today as the standard database language. From its obscure beginning as an IBM research project, SQL has come of age as an important computer technology and a powerful market force. Some of the characteristics of SQL are briefed below.

- **Standard Independent Language:** The universal rules of the SQL have been established by ANSI and ISO. Therefore, the SQL is an open language, implying it is not owned or controlled by any single company. Today, SQL is offered by all the leading DBMS vendors. SQL is not just for a particular product; it works with Oracle, Microsoft SQL Server, and Sybase, just to name a few. Each vendor may have their proprietary extensions to SQL, but the basics of SQL are almost identical across all the database vendors. It means, regardless of the DBMS, the result of an SQL query will be the same. This vendor independence has enabled programmers to develop truly independent database applications.

- **Cross-Platform Abilities:** One of the biggest hardships of using a programming language to access database is that it seldom produces a “true” cross-platform application. SQL may be relatively new to some of the most popular programming languages like COBOL or C, but it has been in use on different hardware platforms for years. In most cases, the same SQL statement can be used on a desktop, a server, or a mainframe. An SQL-enabled database and the programs that use it can be moved from one DBMS to another vendor's DBMS with minimal (or no) conversion effort.
- **Easy to Learn and Use:** SQL statements resemble simple English sentences, making SQL easy to learn and understand. SQL has been created such that it is intuitive, simple, non-procedural (one need not specify step-by-step instructions to execute certain actions), and maps human's cognitive model. Being non-procedural in nature, the user just has to type a single SQL declaration and hand it to the DBMS. The DBMS then executes internal code (hidden from the user) and returns a set, which is a group of data that is logically defined.
- **Less Programming:** SQL allows extraction, manipulation, and organization of data with less programming as compared to traditional methods. SQL allows for record sets that are more complex to be created but requires less coding. It also allows tables to be joined together in ways that are not easily accomplished in other programming languages. Since SQL is a text-based programming language (that is, no executables are created for SQL-only programs), the SQL statement can also be constructed at run-time to give programs amazing flexibility. Furthermore, SQL can be used from within most programming languages. Whether you code in C++ or Java, SQL is the method of choice for accessing data.
- **Universality:** An alternative to using SQL statements is to write code in a procedural language like C++. The problem with this approach is that you are then closely tied to the procedural language, the metadata, and the specific DBMS. If there is a change in the structure of the tables, the code must be changed. If a new DBMS is installed, most of the codes must be revised to mesh with the new DBMS system of pointers, record set definitions, and so on. However, by using SQL statements, almost all changes are handled by the DBMS, behind the scenes from the programmer.
- **Scaling:** Many applications are developed for desktop environments. As they grow, a number of problems become apparent with these systems. For example, database software designed for the desktop generally fails when numerous people try to use the data at the same time. Therefore, organization has to move from desktop systems to a more robust DBMS. Almost all these heavy-duty DBMS rely on SQL as the main form of communication to scale well.
- **Speed:** Over the last 10 years, SQL data engines have been the focus of an intense effort to improve performance. The intense competition among database vendors has resulted in faster, more robust DBMSs that work at lower costs per transaction. Although SQL itself does not alleviate speed problems, the implementation of faster DBMS does, and those faster DBMSs require communications in SQL.
- **Cost Effective:** It is more expensive to run a server-centric DBMS (that only speaks SQL) than it is to use a desktop system (which does not necessarily need to use SQL). First, the software is more expensive. Second, in most cases, operating systems that are more expensive are needed to run in order to support the DBMS. Third, the operating system is required to be tuned differently to optimize for a DBMS than

for other applications. Last, personnel with more expensive qualifications are engaged. However, the price of hardware and software is frequently the smallest item in an IT budget. As the data centre grows, at some point, the reliability, performance, and standardization benefits of a DBMS that uses SQL outweigh the cost.

Nowadays, most commercial DBMSs allow SQL to be used in two different ways. First, SQL commands can be typed at the command line directly. The DBMS interprets and processes the commands instantly and displays the retrieved result rows (if any). This method is known as *interactive* SQL. The second method is called *embedded* SQL (or programmatic SQL), where SQL statements are embedded in a host language like COBOL or C. The host language provides more flexibility in SQL because SQL does not have any statements or constructs that allow a program to "branch" or "loop." The host language provides the necessary looping and branching structures, and the interface with the user, while SQL provides the statement to communicate with the DBMS. As a language, SQL plays many different roles:

- It is an *interactive language* because the user gains access to the data by issuing the desired query. This provides a convenient, easy-to-use tool for ad-hoc database queries.
- It is a *database programming language* because SQL commands can easily be embedded into application programs to access the data from a database.
- It is a *database administration language* because the database administrator uses SQL to define the database structure and control access to the stored data.
- It is a *client/server language* because it can be used to access remote (over a network) databases.
- It is a *distributed database language* because SQL helps in distribution of data across many connected computer systems.

SQL is a comprehensive language for controlling and interacting with a DBMS. However, it is not a program or a development environment such as Visual Basic or C++. There is no front-end built into SQL, neither does SQL have a back-end. There are no tools intrinsic to the language that can actually store data. SQL is only a standard means of communicating with software products that can hold data. As discussed previously,

SQL is non-procedural in nature. Consequently, it can execute only one command at a time. SQL is not a particularly structured language, especially when compared to highly structured languages such as C or Java. SQL does not provide decision-making constructs such as IF-THEN-ELSE, DO-WHILE, and so on. Therefore, the user cannot use the results of an action to perform another action. In addition, SQL does not provide any error-handling procedures. Hence, to get the best out of SQL, it is embedded into other languages, such as COBOL or C, to extend that language for use in database access. Besides, SQL itself is not a DBMS, nor a stand-alone product. SQL cannot be bought “off-the-shelf.” Instead, it is an integral part of a DBMS, and it acts as a tool for communicating with the DBMS. SQL has emerged as the standard language for using relational databases. Its overwhelming features make it the most popular and powerful database language available today.

21.1.3 Brief History of SQL

The history of SQL began in an IBM laboratory in San Jose, California, where SQL was developed in the 1970s. The publication of Dr Edgar Codd's rules resulted in a considerable amount of relational database research done in the early 1970s. By 1974, the IBM had surfaced with a prototype of a relational database called System/R. Although the System/R project ended in 1979, two significant accomplishments were accredited to that project. The relational data model's viability was sufficiently proven to the world and the project included significant work on a database query language. By the end of the System/R project, IBM had implemented a language that enabled the programmers to define, query, or manipulate database with simple English-like statements. This language was named as the **s**tructured **E**nglish **que**ry **l**anguage (SEQUEL). The name later was shortened to structured query language (SQL). Today,

many users still pronounce the abbreviation as “sequel” because of these early roots.

THINGS TO REMEMBER

History of SQL

A group of engineers watching the System/R project realized relational database's potential and formed a company named Relational Software Inc. In 1979, they produced the first commercially available RDBMS (called Oracle) and implemented SQL as its query language. Since then, many RDBMSs have come to market—all supporting SQL as their primary language.

SQL Standards: A database language standard specifies the semantics of various components of a DBMS. It defines the structures and operations of a data model implemented by the DBMS, as well as other components that support data definition, data access, security, programming language interface, and data administration. The SQL standard specifies data definition, data manipulation, and other associated facilities of a DBMS that supports the relational data model.

As discussed previously, SQL is a standard, open language without corporate ownership. The commercial acceptance of SQL was begun by the formation of SQL Standards Committees by the ANSI and the ISO. ANSI standardized SQL in 1986 and ISO standardized it in 1987. In 1989, a revised standard known commonly as **SQL89** (also called SQL1) was published. However, due to conflicting interests from commercial vendors, much of the SQL89 standard was intentionally left incomplete, and many features were labelled implementer defined. In

order to strengthen the standard, the ANSI committee revised the previous standard in 1992, and **SQL92** (also called SQL2) was conceived. Afterwards, various SQL standards were released including **SQL99, SQL2003**, and **SQL2006.** The latest standard is **SQL2008**, which is the sixth standard of SQL (Table 21.1).

Table 21.1 SQL Standards

Year	Name	Description
1986	SQL86	First published by ANSI in 1986 and ratified by ISO in 1987
1989	SQL89	Some minor revision
1992	SQL92	Major revisions were made
1999	SQL99	Added regular expression matching, recursive queries, triggers, non-scalar types, some object-oriented features
2003	SQL2003	Introduced XML-related features, standardized sequences, and columns with auto-generated values
2006	SQL2006	Defined ways of importing and storing XML data in an SQL database, manipulating it within the database and publishing both XML and conventional SQL data in XML form
2008	SQL2008	Added INSTEAD OF triggers and the TRUNCATE statement

21.2 GETTING STARTED WITH SQL

Previously, we defined SQL as a query language that allows access to data residing in a DBMS. Users execute queries to extract information from the database using criteria that are defined by the user. The main objectives of SQL queries are:

- Defining database and its objects such as tables, attributes, data type, relational keys (primary key, foreign key, etc.), and indexes.
- Defining data access controls on tables through Grants and Views.

- Altering, dropping, and replacing the tables and others database objects.
- Inserting, updating, and deleting data in the tables.
- Performing queries that involve table joins; various types of conditions and nested queries.

The first three objectives are generally called *data definition language (DDL)* since they involve definition and management of the basic structure of the database. DDL commands are used primarily by database administrators during the set up and removal phases of a database project. The last two objectives are referred as *data manipulation language (DML)*. DML is used to retrieve, insert, and modify database information and DML commands will be used by all database users during the routine operation of the database.

21.2.1 Data Types

SQL is considered a strongly typed language. This means that any piece of data represented by a table's field has an associated data type, that is, a set of rules describing a specific set of information, including the allowed range and operations and how information is stored. When the table is defined, every field in it is assigned a data type such as character, numeric, and so on. The type of a data value both defines and constraints the kinds of operations, which may be performed on it. SQL has numerous data types, such as *Char*, *Numeric*, and *Date*. Some of the most commonly used SQL data types are listed in Table 21.2.

Table 21.2 Common SQL Data Types

Data Type	Description
Char*(Size)*	It defines a fixed *Size*-length character string (can contain letters, numbers, and special characters), where *Size* can be a maximum of 255
Varchar*(Size)*	It defines a variable-length character string (can contain letters, numbers, and special characters) of up to *Size* characters
Integer	It defines an integer-only data type (usually a 32-bit signed integer)
Numeric*(S,D)* or Decimal*(S,D)*	It holds numbers with fractions. The maximum numbers of digits are specified in *S*. The maximum number of digits to the right of the decimal is specified in *D*
Date	This data type is used to store date. By default, the format is YYYY-MM-DD
Time	This data type is used to store time. By default, the format is HH:MM:SS
Boolean	This data type accepts a single value that can be *true* or *false*

21.2.2 SQL Syntax

SQL input consists of a sequence of commands, and each command is composed of a sequence of tokens, terminated by a semicolon (“;”), although it is not essential. The end of the input stream also terminates a command. The validity of tokens depends on the syntax of the particular command. A token can be a keyword, an identifier, a literal (or constant), or a special character symbol. Generally, tokens are separated by white space (space, tab, or new line). For example, the following is (syntactically) a valid SQL input:

```
SELECT * FROM EMPLOYEE;
UPDATE EMPLOYEE SET Salary = 10000;
```

This is a sequence of two commands, one per line (although it is not mandatory). More than one command can be on a line, or commands can be split across

multiple lines. Tokens such as `SELECT` and `UPDATE` in the above example are known as *keywords*, that is, words that have a fixed meaning in the SQL language. The tokens *EMPLOYEE* and *Salary* are examples of *identifiers*. They identify names of tables, columns, or other database objects. Note that, SQL identifiers and keywords must begin with a letter (a–z or A–Z) or underscore ("_"). Subsequent characters in an identifier or keyword can be digits (0–9), letters, or underscores. Remember that an identifier can never be a keyword.

There is a second kind of identifer: the *delimited identifer* (or quoted identifer). This identifer is formed by enclosing a sequence of characters in double quotes (" "). Therefore, "*SELECT*" could be used to refer to a column or table named "*EMPLOYEE*" whereas an unquoted *SELECT* would be taken as a keyword and would, therefore, provoke a parse error when used where a table or column name is expected. The example can be written with quoted identifers like this:

```
UPDATE "EMPLOYEE" SET "Salary" = 10000;
```

Quoted identifers can contain any character other than a double quote itself. This allows the user to construct table or column names that would otherwise not be possible, such as the ones containing spaces or ampersands (&).

Note: *In SQL, keywords, function names, and column names are not case sensitive. They can be given in any letter case. For example, SELECT DATE(), select date(), and SeLeCt DaTe() have same meaning.*

21.3 DDL COMMANDS

DDL consists of those commands in SQL that directly create or delete database objects such as tables and indexes, specify links between tables, and impose

constraints between database tables. The most important DDL statements in SQL are:

- **`CREATE TABLE`:** To create a new table.
- **`ALTER TABLE`:** To modify the structure of a table.
- **`DROP TABLE`:** To delete a table.

21.3.1 `CREATE TABLE` Command

The `CREATE TABLE` is used to define the structure of the table.

Syntax:

```
CREATE TABLE table-name(
<column1> <data-type>,
<column2> <data-type>,
... ... ... ... ...
<columnN> <data-type>);
```

Example:

```
CREATE TABLE EMPLOYEE(
Empid INTEGER,
Dept CHAR(10),
Empname CHAR(15),
Address CHAR(25),
Salary DECIMAL(8,2));
```

The above SQL statement creates a database table on disk and assigns it a name *EMPLOYEE*. Note that the table and column names must start with a letter, which can be followed by letters, numbers, or underscores.

When a table is created, each of the columns must be defined as a specific data type. For example, the *Empid column* is defined as an INTEGER and *Empname* is defined as CHAR(15). This means that when data are added to the table, the *Empid* column will only hold integers and the *Empname* column will hold character string values up to a maximum of 15 characters. Again, the column name must be unique within the table definition. Each column definition should be separated with a comma. Finally, use an opening parenthesis before beginning the table and a closing parenthesis after the end of the last column definition.

21.3.2 ALTER TABLE Command

The ALTER TABLE command allows a user to change the structure of an existing table. New columns can be added with the ADD clause. Existing columns can be modified with the MODIFY clause. Columns can be removed from a table by using the DROP clause.

Syntax:

```
ALTER TABLE table-name
<ADD | MODIFY | DROP column(s)<;
```

Example:

```
ALTER TABLE EMPLOYEE ADD Email CHAR(25);
ALTER TABLE EMPLOYEE MODIFY Empname CHAR(25);
ALTER TABLE EMPLOYEE DROP Dept;
```

The first `ALTER` command will add a new column named Email having a maximum width of 25 characters in the *EMPLOYEE* table. The second command will change the Empname to have a maximum width of 25 characters in the *EMPLOYEE* table. The last command will delete the Dept column from the *EMPLOYEE* table.

Note: *The `ALTER TABLE` command is not part of the ANSI/ISO standard. As a result, different vendors have different rules for this command.*

21.3.3 `DROP TABLE` Command

The `DROP TABLE` command removes the table definition (with all records).

Syntax:

```
DROP TABLE table-name;
```

Example:

```
DROP TABLE EMPLOYEE;
```

The above SQL command will delete the *EMPLOYEE* table.

21.4 DML COMMANDS

DML consists of those commands, which operate on the data in the database. These include statements that add data to the table as well as those statements that are used

to query the database. The most important DML statements in SQL are:

- **`INSERT`:** To insert data into a table.
- **`UPDATE`:** To update data in a table.
- **`DELETE`:** To delete data from a table.
- **`SELECT`:** To retrieve data from a table.

21.4.1 `INSERT` Command

In the previous sections, we discussed how to create, alter, and delete a table. Nevertheless, we did not discuss how to insert rows (records) in a table. We can add records in a table by using the INSERT command.

Syntax:

```
INSERT INTO table-name (
column1, column2, ..., columnN)
VALUES (value1, value2, ..., valueN);
```

Example:

```
INSERT INTO EMPLOYEE (
Empid, Dept, Empname, Address, Salary)
VALUES (101, 'RD01', 'Prince', 'Park Way', 15000)
```

The above example will add a new record at the bottom of the *EMPLOYEE* table consisting of the values in the parenthesis. Note that for each of the listed columns, a matching value must be specified. If one is inserting values corresponding to all columns of the table, then the

column list can be ignored. That is, in the above example, we can ignore the list of the columns since the order and the number of attributes match the *EMPLOYEE* table structure and the values correspond to that order. Hence, the above example can be rewritten as follows (refer to section 21.3.1 for *EMPLOYEE* table's structure):

```
INSERT INTO EMPLOYEE
VALUES (101, 'RD01', 'Prince', 'Park Way', 15000);
```

21.4.2 UPDATE Command

The `UPDATE` statement is used to make changes to the data in the database.

Syntax:

```
UPDATE table-name
SET column1 = value1
[, column2 = value2]
... ... ... ... ...        ——— Optional
[, columnN = valueN]
[WHERE condition];
```

Example 1:

```
UPDATE EMPLOYEE
SET Salary = Salary + 1000;
```

Example 1:

```
UPDATE EMPLOYEE
SET Salary = Salary + 1000;
```

The above command will update (in our case, increments) the Salary field with 1,000 for all the records.

Example 2:

```
UPDATE EMPLOYEE
SET Salary = Salary + 1000;
WHERE Dept = 'RD01';
```

The `UPDATE` command in Example 2 checks the `WHERE` clause first (`WHERE Dept = 'RD01'`). For all records in the given table in which the `WHERE` clause evaluates to true, the corresponding value is updated. In the example, the `WHERE` condition states that the `UPDATE` command will increase the value of *Salary* field by 1,000 for only those records where department (Dept) is "RD01." Notice in the `SET`, 1000 does not have quotes, because it is a numeric data type. On the other hand, "RD01" is a character data type, which requires the quotes.

The equal to (=) sign, used with the `WHERE` condition, is known as the *comparison operator*. Table 21.3 lists all the comparison operators used in SQL.

Table 21.3 Comparison Operators

Operator	Description
=	Equal to
>	Greater than
<	Less than
>=	Greater than or equal to
<=	Less than or equal to
<>	Not equal to

Example 3:

```
UPDATE EMPLOYEE
SET Salary = Salary + 1000, Dept = 'RD01'
WHERE Empname LIKE 'P%';
```

The above statement is an example of multi-column update. The LIKE pattern matching operator is used in the conditional selection of the WHERE clause. It is a very powerful operator that allows the user to select only rows that are *like* the specified condition. The percent sign "%" is used as a wildcard to match any possible character that might appear before or after the characters specified. Hence, any *Empname* that begins with "P" (for example, Paul, Prince, Priya) will be considered for the WHERE condition and the salary of all these *EMPLOYEE*s will be incremented with 1,000 and their *Dept* will be updated to "RD01."

21.4.3 DELETE Command

The `DELETE` statement is used to delete all or selected records from the specified table.

Syntax:

```
DELETE FROM table-name
[WHERE condition]; ← Optional Clause
```

Example:

```
DELETE FROM EMPLOYEE
WHERE Salary > 8000;
```

Example:

```
DELETE FROM EMPLOYEE
WHERE Salary > 8000;
```

The above statement deletes all the records from the *EMPLOYEE* table, which satisfy the WHERE condition. That is, records for all the *EMPLOYEE*s whose *Salary* is more than 8,000 will be deleted. Note that, if the `WHERE` condition is not used then all the records from the specified table will be deleted.

21.4.4 SELECT Command

Until now, we have discussed how to create, alter, and delete table. We also discussed how to insert, alter, and delete record(s) from a given table. Now comes the most important SQL command, `SELECT`. In fact, this command truly justifies the name of SQL. This command actually queries the database and retrieves the requested result set. `SELECT` allows users to retrieve data from one or more tables with various conditions. The `SELECT` statement allows the user to specify the desired data to be retrieved, the order to arrange the data, calculations to be performed on the data set, and many such

operations. The `SELECT` statement has a well-structured set of clauses. We will discuss some of the most commonly used variations of `SELECT` in the next few sub-sections.

Syntax:

```
SELECT column(s) or computed attribute(s) [AS new-name]
FROM table-name
[WHERE condition]
[GROUP BY column(s)]
[HAVING condition]
[ORDER BY column(s) [ASC|DESC]];
```

Optional

Before moving ahead with SELECT, let us first create a sample table with few records to execute our queries against it. Assume that the table name is *EMPLOYEE* and it has the following structure (Figure 21.2).

EMPID	DEPT	EMPNAME	ADDRESS	SALARY
101	RD01	Prince	Park Way	15000
102	RD01	Harry	Pebble Street	12000
103	RD02	Tom	Rose Garden	11000
104	RD02	Susan	Model Town	10000
105	ED01	Mark	Victor Crescent	10000
106	AD01	Francis	Chelmsford Park	13000
107	GR01	Robert	Downtown Cross	14000
108	RD03	Philip	Park Avenue	15000
109	RD03	Henry	Malibu Beach	14000
110	AD01	Frank	St. Peters Lane	7000

Figure 21.2 EMPLOYEE Table

Example 1:

```
SELECT Empid, Dept, Empname, Address, Salary
FROM EMPLOYEE;
```

The above example will display all the records of *EMPLOYEE* table as shown in Figure 21.2. Note that it is not mandatory to type the entire field list. You can use an asterisk (*) to substitute the field list.

```
SELECT * FROM EMPLOYEE;
```

The above `SELECT` command will also produce exactly the similar result, as it would do for Example 1. Remember, * is used only if one wants to display the data from all the fields. In case only selected fields are to be considered, the user has to give the name of the desired fields. For example, if one wants to display data from only two fields (*say Empname and Salary*), the following `SELECT` command must be used.

```
SELECT Empname, Salary FROM EMPLOYEE;
```

Example 2:

```
SELECT Empname, Salary
FROM EMPLOYEE
WHERE Salary > 10000 AND Dept = 'RD01'
ORDER BY Empname DESC;
```

Output:

```
Empname   Salary
-------   ------

Prince    15000
Harry     12000
```

The above example illustrates conditional retrieval of data. Let us analyse the SQL statement. Notice that the `WHERE` condition contains an additional keyword, `AND`. Just like its name, `AND` is used to specify more than one condition along with `WHERE`. It tests for multiple conditions, that is, all the conditions must be true for the execution of the query. In our example, the `WHERE` condition specifies that only those records should be fetched whose Salary field contains a value greater than 10,000, and such records should also have "RD01" as Dept field's value. One can also use the OR logical operator to join two or more conditions, and if any of the conditions is true, the result will be displayed. For example, if OR is used in the above example, then all those records will be retrieved whose Salary field contains a value greater than 10,000, or whose Dept field's value is "RD01."

The last clause in the above `SELECT` command is `ORDER BY`. It specifies that the fetched record set should be arranged (sorted) according to Empname field's value. In all the queries we have discussed so far, the rows of the results table have not been ordered in any way. SQL just retrieves the rows in the order in which it found them in the table. The `ORDER BY` clause facilitates imposing an order on the query results. By default, `ORDER BY`

arranges the result set in ascending order (whether one uses ASC or not). Nevertheless, one can fetch records in descending order as well by adding DESC clause along with `ORDER BY`, as we have done in the aforementioned example. Moreover, `ORDER BY` can be used with more than one column name to specify the ordering of the query result.

Using Aggregate Functions: By their very nature, the databases contain large volume of data. Previously, we explored various methods of extracting specific data using the SQL. Using such queries, one can find answers for certain questions like "What are the names of all *EMPLOYEE*s who have a salary greater than 12,000?" Frequently, users are also interested in summarizing the data to determine trends or produce certain reports. For example, the CEO of the company may not be interested in a listing of all employee salaries, but may simply want to know the sum of the salary paid. Fortunately, ANSI/ISO SQL provides five aggregate functions to assist with the summarization of large volumes of data. In this section, we will look at functions that can be used to add and average data, count records meeting specific criteria, and find the largest and smallest values in a table.

Aggregate functions are functions that take a collection of values as input and return a single value as the output. These functions summarize the result of a query instead of reporting each detail record. SQL has five built-in aggregate functions, which are listed in Table 21.4.

Table 21.4 Aggregate Functions

Function	Description
COUNT(column name)	Returns the number of rows of a column
COUNT(*)	Returns the number of selected rows
AVG(column name)	Returns the average or mean value of a column
MAX(column name)	Returns the highest value of a column
MIN(column name)	Returns the lowest value of a column
SUM(column name)	Returns the total summation of a column

Note: *Aggregate functions are also referred to as group functions.*

Example 1:

```
SELECT COUNT(*) FROM EMPLOYEE;
```

Output:

```
COUNT(*)
————————

10
```

The output of the above query is 10 because there are ten records in the *EMPLOYEE* table. The `COUNT(*)` total includes all the rows addressed by the query. In case the user is only interested in obtaining the number of employees who draw a salary of more than 12,000, the `WHERE` clause has to be used to retrieve the specific records. For example,

```
SELECT COUNT(*)
FROM EMPLOYEE
WHERE Salary > 12000;
```

Example 2:

```
SELECT AVG(Salary) AS Average_Salary FROM EMPLOYEE;
```

Output:

```
Average_Salary
______________

12100
```

The `AVG ()` function calculates the average or arithmetic mean of values in a column. Note that `AVG()` is only applied to numeric type of columns and it outputs a numeric value. SQL calculates the average by adding up all the values in the column, then dividing the total by the number of values.

Notice in the above example, we have used `AS` clause with `AVG ()` function. This clause is used to specify the heading for a column in the query output. The main reason behind specifying a new heading is that all too often, the column names or the calculated expression's heading contains cryptic names that do not make sense to users outside of a company's IT department. Hence, the AS clause provides the mechanism, which allows the user to change the headings displayed in query output for enhanced readability. Notice in Example 1, the heading of the query result is “COUNT(*).” This heading may not bother the query executor, but if this result is shown to anybody else (who many not have the same SQL awareness), there may be a problem of readability. However, if the heading reads something similar to that of Example 2 (“Average_ Salary”), then anyone could easily comprehend that the figure shown is the average salary.

Note: *The name used with the AS clause can be an expression but cannot contain characters (for example, spaces) that are not permitted in table field names.*

Grouping Data Using GROUP BY: The aggregate functions described previously have been used to obtain grand totals. Values output by them are just like the totals that appear at the end of each column listing in a report. These functions can also be used to output subtotal values by using the GROUP BY clause in the SELECT statement. This clause lets the user to split up the values in a column into subsets. That is, the GROUP BY clause divides a table into groups of rows. This clause specifies how the retrieved records are grouped for aggregate functions, and without the GROUP BY function it would have been impossible to find the sum for each individual group of column values. Suppose the user wants to count the total number of employees department wise, then the GROUP BY clause should be used. One can group by column name when selecting character-based data, or by the results of computed columns when using numeric data types. A GROUP BY clause is most useful in statements that include aggregate functions, in which case the aggregate function produces a value for each group. To simplify, let us discuss this clause with the help of an example.

Example:

```
SELECT Dept, COUNT(Dept) AS No_of_Employees FROM E
GROUP BY Dept;
```

Output:

```
Dept No_of_Employees
AD01 2
ED01 1
GR01 1
RD01 2
```

```
RD02 2
RD03 2
```

The above query first groups the rows in the *EMPLOYEE* table by the values in Dept. The `COUNT()` function then operates on each group, that is, the `COUNT()` function is performed for each department. The aggregate function `COUNT(Dept)` does not only apply to a single set of rows, but also to a group of sets of rows. The `GROUP BY` clause will form the group by the Dept and the `COUNT(Dept)` will count the number of employees with the same Dept. The `GROUP BY` clause is quite useful when an aggregate function is required to apply not only to a single set of rows, but to a group of sets of rows.

Note: *All fields in the `SELECT` field list must either be included in the `GROUP BY` clause or be included as arguments to an SQL aggregate function.*

Filtering Groups of Data Using HAVING:
Previously, we have used the `WHERE` clause to impose conditions to filter rows. However, sometimes we may want to state a condition that applies to groups rather than rows. For example, the user may be interested only in those departments in which the average salary of employees is greater than 13,000. This can be done by using the `HAVING` clause with the `GROUP BY` to select groups according to some condition. Adding a `HAVING` clause to a `SELECT` statement sets conditions for the `GROUP BY` clause in a similar way the WHERE statement sets conditions for the `SELECT` clause. The `HAVING` search conditions are almost identical to `WHERE` search conditions. The only difference is that `WHERE` search conditions cannot include aggregate functions and `HAVING` search conditions often include these functions. Note that, `HAVING` can only be used with `GROUP BY` Without the `HAVING`, it would not be possible to test for function result conditions. It can include as

many filter conditions as needed, connected with the `AND` or `OR` operator.

Example:

```
SELECT Dept AS Department, AVG(Salary) AS Average_
FROM EMPLOYEE
GROUP BY Dept
HAVING AVG(Salary) > 13000
ORDER BY AVG(Salary);

Output:
Department Average_Salary
__________ ______________
RD01       13500
GR01       14000
RD03       14500
```

In the above SQL query, firstly, the rows will be grouped according to *Dept* and the average *Salary* of each group will be calculated. Then the `HAVING` clause will set the filter condition that groups must meet to be included in the query results, that is, the average *Salary* of each *Dept* should be greater than 13,000. Finally, the results will be displayed according to the ascending order of the average *Salary*.

Using Character Functions: There are certain situations in which the user might wish to format the data in a particular manner. For character-type data, there are several character and string functions available, such as concatenation, trim, and substring functions. Note that different DBMS vendors have different string function implementations. Table 21.5 covers the most common character functions.

Table 21.5 Character Functions

Function	Description
LEN(expression) or LENGTH(expression)	Returns number of characters in the specified expression
INITCAP(expression)	Converts alphabetical character values to upper case for the first letter of each word, all other letters in lower case
LOWER(expression) or LCASE(expression)	Converts alphabetical character values to lower case, that is, the characters A–Z will be translated to the characters a–z
UPPER(expression) or UCASE(expression)	Converts alphabetical character values to upper case, that is, the characters a–z will be translated to the characters A–Z
LTRIM(expression)	Removes spaces from the beginning of expression
RTRIM(expression)	Removes spaces from the end of expression
CONCAT(first_string, second_string)	Joins two character expressions into a single string value
REPLACE(searched_string, search_key, replacement_string)	Replaces all occurrences of *search_key* in *searched_string* with *replacement_string*. Note the *replacement_string* is optional and if it is left out, each occurrence of the *search_key* on the string to be searched is removed and is not replaced with anything
SUBSTR(target_string, start_pos, end_pos) or SUBSTRING (target_string, start_pos, end_pos)	Extracts a specified portion of the character expression (*target_string)* on which it is operating. start_pos is the position of the first character to be output and end_pos is the number of characters to show

Example 1:

```
SELECT Empname, LEN(RTRIM(Empname)) AS Name_Length
FROM EMPLOYEE
WHERE Dept = 'RD03';
Output:
Empname Name_Length
_______ ___________
```

```
Philip    6
Henry     5
```

In the above example, `LEN()` function is used to ascertain the length of *Empname* column's values. We have used the `RTRIM()` function along with `LEN()` because while creating the *EMPLOYEE* table, the *Empname* column was stated as fixed-length character data type, that is, CHAR(15) (refer Section 21.3.1). Consequently, this column will automatically add spaces if the inserted value's length is not equal to 15. To be precise, if the value is “Prince,” actually it is stored in the table as “Prince-------” (- denotes space). In such a case, if only `LEN()` command is used, it will display 15 as output for all the rows irrespective of the actual length of *Empname*'s value. However, had we created the column with the VARCHAR, we would not have to use the `RTRIM()` function. In that case, simply using the `LEN()` would have accomplished the same output.

Example 2:

```
SELECT UCASE(Empname)
FROM EMPLOYEE
WHERE Salary > 13000;

Output:

UCASE(Empname)
PRINCE
ROBERT
PHILIP
HENRY
```

In the above example, `UCASE()` function is used to display the values of Empname column in upper case. Note that since `UCASE()` is used with `SELECT`, it actually does not convert the column's value in the table. For that, you will have to use the `UPDATE` command.

Example 3:

```
SELECT Empname FROM EMPLOYEE
WHERE SUBSTR(Empname,1,4) = 'Fran';

Output:

Empname

_______
Francis
Frank
```

In the above example, the `WHERE` condition selects only those rows from the *EMPLOYEE* table for which the employee name (Empname) starts with the word "Fran." Here 1 specifies the position in the string from where to start the substring and 4 specifies the length of the substring.

Using Mathematical Functions: Just as the character functions, SQL has some inbuilt mathematical functions that are used to retrieve data which involve mathematics. Most DBMSs provide arithmetic functions similar to the functions covered in Table 21.6.

Table 21.6 Mathematical Functions

Function	Description
ABS(expression)	Returns the absolute value of the expression
CEILING(expression) or CEIL(expression)	Returns the smallest integer value greater than or equal to expression
FLOOR(expression)	Returns the largest integer value less than or equal to expression
ROUND(expression, length)	Returns the expression rounded to length places right of the decimal point. If length is negative, expression is rounded to the absolute value of length places to the left of the decimal point
SQRT(expression) or SQR(expression)	Returns the square root of the expression
POWER(expression, p)	Raises the expression to the power of p
EXP(expression)	Returns the exponential function of expression
LOG(expression)	Returns the natural logarithm of expression
LOG10(expression)	Returns the base-10 logarithm of expression
SIN(expression), COS(expression), TAN(expression)	Returns the sine, cosine, and tangents of the specified expression, respectively

Example 1:

```
SELECT SQRT(10), POWER(10,2);
```

Output:

```
SQRT(10)          POWER(10,2)
_______           ___________

3.16227766017     100
```

In the above example, the `SQRT()` function evaluates the square root value of 10 and the `POWER()` functions raise the value of 10 to the power of 2.

Example 2:

```
SELECT Empname, ROUND(Salary/12.3,2) AS PF
FROM EMPLOYEE
WHERE ABS(Salary) > 13000;
```

Output:

```
Empname  PF
-------  --
Prince   1219.51
Robert   1138.21
Philip   1219.51
Henry    1138.21
```

The above example illustrates the use of mathematical functions on a table's column. The `ROUND()` function rounds off the specified numerical expression (*Salary*/12.3) to 2.

21.5 QUERYING MULTIPLE TABLES (SQL JOINS)

All the queries until this point were based on a single table only. However, SQL possesses a very powerful feature of gathering and manipulating data from across several tables. Without this feature, the database creator would have to store all the data elements necessary for each application in one table. Moreover, without common tables, the same data may have to be stored in several tables To solve this problem, SQL introduced the concept of *joins*, which enables the database designer to design smaller, more specific tables that are easier to maintain than a single larger table.

THINGS TO REMEMBER

SQL Joins

A join combines records from two or more tables in a relational database. In other words, join makes

relational database systems “relational.” There are three types of joins, namely, inner join, left outer join, and right outer join. These joins determine records that appear when a query is executed to show information from two tables. All these joins work by linking records on a key field.

Remember, we discussed the concept of foreign keys in the last chapter. A foreign key is a column in a table where that column is a primary key of another table, which means that any data in a foreign key column must have corresponding data in the other table where that column is the primary key. In DBMS terms, this correspondence is known as *referential integrity*. The join condition actually puts the concept of foreign key in practical use to link data from two or more tables together into a single query result—from one single SELECT statement. Good database design suggests that each table lists data only about a single *entity*, and detailed information can be obtained in a relational database, by using the join operation with other tables. A “join” can be recognized in a SELECT statement if it has more than one table after the FROM keyword.

Syntax:

```
SELECT list-of-columns
FROM table-name1 [AS alias-name], table-name2 [AS
[WHERE table1_keyfield = table2_foreign_keyfield],
```

Before going ahead with joins, let us create two tables so that we can “join” them to retrieve data (Figure 21.3).

EMPID	DEPT	EMPNAME	SALARY
101	RD01	Prince	15000
102	RD01	Harry	12000
103	RD02	Tom	11000
104	RD02	Susan	10000
105	ED01	Mark	10000

DEPT	JOB
RD01	Coding
RD02	Testing
ED01	Documentation

Figure 21.3 EMPLOYEE and DEPARTMENT Tables

Now, you might think that a natural extension of the SELECT statement is the following:

Example 1:

```
SELECT *
FROM EMPLOYEE, DEPARTMENT;
```

Output:

```
  Empid EMPLOYEE.Dept Empname Salary DEPARTMENT.Dept Job
  ----- ------------- ------- ------ --------------- ---
✓ 101   RD01          Prince  15000  RD01            Coding
  101   RD01          Prince  15000  RD02            Testing
  101   RD01          Prince  15000  ED01            Documentation
✓ 102   RD01          Harry   12000  RD01            Coding
  102   RD01          Harry   12000  RD02            Testing
  102   RD01          Harry   12000  ED01            Documentation
  103   RD02          Tom     11000  RD01            Coding
✓ 103   RD02          Tom     11000  RD02            Testing
  103   RD02          Tom     11000  ED01            Documentation
  104   RD02          Susan   10000  RD01            Coding
✓ 104   RD02          Susan   10000  RD02            Testing
  104   RD02          Susan   10000  ED01            Documentation
  105   ED01          Mark    10000  RD01            Coding
  105   ED01          Mark    10000  RD02            Testing
✓ 105   ED01          Mark    10000  ED01            Documentation
```

Although the above query is perfectly correct, but obviously, it did not retrieve the desired result. Every single row of the first table has been joined with each row of the second table, not just the rows that we think should correspond (the marked records). This is called a Cartesian join or cross-join and can rapidly generate enormous tables. According to this join, if there are two tables, each with a thousand rows, then 1,000 × 1,000 = 1,000,000 rows are displayed. Hence, to reduce the number of rows, and to get the desired result, simply use a WHERE clause to join the tables on the matching column, thus:

Example 2:

```
SELECT *
FROM EMPLOYEE, DEPARTMENT
WHERE EMPLOYEE.Dept = DEPARTMENT.Dept;
```

Output:

Empid	EMPLOYEE.Dept	Empname	Salary	DEPARTMENT.Dept	Job
101	RD01	Prince	15000	RD01	Coding
102	RD01	Harry	12000	RD01	Coding
103	RD02	Tom	11000	RD02	Testing
104	RD02	Susan	10000	RD02	Testing
105	ED01	Mark	10000	ED01	Documentation

In the above query, there are mainly two issues. Firstly, the matching columns (Dept) are used along with their table names with a dot operator (.). This is done to clearly mention the table to which the columns belong. In any case, if the table names are not mentioned along with the column, SQL will generate an error. Secondly, an asterisk (*) is used to display all the columns from

both joined tables to be listed. However, just like normal SELECT statements, we can use specific column names as well.

Example 3:

```
SELECT Empid, EMPLOYEE.Dept, Empname, Job
FROM EMPLOYEE, DEPARTMENT
WHERE EMPLOYEE.Dept = DEPARTMENT.Dept;
```

Observe that in the above query, the table names are not used with those columns that are not common in joined tables. Although the uncommon column names can also be listed along with their table names, it is up to the user discretion. Nevertheless, the common column(s) must be listed along with their table names. Thus, the above query can also be written as:

```
SELECT EMPLOYEE.Empid, EMPLOYEE.Dept, EMPLOYEE.Empname
FROM EMPLOYEE, DEPARTMENT
WHERE EMPLOYEE.Dept = DEPARTMENT.Dept;
```

Just as we have changed the column heading using the AS clause, a table can also have an alias name in the FROM clause. The alias is used wherever the table or column name would be used in other parts of the query. When an alias is defined, we can use the alias (instead of relation name) to refer to whose attribute we want to use in the query. For example, the previous join query can be rewritten as:

```
SELECT E.Empid, E.Dept, E.Empname, D.Job
FROM EMPLOYEE E, DEPARTMENT D
```

```
WHERE E.Dept = D.Dept;
```

21.5.1 Combining Queries Using UNION

There are occasions where the user might want to see the results of multiple queries together, combining their output. In such situations, UNION can be used. UNION is somewhat similar to “join” as they are both used to relate information from multiple tables. UNION adds the result of a second SELECT clause to the table created by the main SELECT command. One restriction of UNION is that all corresponding columns need to be of the same data type (numeric columns must correspond to numeric columns, string columns with string columns).
Moreover, when using UNION, only distinct values are selected. For example, to get the list of department IDs, the following query is used.

Example:

```
SELECT Dept FROM EMPLOYEE
UNION
SELECT Dept FROM DEPARTMENT;

Output:

Dept
____
RD01
RD02
ED01
```

Note: *With UNION, only distinct values are selected. To select all the values, use UNION ALL.*

A subquery is SELECT statement that nests inside the WHERE clause of another SELECT statement. In relational databases, there may be many situations when the user has to perform a query, temporarily store the result(s), and then use this result as part of another query. This nesting of queries is known as *subquery*. The basic idea behind it is that instead of a "static" condition, the user can insert a query as part of a WHERE clause. The main reason of using subquery is that it breaks down a complex query into a series of logical steps, and as a result, solves a problem with a single statement.

Syntax:

```
SELECT list-of-columns FROM table-name
WHERE attribute conditional operator
(SELECT list-of-columns FROM table-name
[WHERE condition]);
```

While creating a subquery, the following guidelines should be considered.

- The subquery must be on the right side of the conditional operator.
- The subquery must be enclosed within parenthesis.
- Multiple select subqueries can be combined (using AND, OR) in the same statement, but it is not advisable.
- When a subquery is part of a WHERE condition, the SELECT clause in the subquery must have columns that match in number and type to those in the WHERE clause of the outer query.

Example:

```
SELECT Empname, Salary FROM EMPLOYEE
WHERE Salary = (SELECT MIN(Salary) FROM EMPLOYEE);
```

```
Output:
Empname Salary
_______ ______
Susan 10000
Mark 10000
```

In the above example, the subquery `(SELECT MIN(Salary) FROM EMPLOYEE)` will return 10,000 as the result. This result will then be used by the main query to display the queried output.

LET US SUMMARIZE

1. SQL is an ANSI/ISO standardized computer language for Defining and manipulating data sets from a database. Using SQL, the user can retrieve data from a database, create databases and database objects, add data, modify existing data, and perform other, more complex functions.
2. SQL is a medium of communicating to the DBMS what the user wants it to do. SQL statements describe what data to be retrieved, rather than specifying how to find the data. The DBMS takes care of locating the requested data from the database.
3. Typically, there are two parties (*front-end and back-end*), which together use the SQL to process the requested query. The front-end sends an SQL statement (containing instructions to create, read, change, or delete data) across a connection to a back-end, which holds the data.
4. SQL can be used in two different ways: either SQL commands can be typed at the command line directly or it can be embedded in a host language like COBOL or C. The first method is known as interactive SQL and the second method is called embedded SQL.
5. The main characteristics of SQL are that it is an ANSI/ISO standard computer language and not owned or controlled by any single company. It is a cross-platform language, that is, the same SQL statements can be used on a multitude of hardware/software platforms. SQL is easy to learn and understand and requires less programming since most of the queries are of single line. However, SQL is not a complete programming environment because there is no front-end built into it; neither does it have a back-end.
6. The SQL standard specifies data definition, data manipulation, and other associated facilities of a DBMS that supports the relational data model. ANSI standardized SQL in 1986, and ISO standardized it in 1987. After that, SQL was revised as SQL89 (or SQL1), SQL92 (or SQL2), SQL99 (or SQL3), SQL2003, SQL2006, and SQL2008.
7. SQL input consists of a sequence of commands, and each command is composed of a sequence of tokens (keyword,

identifer, or literal), terminated by a semicolon. Note that, SQL is not case sensitive.

8. The `CREATE TABLE` command is to define the table's structure.

```
CREATE TABLE table-name(
<column1> <data-type>,
<column2> <data-type>,
... ... ... ...
<columnN> <data-type>);
```

9. The `ALTER TABLE` command is used to change the structure of an existing table.

```
ALTER TABLE table-name
<ADD | MODIFY | DROP column(s)>;
```

10. The `DROP TABLE` command removes the table definition (with all records).

```
DROP TABLE table-name;
```

11. The `INSERT` command is used to add records in a table.

```
INSERT INTO table-name (
column1, column2, ..., columnN)
VALUES (value1, value2, ..., valueN);
```

12. The `UPDATE` command is used to make changes to the data in the database.

```
UPDATE table-name
SET column1 = value1
[column2 = value2,
... ... ... ...
columnN = valueN]
[WHERE condition];
```

13. The `DELETE` command is used to delete all or selected records from the specified table.

```
DELETE FROM table-name
[WHERE condition];
```

14. The `SELECT` command is used to query the database to retrieve the requested result set. It allows users to retrieve data from one or more tables with various conditions. It also allows the user to arrange the data, to perform calculations on the data set, and many such operations.

```
SELECT column(s) or computed attribute(s) [AS new-name]
FROM table-name
[WHERE condition]
[GROUP BY column(s)]
```

```
[HAVING condition]
[ORDER BY column(s) [ASC | DESC]];
```

15. ANSI/ISO SQL provides five aggregate functions to assist with the summarization of large volumes of data.
    1. `COUNT()` returns number of rows of a column.
    2. `AVG()` returns the average or mean value of a column.
    3. `MAX()` returns the highest value of a column.
    4. `MIN()` returns the lowest value of a column.
    5. `SUM()` returns the total summation of a column.
16. SQL provides several character and string functions, such as:
    1. `LEN()/LENGTH()` returns the length of the specified string.
    2. `LCASE()/LOWER()` converts alphabetical character values to lower case.
    3. `UCASE()/UPPER()` converts alphabetical character values to upper case.
    4. `INITCAP()` converts alphabetical character values to upper case for the first letter of each word, all other letters in lower case.
    5. `LTRIM()` removes spaces from the beginning of expression.
    6. `RTRIM()` removes spaces from the end of expression.
    7. `CONCAT()` joins two character expressions into a single string value.
    8. `REPLACE()` replaces all occurrences of a search key from the searched string with a replacement key.
    9. `SUBSTR()/SUBSTRING()` extracts a specified portion of the character expression on which it is operating.
17. SQL provides several mathematical functions, such as:
    1. `ABS()` returns the absolute value of the expression.
    2. `CEILING()/CEIL()` returns the smallest integer value greater than or equal to expression.
    3. `FLOOR()` returns the largest integer value less than or equal to expression.
    4. `ROUND()` returns the expression rounded to length places right of the decimal point.
    5. `SQRT()/SQR()` returns the square root of the expression.
    6. `POWER()` raises the expression to the specified power.
    7. `EXP()` returns the exponential function of expression.
    8. `LOG()` returns the natural logarithm of expression.
    9. `LOG10()` returns the base-10 logarithm of expression.
    10. `SIN(), COS(), TAN()` returns the sine, cosine, and tangents of the specified expression, respectively.
18. SQL possess a powerful feature of gathering and manipulating data from across several tables by using joins. A *join* can be recognized in a SELECT statement if it has more than one table after the FROM keyword.

```
SELECT list-of-columns
FROM table-name1 [AS alias-name], table-name2 [AS alias-
[WHERE table1_keyfield = table2_foreign_keyfield];
```

19. `UNION` adds the result of a second SELECT clause to the table created by the main `SELECT` command.
20. A subquery is `SELECT` statement that nest inside the `WHERE` clause of another `SELECT` statement.

```
SELECT list-of-columns FROM table-name
WHERE condition
SELECT list-of-columns FROM table-name
[WHERE condition]
```

EXERCISES

Fill in the Blanks

1. SQL-based queries collect rows having same value in the.....clause in one cluster.
2. The language used for data modifcation is known as.....
3. .....clause is used to eliminate some groups from a SELECT query.
4. The technique of passing the result of inner query to outer query is known as.....
5. When SQL statements are incorporated into programs written in general-purpose programming language, it is known as.....
6. The part of a software program that a user sees and interacts with is known as.....
7. SELECT statements describe.....to retrieve, rather than.....to retrieve.
8. The three main SQL standards are....., ..., and .....
9. Group functions are also known as ..... functions.
10. Instead of specifying entire column list, ...can also be used to select all columns.

Multiple-choice Questions

1. You want to run a SELECT command that lists all employees by the ascending order of their names. What clause would you include in the query to do the ordering?
    1. LIST ASC BY EMPLOYEE.Name
    2. ORDER BY EMPLOYEE.Name
    3. SORT BY EMPLOYEE.Name
    4. WHERE EMPLOYEE.Name ASC
2. A language that is used to specify the database scheme through a set of definitions is called.....
    1. SQL
    2. DML
    3. DDL
    4. ISO
3. Which of the following is an aggregate function?
    1. PLUS
    2. ABS
    3. MINIMUM
    4. COUNT
4. SQL is.....type of language.
    1. Low level
    2. Procedural
    3. High level

4. Non-procedural

5. The query: `SELECT * FROM STUDENT WHERE Course LIKE '%A';` will.....

    1. Display all the courses that begin with the letter "A"
    2. Display all the courses that end with the letter "A"
    3. Display only those courses that contain numeric digits in the beginning
    4. Produce an error because the query is incorrect

6. What does SQL stand for?

    1. Structured question language
    2. Structured query language
    3. Structured question language
    4. Structured quarantine language

7. Which SQL command inserts new data in a database?

    1. INSERT
    2. NEW
    3. ENTER
    4. ADD

8. .....command extracts data from a database.

    1. OPEN
    2. EXTRACT
    3. GET
    4. SELECT

9. To select all the records from a table named "STUDENT," where the First_Name is "Mark" and the Last_Name is "Edward," which of the following SQL statement should be used?

    1. SELECT First_Name="Mark" Last_Name="Edward" FROM STUDENT
    2. SELECT * FROM STUDENT WHERE First_Name="Mark" AND Last_Name= "Edward"
    3. SELECT "Mark" "Edward" FROM STUDENTFirstName, STUDENTLastName
    4. It is not possible using SQL query.

10. .....function returns the smallest integer value greater than or equal to the expression.

    1. FLOOR(expression)
    2. CEILING(expression)
    3. ABS(expression)
    4. EXP(expression)

State True or False

1. In SQL, columns of different tables can have the same name.
2. ODBC is an example of back-end.
3. SQL join creates a new table by joining two or more tables.
4. SELECT Name Item Amount FROM CUSTOMER; is an invalid query.
5. SQL is owned jointly by IBM and Oracle.
6. The maximum limit of a VARCHAR data type is 256 characters.
7. In SQL SELECT, if there are more than one table after the FROM keyword, it represents a join.
8. SQL is case-insensitive language.
9. The AND operator displays a record if any conditions listed are true. The OR operator displays a record if ALL of the conditions listed are true.
10. The SQL REMOVE command deletes data from a database.

Descriptive Questions

1. How SQL is used in a DBMS to retrieve the data?
2. Discuss the characteristics of SQL in brief.
3. Consider the following tables:

Structure of EMPLOYEE Table

Field Name	Data Type	Description
EmpID	Numeric	Employee's ID number
DeptID	Numeric	Department's ID number
Emp_Name	Character	Employee's name
Employee's Salary	Numeric	Salary

Structure of DEPARTMENT Table

Dept_ID	Data Type	Description
DeptID	Numeric	Department's ID number
Dept_Name	Character	Department's name
Supervisor	Character	Department supervisor's name

Give the SQL expressions for the following queries with respect to the above tables.

1. Find names for all employees who work for the “Accounts” department.
2. For each employee, retrieve the employee's name and the name of his/her supervisor.
3. How many employees work in the “Accounts” department?
4. What are the maximum, minimum, and average salary in the “Accounts” department?
5. Retrieve the department names for each department in which more than two employees work.
6. Modify the database so that “Michael” is assigned to the “Accounts” department.
7. Give all employees of “Sales” department a 10% rise in salary.
8. Delete all rows for employees of “Apprentice” department.

4. There is a table named “PEOPLE” that has four columns: Country, Continent, Area, and Population. The first two columns are character types and the last two columns are numeric types. Now write the appropriate SQL queries for the following:

1. List the countries that have at least 5,00,000 people. [Show Country, Population]
2. Identify the countries, which have the word “United” included in their names.
3. List each country name where the population is larger than “United States.”
4. For each continent, show the number of countries with populations of at least 10 million.
5. List the continents with total populations of at least 100 million.

Fill in the Blanks

1. GROUP BY
2. DML
3. HAVING
4. Subquery
5. Embedded SQL
6. Front-end
7. What, How
8. SQL89, SQL92, SQL99
9. Aggregate
10. Asterisk (*)

Multiple-choice Questions

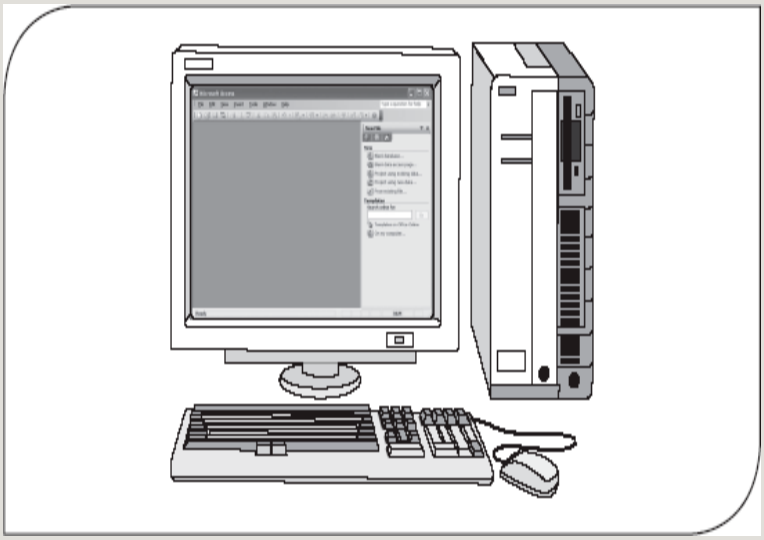
1. (b)
2. (c)
3. (d)
4. (d)
5. (b)
6. (b)
7. (a)
8. (d)
9. (b)
10. (b)

State True or False

1. True
2. False
3. False
4. True
5. False
6. False
7. True
8. True
9. False
10. False

Chapter 22

Microsoft Office Access 2003



This chapter provides a foundation to harness the power of databases using Microsoft Office Access 2003. Access is Microsoft's desktop relational database management application. It enables the user to effectively use the data stored in databases, create storage structures for data, manipulate and analyze that data, and format the data for output. In this hands-on chapter, you will learn to create simple databases, tables, queries, forms, and reports for viewing the contents of a database.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- Microsoft Access 2003 that allows the user to create a database
- Steps involved in creating a database
- Tables, which are used to store and organize data in the database
- Query, which is the basic tool used to retrieve information from the database
- Forms, which are used to create a graphical representation of the table
- Reports, which are designed to group, sort, and manipulate the data in a user-friendly manner

22.1 INTRODUCTION

Microsoft Access is a database management system that allows the user to create a relational database and then access, retrieve, and use the data in it. Access facilitates creation of database in several ways, manage database structures, import and export databases from/to various data sources and file formats, backup, restore, and perform other operations in easy-to-use graphical user interface.

Access is a part of Microsoft Office suite of applications, and of all the Office applications, only Access has the facility to be used as a relational database. One can store all kinds of information in an Access database, use queries to analyse the data, and create professional reports. Access also integrates well with other Office packages, and data transfer between Access and the other Office components is relatively easy. In addition, against the other desktop databases, Access is both rich in features and powerful. Lastly, Access can be used by almost any user. Beginners can learn to use the wizards and the easy-to-understand interface while developers can push it to its limits and do some extraordinary things with it. It is an application for everyone created with a mixed user level in mind. This chapter discusses the

basics of designing and creating an efficient and productive database using Access.

22.2 STARTING MICROSOFT OFFICE ACCESS 2003

To open Microsoft Access, perform *any* one of the following steps:

- Double-click the **Microsoft Access** icon located on the desktop.
- Click **start**, point to **All Programs**, then point to **Microsoft Office**, and then select **Microsoft Office Access 2003** (Figure 22.1).

Figure 22.1 Starting Microsoft Access

22.3 WORKING WITH ACCESS DATABASE

Whenever you start Microsoft Office Access 2003, you will see the Getting Started task pane, which enables to access online help and content, open existing database, and create a new database. To create a new database, click **Create a new file** on the Getting Started task pane. This displays the **New** File task pane which provides various options for creating a blank database, opening an existing database, or navigating to any database stored on your computer or on a network drive (see Figure 22.2). The New section in the New File task pane provides the following options:

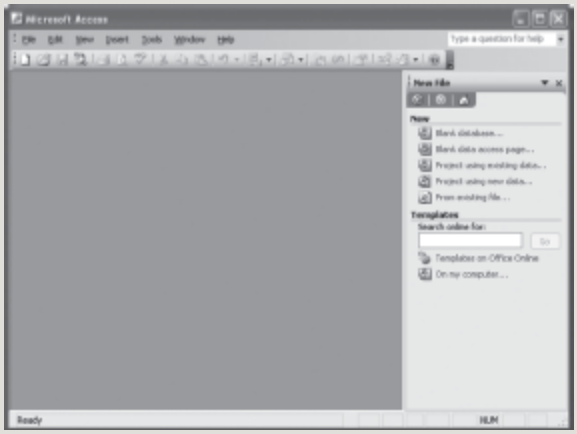


Figure 22.2 Access Opening Window

- **Blank Database:** To create a new database from scratch.
- **Blank Data Access Page:** To create a new data access page.
- **Project Using Existing Data:** To create database project using existing data.
- **Project Using New Data:** To create a database project using new data.
- **From Existing File:** To select the most recently used database from the list of existing databases.

22.3.1 Creating and Saving a Database

To create a new database, follow the steps given below:

1. Select **Blank database** from the **New** section in the New File task pane to display the File New Database dialog box (see Figure 22.3).

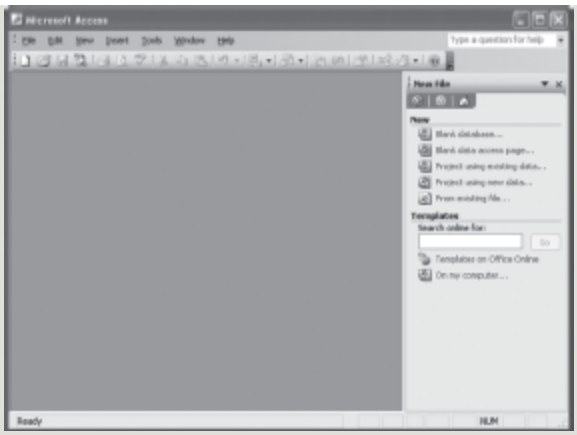


Figure 22.3 File New Database Dialog Box

2. Click the **Save in** list box to specify the location to save the new database. Enter the desired name of the new database in the **File name** text box (by default, Access names the database as db1.mdb). Click the Create button to **create** the new database.

Note: *Unlike other Microsoft Office applications, an Access database must be saved before one starts working on it.*

22.3.2 Opening a Database

To open an existing database, follow the steps given below:

1. Select **Open** from the **File** menu to display the Open dialog box. You can also open an Access database by clicking the **Open** button ( ) on the Database toolbar.
2. Specify the name of the desired file in the **File name** drop-down box. If the desired file is not in the current location, you can locate it by navigating through **Look in** list box. Select the desired file and click the **Open** button (Figure 22.4).

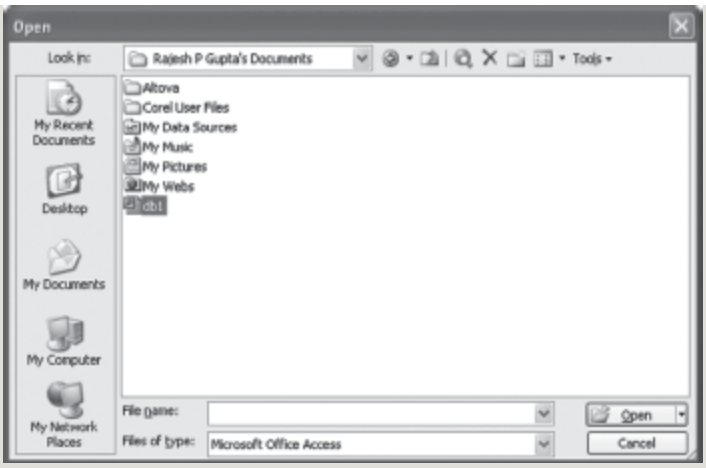


Figure 22.4 Open Dialog Box

Note: *Access database can also be opened by double-clicking the Access file icon placed on the specified location.*

22.3.3 Closing a Database and Exiting Microsoft Access

After finishing all the work in Access database, you may want to close the database and exit Access. For this, perform any of the steps given below:

- Select **Close** from the **File** menu to close the database and select Exit from **File** menu to close the Access program (Figure 22.5).

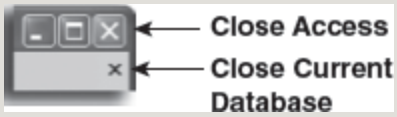


Figure 22.5 Closing a Database and Exiting Microsoft Access

- Click the **Close** button at the top-right corner of the Database window to close the current database. Note that when a database is opened in Access then there are two close buttons. One is located at the right-most corner on the title bar of the Access application window, which closes the Access application while the second at the top-right corner of the Database window which closes the current file.

Note: *The extension of an Access database is .mdb. When you are working in Access, a temporary file is created with .ldb extension. This file can be ignored as it disappears as soon as the database is saved and closed. In addition, a user cannot proceed further if the selected database has not been saved.*

When an Access database is created or opened, the main screen of the program appears (see Figure 22.6). This main window has many different parts, which are described in detail below.

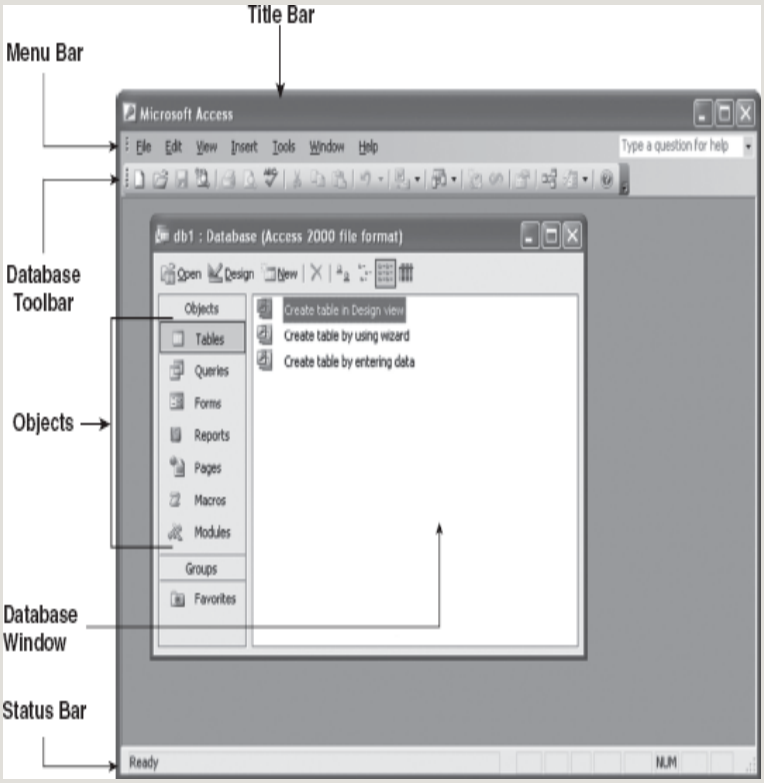


Figure 22.6 Parts of an Access Window

Title Bar: The bar across the top of the application window is the title bar. It displays the name of the

application. On the far right of the title bar are the **Minimize, Maximize/ Restore**, and **Close** buttons; these buttons are used to change the size or close the application window.

Menu Bar: Below the title bar is the menu bar, which contains different drop-down menus that allow the user to perform different functions in Access.

Toolbar: Generally, one or more toolbar appears just below the menu bar. The toolbar contains command buttons that provide single-click entry to carry out many commonly performed tasks. By default, Database toolbar is displayed in the Access environment. Additional toolbars like Query Design and Relationship toolbars can be added by right-clicking on the menu bar and selecting the **Customize** option from the pop-up menu.

Status Bar: The bar at the bottom of the Access window is the status bar which displays information specific to the task you are currently performing.

Database Window: In the middle of the Access desktop window is the Database window, the command centre of Access environment. The purpose of this window is to provide access to the objects in the database and allow the user to create new ones. At the top of the Database window is a bar which displays the name of the active database. The left side of the Database window has **Objects** and **Groups** section containing the list of objects that Access supports. It provides direct access to every object in the database. The options under **Objects** section are used to create, edit, or view the seven main object types in the Access database such as tables, reports, queries, and so on. Under **Groups** section, one can create shortcuts to the Objects in the database.

Database Objects: An Access database consists of a collection of tables, queries, forms, reports, pages,

macros, and modules which are used to manage and present the data. These components are also called **database objects**.

1. **Tables:** A table holds a collection of information stored in the form of fields and records. All the information necessary for the queries, forms, and reports in the database are based on the data in the table. One can sort and filter the information in many different ways. It is a place, which permits addition, deletion, or modification of records, fields, and field properties.
2. **Queries:** A query facilitates retrieval of information from multiple tables, which can be put together for viewing and analysis purposes. Using queries, one can select fields or records from different tables, sort the results, summarize, and perform calculations with the data.
3. **Forms:** Forms are used as the front-end for entering data. It is an object in Access, which allows the database users to edit and view information, one record at a time. All the information entered in a form will then be inserted into the corresponding table. Forms may be customized to have all or only some of the fields in a table. Forms are optional, but they help in making the data entry easier, especially if several people are entering data into the database.
4. **Reports:** A report provides total control of the presentation of the data. When information from the tables and queries are ready to be displayed, reports help in presenting this information in a formatted manner. One can add colours, put in summaries and calculations, add in appropriate titles, only show certain fields, and perform many other formatting details.
5. **Pages:** Pages (also known as data access pages) are web pages that contain data from the database and are published on the Web. A data access page can be used to view, add, or edit data stored in the database.
6. **Macros:** Macros are used to speed up the way a database works, for example, you could create a macro which automatically starts your database on a specific form.
7. **Modules:** Modules perform the same function as macros. They allow you to write small programs which can run within Access to speed up your work.

22.5 WORKING WITH TABLES

A *table* is a collection of data and is the first object that is added to the database. It contains rows and columns. The data entered in a row is called a *record*. For example, in a table EMPLOYEES, there are multiple rows and each row has a specific record of an employee. In Access, a table stores the information. Before we input the information, we must first create tables. There are quite a few ways to create Access tables. However, tables

can be mainly created in *Design view* or using the *Table Wizard*.

22.5.1 Creating a Table in Design View

Creating a table in Design view is very common because it offers several benefits. Design view allows you to define the fields in the table before adding any data to the datasheet. The table's Design view window is divided into two parts: a top pane for entering the field name, data type, and an optional description of the field, and a bottom pane for specifying field properties. To create a table in Design view, follow the steps given below:

1. Select **Tables** from the list of Objects in the Database window to display various options for creating tables in the right side of the Database window.
2. Double-click **Create table in Design view** to display the Table1 : Table window showing the design view of the table (Figure 22.7).

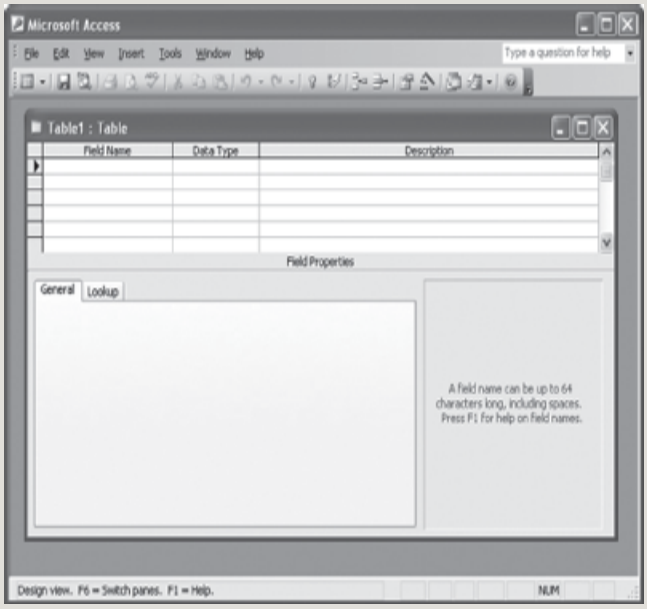


Figure 22.7 Design View of a Table

3. In the first row of the **Field Name** column, type in the name of the field and press Tab or Enter to move to the next column.

(Note that field names can contain letters, numbers, and/or spaces, and other characters, excluding periods, exclamation points, accents, and brackets.)

4. In the **Data Type** column, click the drop-down arrow to the right of the cell to display a drop-down list containing Access data types. Select an appropriate data type from the list and press **Tab or Enter** to move to the next column.
5. The **Description** column, while optional, is good to use for documentation purposes as well as for giving directions to the person entering the data.
6. Continue steps 3-5 until all the desired fields are created in the table (Figure 22.8).

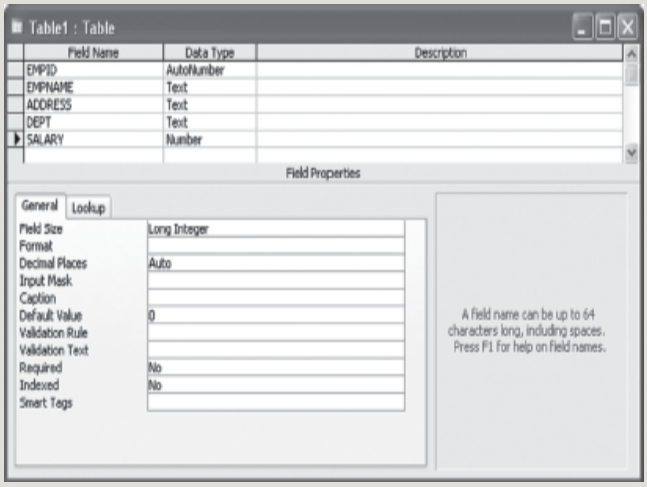


Figure 22.8 Creating Fields of a Table

7. Once the table is created, to save the table, click the **Save** button () on the Table Design toolbar or select **Save** from the **File** menu to display the Save As dialog box (Figure 22.9).

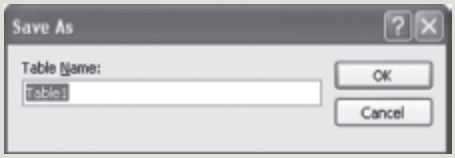


Figure 22.9 Save As Dialog Box

8. Enter the desired name of the table in the **Table Name** text box and then click **OK**. Microsoft Office Access message box appears

which prompts you to define a primary key for the table. Click **Yes** to create the primary key, or click **No** if no primary key is required, or click **Cancel** to remain in the Design view (Figure 22.10).

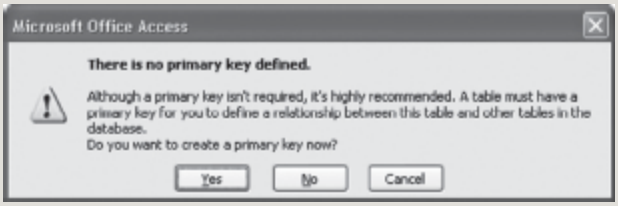


Figure 22.10 Microsoft Office Access Message Box

Setting Primary Key: A primary key is a field whose value uniquely identifies each record of the table. In other words, when a field is set as a primary key it ensures that no two records in a table can have the same value. While it is optional to set a primary key, it is a good practice to create one. For example, to set EMPID as the primary key of the above table, follow the steps given below:

1. Open the table in Design view. (Note that to open a table in Design view, select the desired table from table list, and then click the **Design** button (Design) on the Database window toolbar.)
2. Place the cursor in the EMPID field row and right-click to display a pop-up menu (Figure 22.11).

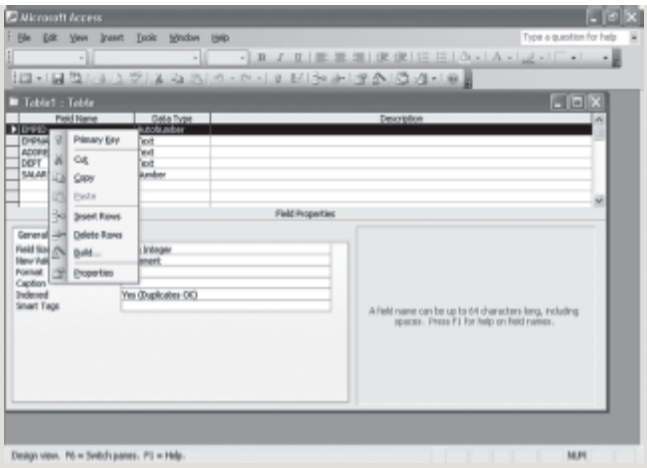


Figure 22.11 Setting Primary Key

3. Select the **Primary Key** option from the pop-up menu. This displays a key symbol ( ) in the record selector area adjacent to the field name EMPID which confirms that EMPID has been set as the primary key for the table (Figure 22.12).

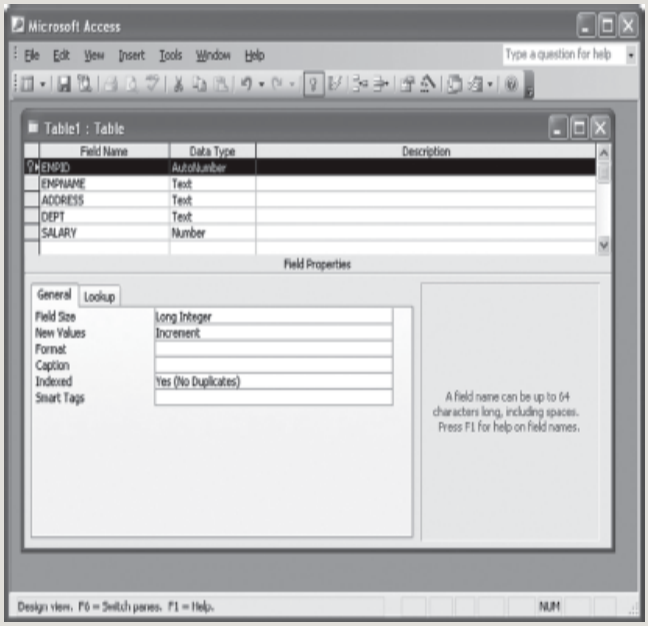


Figure 22.12 Design View after Setting EMPID as Primary Key

4. Click the **Close** button ( ) on the Table1 : Table window, Microsoft Office Access message box to confirm the changes made to the table's design appears. Click Yes to save changes and close the table's design view. Access takes you to the Database window.

Access Data Types: While creating a table, one must assign a data type for each field. The data type determines the type of values one can enter for a field in a table. Table 22.1 describes the data types available with Access 2003.

Table 22.1 Access Data Types

Data Type	Description
Text	It is used to store textual data that include letters, digits, spaces, and special characters. The default size of text field is 50 and it can contain up to 255 characters
Memo	It is used for text entries which are longer than 255 characters. It can contain up to 0 to 65,536 characters
Number	It is used for numeric values and depending on the field size, it can be 1, 2, 4, or 8 bytes
Date/Time	It is used for dates and times. The default size is 8 bytes
Currency	It is used for currency values. Currency fields are similar to the number data type, except that the decimal places and field size are predetermined, and calculation performed using the currency data type is not subject to round-off error. The default size is 8 bytes
AutoNumber	It is used when you want integers or a value automatically inserted in the field as each new record is created. The default size is 4 bytes
Yes/No	It is used for yes/no, true/false, or on/off values. The default size is 1 bit
OLE Object	It is used for data or files that are created in other software applications, such as photographs, video images, graphics, drawings, sound recordings, spreadsheets, word-processing documents, and so on. The size can be 0 bytes to 1 GB (depending upon what is stored within the field)
Hyperlink	It is used for storing text and numbers that can be used as a hyperlink for a web address. The size can be up to 64,000 characters
Lookup Wizard	This is not a data type, but this option creates a field that lets you select a value from another table or from a pre-defined list of values. Once the list is created, Access then sets the data type for you

22.5.2 Creating a Table Using Wizard

The Table Wizard helps the user to organize data by providing sample fields that one can include in a table and it assigns a primary key to each record. Later, if a change is required in the design of a table, one can

modify it using the Design view. To create a table using the Table Wizard, follow the steps given below:

1. Select **Tables** from the list of Objects in the Database window and double-click **Create table by using wizard** to display the Table Wizard dialog box.
2. The wizard has **sample tables**, which are categorized as, Business or Personal. The **Business** category contains tables like Customers, Products, Students, Employees, and so on. The **Personal** category contains tables like Household Inventory, Recipes, Photographs, and so on. Select an appropriate category to display the sample tables related to that category (Figure 22.13).

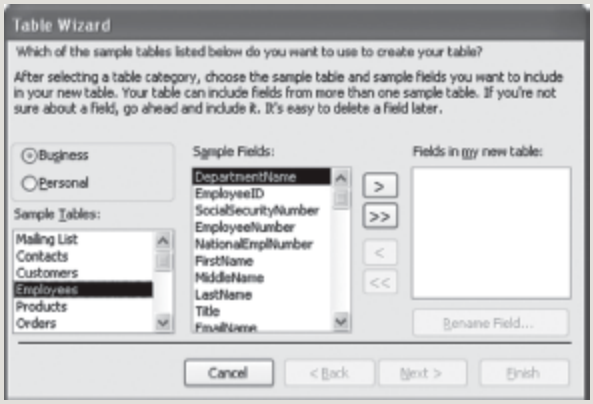


Figure 22.13 Selecting a Category

3. Select the appropriate table from the **Sample Tables** list box. The **sample fields** related to the selected sample table appears in Sample Fields list box.
4. Now, select the desired fields that you require in your table from the **Sample Fields** list box. To select the required field, either double-click the field, or simply select the field and click the (>) button. To select all the fields listed, click the (>>) button. You can click t(<) the button to remove a field from the list, or (<<) button to remove all the fields. The **Fields in my new table** list box displays the selected field(s) (Figure 22.14).

Figure 22.14 Selecting the Sample Fields

5. To change the name of a field, select the desired field in the **Fields in my new table** list box and click the Rename **Field** button to display the **Rename field** dialog box. Type the new name in the Rename field text box and click **OK**. Once the fields for the new table are selected, click **Next** to proceed (Figure 22.15).

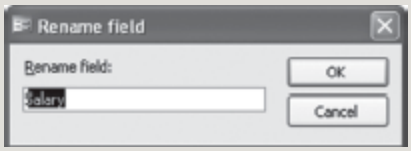


Figure 22.15 Rename Field Dialog Box

6. Enter the desired name for the table in **What do you want to name your table?** text box. In addition, this dialog box can also set a primary key for the table. Select **Yes, set a primary key for me** to create the primary key by Access, or select **No, I'll set the primary key** for manual assignment of the primary key and click **Next** (Figure 22.16).

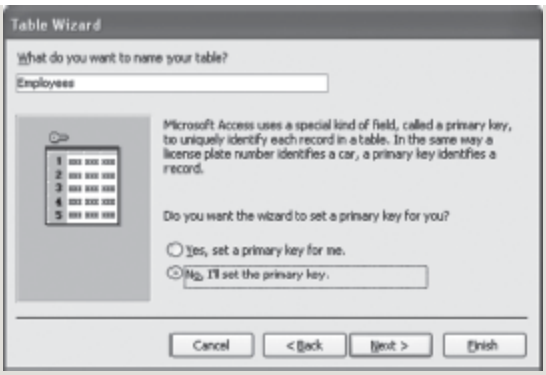


Figure 22.16 Entering Name and Selecting an Option for Setting an Primary Key

7. Note that when you select **No, I'll set the primary key**, a drop-down box containing the fields (that are selected by you for the table) appears. From the drop-down box, select the field to be set as the primary key and click **Next** (Figure 22.17).

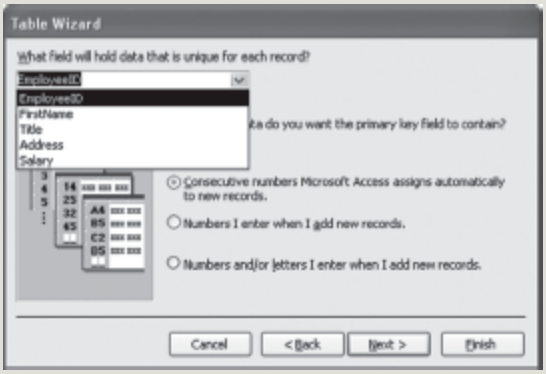


Figure 22.17 Setting Primary Key Manually

8. The last dialog box of the Table Wizard provides options to modify the newly created table in Datasheet view, or enter data manually, or via a form. Select **Enter data directly into the table** and then click **Finish**. Access creates and saves the table and then opens it in Datasheet view (Figure 22.18).

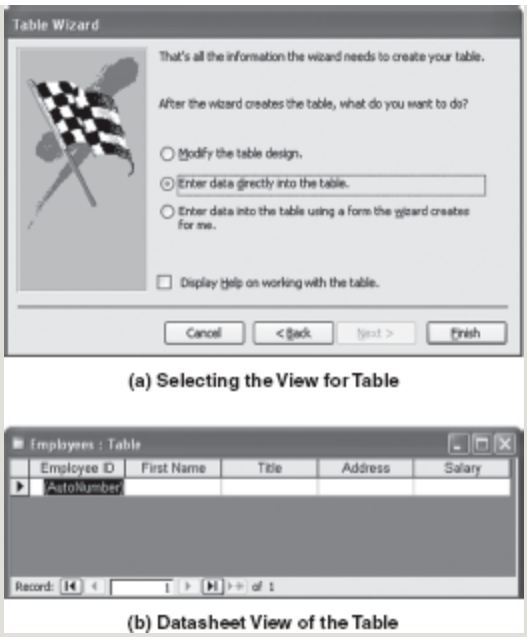


Figure 22.18 Displaying Table in Datasheet View

22.5.3 Adding, Modifying, and Deleting Records

Once the table is created, you may want to add some records to it. To add (or modify or delete) records, you need to open the table. To open a table, double-click the table name or select the table from table list and then click the **Open** button (Open) on the Database window. Note that by default a table opens in Datasheet view.

Adding and Modifying Records: Records can be added to the table in Datasheet view by typing in the record beside the asterisk (*) that marks the new record. To enter data in a table, follow the steps given below:

1. Open the table in the Datasheet view.
2. Enter the necessary information to the first field and press **Tab** to move to the next field. Similarly, add data to all the fields. Once the data for all the fields in a row are entered, press **Enter** to move to the next record. Note that Access automatically saves the new record (Figure 22.19).

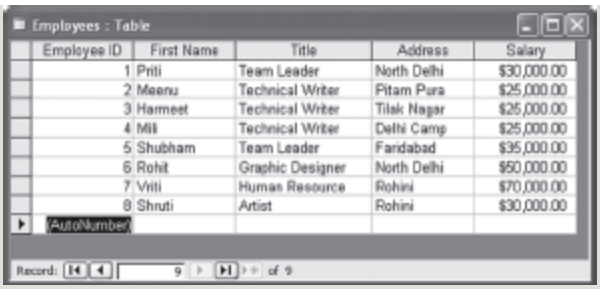


Figure 22.19 Datasheet View of the Table after inserting the records

3. To modify data in a record, place the cursor in appropriate field in the record and make the necessary changes.

Deleting Records: To delete record(s) from a table, follow the steps given below:

1. Open the table in the Datasheet view.
2. Select the record(s) by clicking the record selector area to the left of the record. To select multiple records click and drag within the record selector area.
3. Press the **Delete** key or select **Delete Record** from the **Edit** menu or click the **Delete Record** button ( ) on the Table Datasheet toolbar. An Access message box confirming the deletion appears.
4. Click **Yes** to delete the record(s) or **No** to cancel the deletion.

22.5.4 Navigating a Table

When you open an access database, you see a list of tables contained in that database. In the previous section, you learnt how to add, modify, or delete records in tables. This section explains how you can move around in a table (in Datasheet view).

Use the right arrow key ( ) or press **Tab or Enter** to move between the fields in the table. To move back a field at a time, press the left arrow key ( ) or press the **Shift** and **Tab** keys at the same time. The up ( ) and down arrow ( ) keys can be used to move between records. The record where the cursor is placed displays a triangle in the grey shaded bar at the far left indicating the active row. An asterisk (*) appears at the row where

the next record will be entered. The vertical scroll bar can be used to move up and down the records, and the horizontal scroll bar is used to move left and right around the fields.

Navigation Buttons: At the bottom of the Datasheet window are the Navigation buttons. These buttons are used to move to the first, last, next, previous, or any specific record in the table (Figure 22.20).

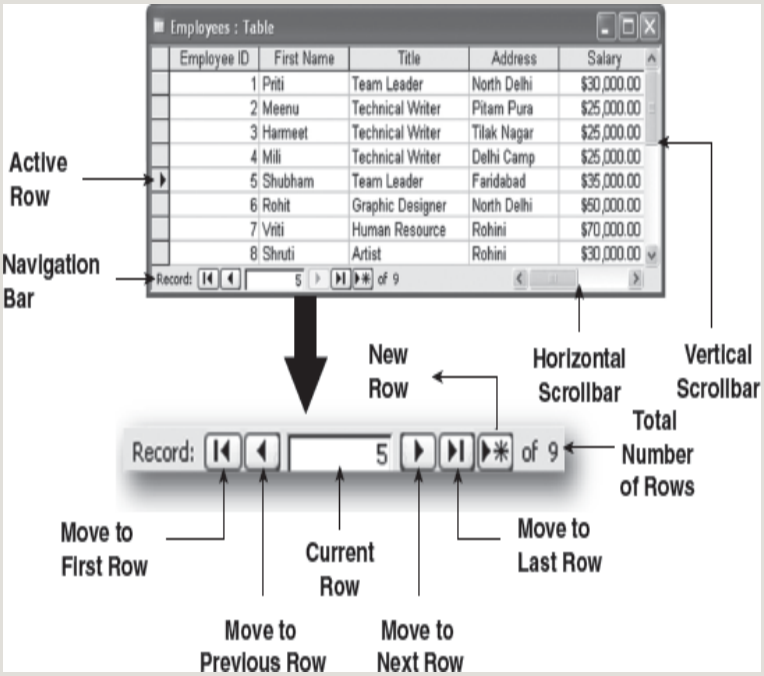


Figure 22.20 Moving around in a Table

Note: *The Datasheet view can also be used to reformat the table by changing the size and order of the field columns.*

22.6 QUERIES

The *query* is the basic tool that Access provides for retrieving information from the database. Without

queries, data would be useless, because they allow the distillation of a large quantity of data into easy-to-understand statements. Queries select records from one or more tables in a database so that they can be viewed, analysed, and sorted on a common datasheet. A datasheet is a temporary table, which the Access creates, to answer queries. It contains a subset of the data in the table serving as the source of the query. This resulting collection of records is known as **dynaset** (dynamic subset). It is saved as a database object and can, therefore, be easily used in the future. Note that whenever the original table is updated, Access automatically updates the query.

22.6.1 Creating a Simple Query Using Wizard

To create a query that displays only a selected set of fields in a table, follow the steps given below:

1. Select **Queries** ( ) from the list of Objects in the Database window and click the **New** button on Database window to display the New Query dialog box (Figure 22.21).

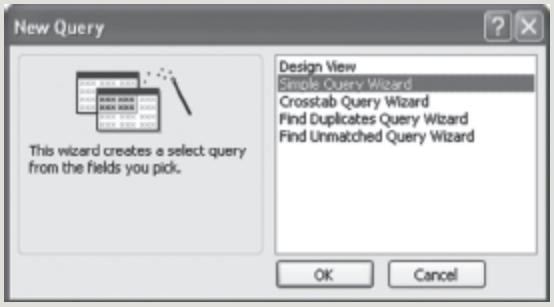


Figure 22.21 New Query Dialog Box

2. Select **Simple Query Wizard** and click **OK** to display the Simple Query Wizard dialog box.
3. Select a table from the **Tables/Queries** drop-down box. The fields related to the selected table appear in **Available Fields** list box (Figure 22.22).

Figure 22.22 Selecting Table for the Query

4. Select the desired fields for the query from the **Available Fields** list box and click **Next** (Figure 22.23).

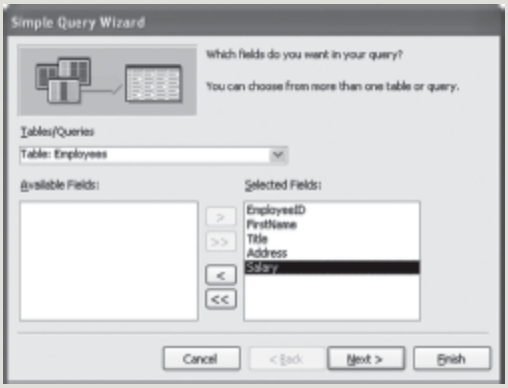


Figure 22.23 Selecting Fields for the Query

5. Based on the fields selected, the next dialog box may give you the option of choosing between a detail and a summary query. Choosing **Detail (shows every field of every record)** shows every field of every record. If you select the **Summary** option, you have additional options for how you would like the data to be summarized. For example, the wizard creates summary fields that do a count or calculate an average on the numeric field(s) (Figure 22.24).

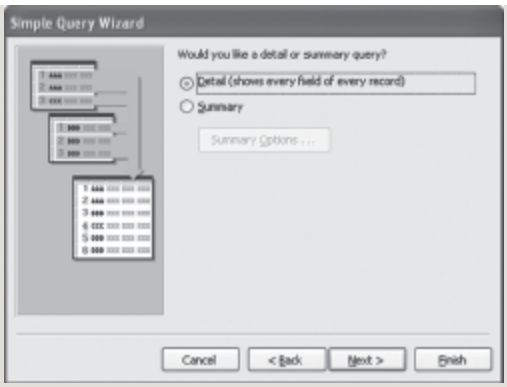


Figure 22.24 Choosing Detail or Summary Query

6. The last step of the Simple Query Wizard prompts for the query name. Enter a name for the query and then click **Finish**. Access creates and saves the query and then opens it in Datasheet view (Figure 22.25).

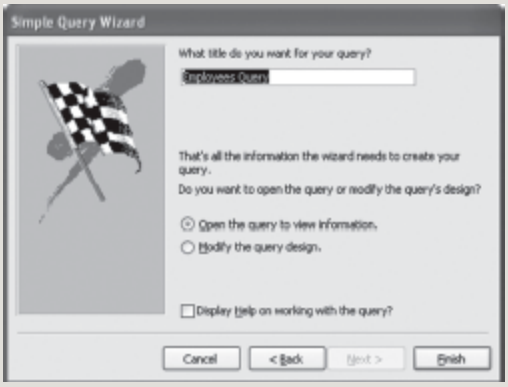


Figure 22.25 Specifying Name for the Query

22.6.2 Running a Query

Once a query is created, you can run the query to view the query result. To run a query, follow the steps given below:

1. Select **Queries** from the list of Objects in the Database window to display the list of available queries.
2. Double-click the query you want to run. Access 2003 runs the query and displays the result in Datasheet view (Figure 22.26).

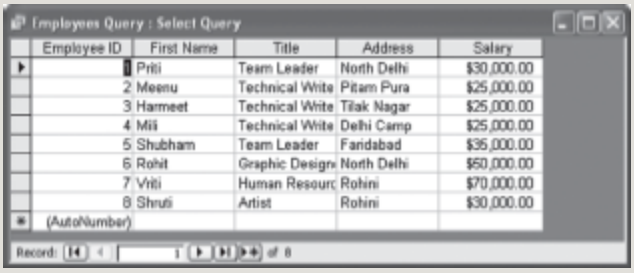


Figure 22.26 Query Result in Datasheet View

22.6.3 Sorting Query Results

Microsoft Access allows the sorting of data retrieved by a query so that it may be displayed in a logical manner, that is, either in ascending or in descending order of given field(s). To sort a query result, follow the steps given below:

1. Select **Queries** from the list of Objects in the Database window and click the desired query name in the query list to select it.
2. Click the **Design** button (Design) on Database window to open the query in Design view (Figure 22.27).

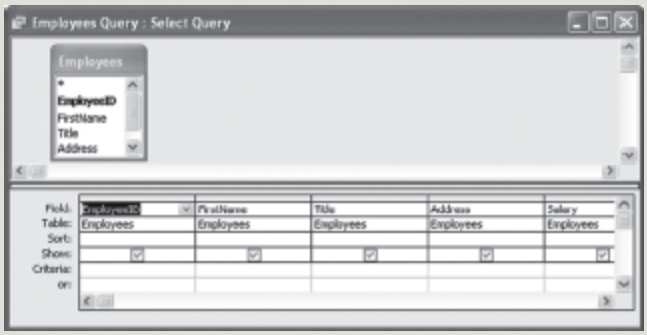


Figure 22.27 Design View of Query

3. Click in the **Sort** row of the field that you want to use for performing the sort and click the drop-down arrow to display the choices for sort order (Figure 22.28).

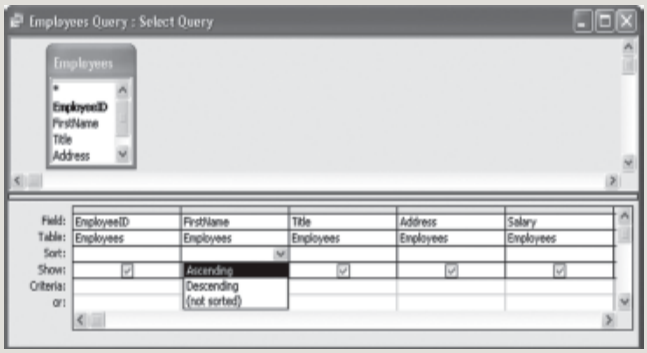


Figure 22.28 Drop-down List Showing Options for Performing the Sort

4. Select the desired sort order, for example, select **Ascending** to sort data in ascending order or select **Descending** to sort data in descending order.
5. Click the Save button ( ) on the Query Design toolbar to save the query and then click the Run button ( ) to view the result of the query in the Datasheet view with desired sort order (Figure 22.29).

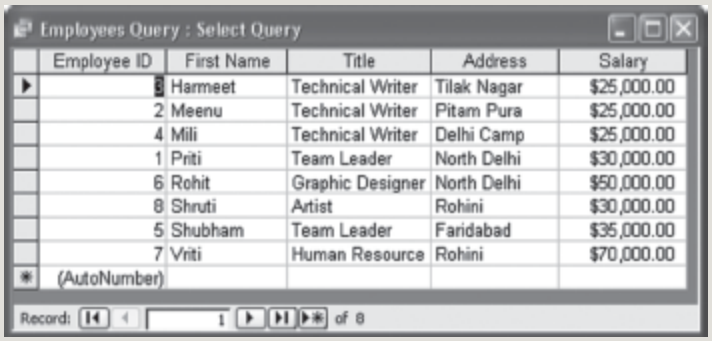


Figure 22.29 Query Result with the Desired Sort Order

22.6.4 Using SQL in Microsoft Access

Apart from generating queries using Wizards or Query Designer, Access also facilitates generation of queries using structured query language (SQL). SQL is a

language often used in querying, updating, and managing relational databases. Each query created in Microsoft Access (created using a wizard or designed in Design view) has an associated SQL statement that defines the actions of that query.

Viewing and Modifying Existing Queries: To view or modify an existing query in SQL View, follow the steps given below:

1. Select the desired query name in the query list and click the **Design** button on the Database window to open the query in Design view (Figure 22.30).

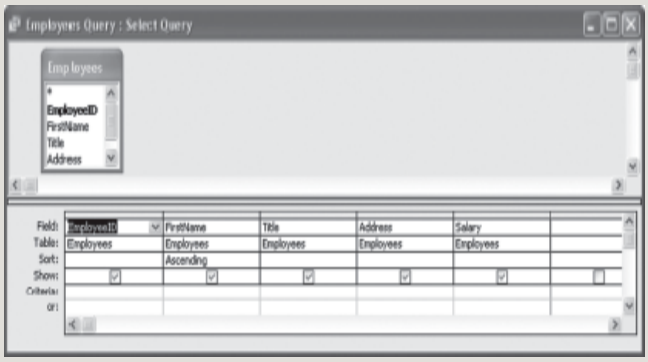


Figure 22.30 Design View of the Employees Query

2. Select **SQL View** from the **View** menu to display the query in Select Query window. This window displays SQL statement that is equivalent to the query that you created in Design view (Figure 22.31).

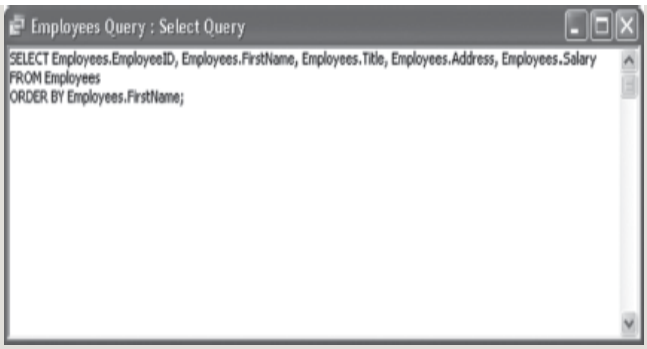


Figure 22.31 Employees Query in Select Query Window

3. If you want to modify or change the query, Access allows you to make the required changes to the SQL statement.
4. Select **Design View** from the **View** menu to return to the Design view.

Creating SQL Queries: To create an SQL query, follow the steps given below:

1. Double-click **Create query in Design view** in the Database window to display the Show Table dialog box with **Tables** as the active tab (Figure 22.32).

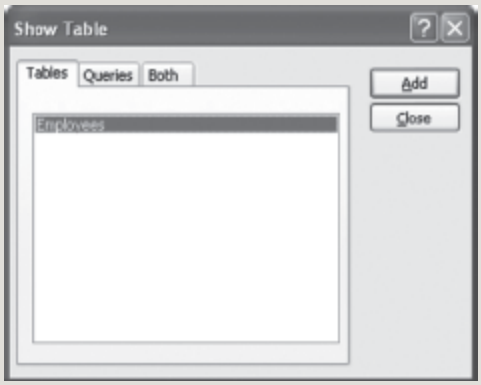


Figure 22.32 Show Table Dialog Box

2. Select the table(s) on which you want to base the SQL query from the table list.
3. Click the **Add** button to add the table(s) in the Design view.
4. Click the **Close** button to close the Show Table dialog box.

5. Select **SQL View** from the **View** menu to display the query in Select Query window (Figure 22.33).

Figure 22.33 Query in Select Query Window

6. Create the desired query and click the **Save** button ( ) on the Query Design toolbar to display the Save As dialog box (Figure 22.34).

Figure 22.34 Creating the Query in Select Query Window

7. Enter the desired name for the query in the **Query Name** text box and then click **OK** to create the query (Figure 22.35).

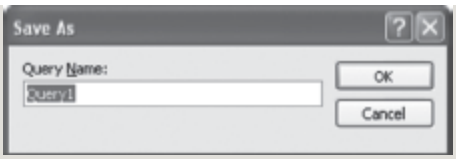


Figure 22.35 Specifying Name for the Query

22.7 FORMS

A *form* is nothing more than a graphical representation of a table. One can add, update, and delete records in the table by using a form. A form is more useful when the table contains numerous fields. This way the user of the form can view all the fields in one screen. If the same situation was to be handled using the Datasheet view, the user would have to keep scrolling to get the desired field.

22.7.1 Creating a Form Using Wizard

The easiest way to build a form is by using the Form Wizard. To create a form using the Form Wizard, follow the steps given below:

1. Select **Forms** from the list of Objects in the Database window and double-click **Create form by using wizard** to display the Form Wizard dialog box.
2. Select the table or query from the **Tables/Queries** drop-down list. The fields related to the selected table (or query) appear in the **Available Fields** list box (Figure 22.36).

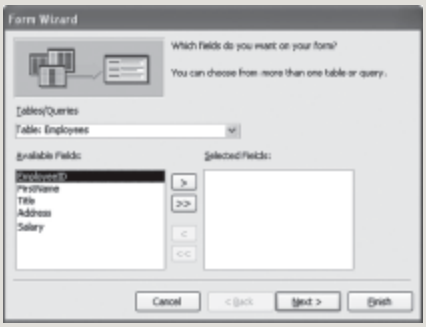


Figure 22.36 Selecting Table or Query

3. Select the desired fields for the form from the **Available Fields** list box and click **Next**. (Note that this displays the fields selected for the form in the **Selected Fields** list box.) (Figure 22.37)

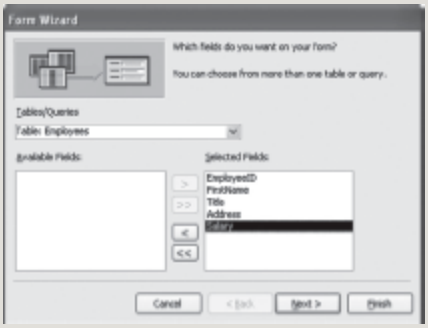


Figure 22.37 Selecting Fields for the Form

4. Select a layout for the form from the available layouts. Various available layouts are listed below:
    1. **Columnar:** A single record is displayed at one time with labels and form fields listed side-by-side in columns.
    2. **Tabular:** Multiple records are listed on the page at a time with fields in columns and records in rows.
    3. **Datasheet:** Multiple records are displayed in Datasheet view.
    4. **Justified:** A single record is displayed with labels and form fields are listed across the screen.
    5. **PivotTable:** Multiple records are displayed as an “interactive table” that allows to summarize a large amount of data quickly.
    6. **PivotChart:** The graphical representation of PivotTable is a PivotChart.

    After selecting the desired layout, click **Next** to proceed (Figure 22.38).

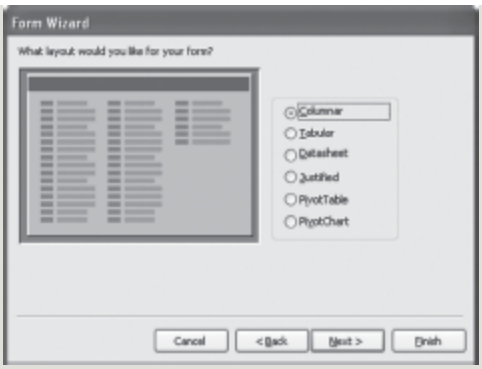


Figure 22.38 Selecting the Form'Layout

5. Select a style for the form from the various pre-defined set of styles provided by Access and click **Next** (Figure 22.39).

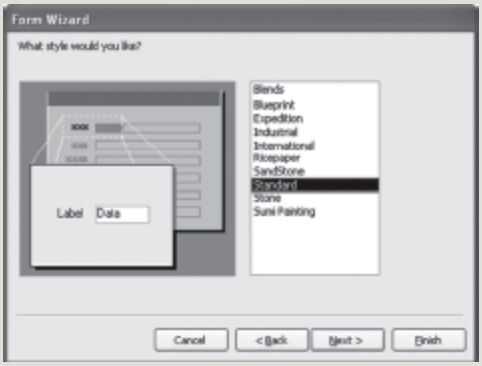
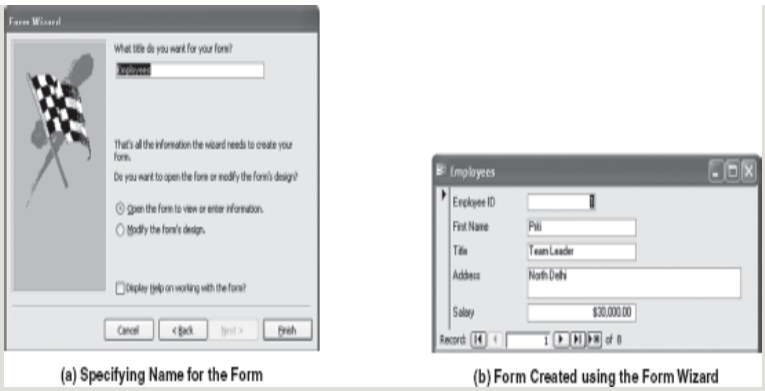


Figure 22.39 Selecting the Visual Style

6. In the last dialog box, enter a name for the form. Select **Open the form to view or enter information** to open the form in Form view or **Modify the form's design** to open it in Design view. Click **Finish** to create the form (Figure 22.40).



(a) Specifying Name for the Form

(b) Form Created using the Form Wizard

Figure 22.40 Displaying the Form

22.8 REPORTS

Although tables, queries, and forms can be printed, they may not look very professional, as they are not designed for printing purposes. Reports, however, are designed to take the data from a table, or selected by a query, and printed in a professional-looking layout. *Reports* allow you to group, sort, and manipulate the data to present them in a friendly manner for those who will be reviewing it. Reports can be made by combining data, images, charts, and even audio/video elements. You can add headers, footers, and page numbers, group the information, and change the background colours, among other things.

22.8.1 Creating a Report Using Wizard

When you design your first report, use the Report Wizard. This is the quickest way of creating a report. The Report Wizard asks you a series of questions to help you design the data exactly as you want. After you have created the report, you can open it in Design view to modify its structure. To create a report using the Report Wizard, follow the steps given below:

1. Select **Reports** from the list of Objects in the Database window and double-click **Create report by using wizard** to display the Report Wizard dialog box.

2. Select a table or query from the **Tables/Queries** dropdown list. The fields related to the selected table (or query) appear in the **Available Fields** list box (Figure 22.41).

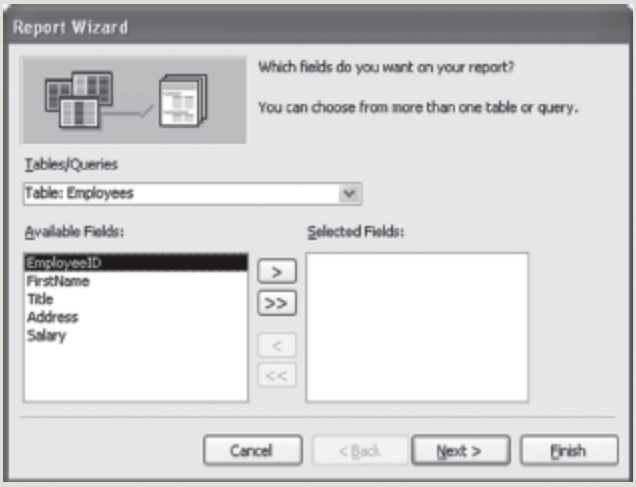


Figure 22.41 Selecting Table or Query

3. Select the desired fields to be included in the report and click **Next** (Figure 22.42).

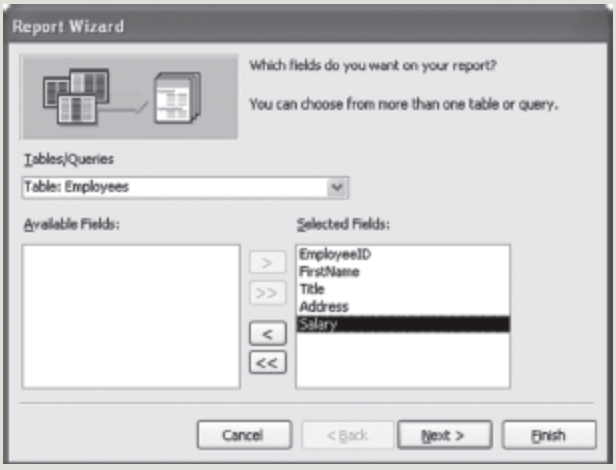


Figure 22.42 Selecting Fields for the Report

4. Based on the fields selected, the next dialog box may give option to add grouping levels for the report. Select fields from the list to

group by fields and click the **Priority** button ( ) to change the order of the grouped fields (if more than one field is selected) and click **Next** (Figure 22.43).

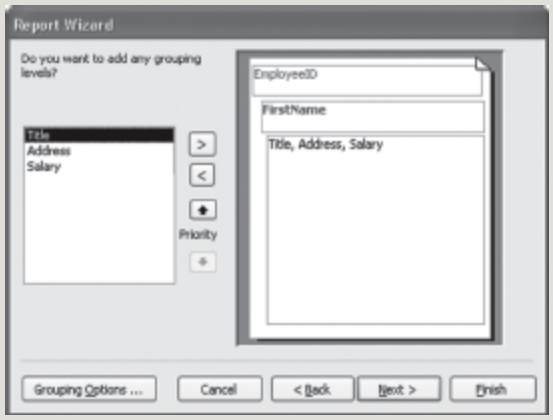


Figure 22.43 Grouping Data for the Report

5. The records in the report can be viewed in a sorted order. In the next dialog box, specify the fields on which you want to base the sorting and select the desired sort order and click **Next** (Figure 22.44).

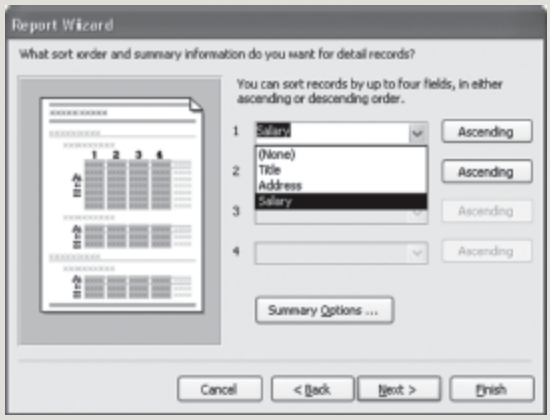


Figure 22.44 Specifying Sort Order

6. Select a **Layout** and an **Orientation** for the report and click **Next** (Figure 22.45).

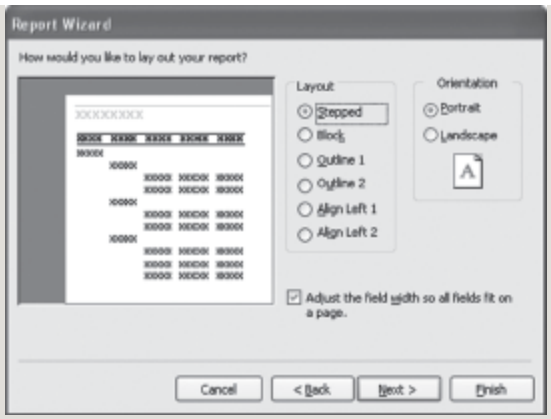


Figure 22.45 Selecting Report's Layout and Orientation

7. Select a graphics style for the report and click **Next** (Figure 22.46).

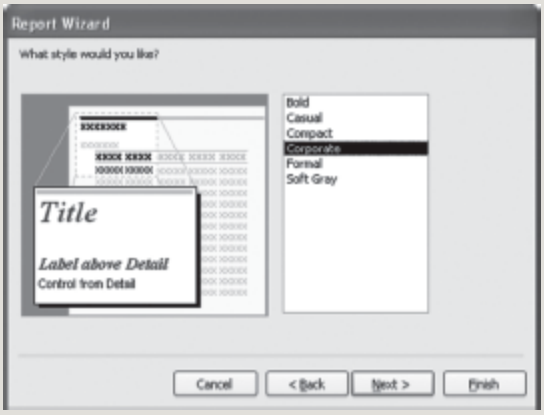


Figure 22.46 Selecting a Style for Report

8. Enter a name for the report in **What title do you want for your report?** text box. In addition, you can also select to open the report in either print preview or design view mode. Click the **Finish** button to create the report (Figure 22.47).

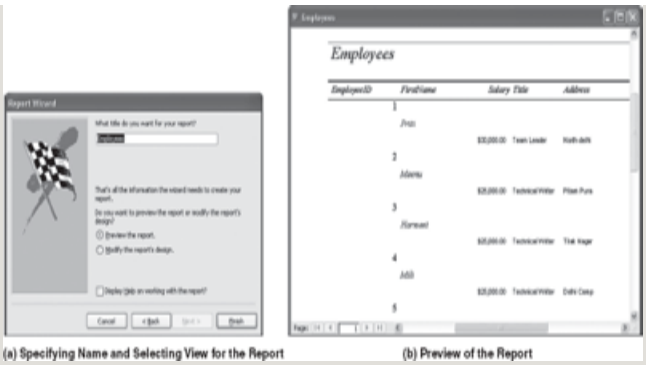


Figure 22.47 Displaying the Report

LET US SUMMARIZE

1. Microsoft Access is a database management system that allows the user to create a relational database and then access, retrieve, and use the data in it.
2. An Access database consists of a collection of tables, queries, forms, and reports that are used to manage and present data. These components are also called database objects.
3. *Table* in Access holds a collection of information stored in the form of fields and records. All the information necessary for the queries, forms, and reports in the database are based on the data in the table.
4. *Query* facilitates retrieval of information from multiple tables, which can be put together for viewing and analysis purposes. One can select records from different tables, sort the results, summarize, and perform calculations using query.
5. Access allows for the sorting of data retrieved by a query so that it may be displayed in ascending or descending order of the given field(s).
6. Access facilitates generation of queries using SQL apart from generating queries using Wizards or Query Designer.
7. *Forms* are the front-end side of an Access database that allows entering of data. It is an object in Access, which allows the database users to edit and view information, one record at a time.
8. *Reports* are the polished form data representation. When information from the tables and queries are ready to be displayed, reports help in presenting this information in a formatted manner.

EXERCISES

Fill in the Blanks

1. Access provides the feature called ..... that helps in generating the tables, queries, and reports automatically.
2. Access is a part of Office suite of applications and is developed by .....

3. ..... data type is used to automatically insert numeric values in a field of a table.
4. The extension of an Access database is .....
5. ..... present at the bottom of a table are used to navigate from one record to another.
6. Specific information can be retrieved from a database using .....
7. Queries can be generated by Query Wizard, Design View, and through .....
8. Forms are ..... representation of data.
9. Reports look very professional, as they are designed for ..... purposes.
10. The Form Wizard allows ..... types of layout.

Multiple-choice Questions

1. Which one of the following is not an object of Access?
    1. Tables
    2. Queries
    3. Layouts
    4. Reports
2. Which of the following statement(s) is true about Access tables?
    1. A database may contain more than one table.
    2. There can be more than one primary key in an Access table.
    3. A database contains only one table.
    4. Access table cannot store images.
3. Which of the following data types can hold up to 65,536 characters?
    1. Text
    2. Currency
    3. Number
    4. Memo
4. In the Form Wizard, which of the following is the default layout of a form?
    1. Datasheet
    2. Columnar
    3. Tabular
    4. Justified
5. Reports can contain .....
    1. Charts
    2. Audio
    3. Images
    4. All of these
6. A report may be based on .....
    1. Table but not query
    2. Neither a table nor a query
    3. Query but not table
    4. Both table and query
7. The Show Table dialog box that appears while creating queries in Design view shows .....
    1. Tables
    2. Queries
    3. Both (a) and(b)
    4. None of these
8. Which of the following operations can be performed in a form?

    1. Adding a record
    2. Modifying a record
    3. Deleting a record
    4. All of these
9. Which of the following view is not supported by Access?
    1. Datasheet View
    2. Slide View
    3. Design View
    4. None of these
10. Which of the following is used to answer a question about a database?
    1. Query
    2. Table
    3. Report
    4. Form

State True or False

1. There are no in-built databases in Access.
2. For creating a query, at a time only one table can be used.
3. It is essential for a user to set the field properties while designing a table.
4. While designing a query, the Sort row is used to sort the fields in ascending or descending order.
5. A user can create queries using SQL commands.
6. While creating a table using the wizard, the user cannot rename a field.
7. It is essential to define a primary key in a table.
8. You can add priority to the fields in a report.
9. In Access, macros and modules are not the objects like tables and queries.
10. A table can be displayed either in a Design view or in a Datasheet view.

Descriptive Questions

1. Write a brief note on various Microsoft Access 2003 objects.
2. Create a query using the Query Wizard. Now write down the steps that must be followed to sort the records based on a key in descending order.
3. Create a report named “Books” that show the books issued according to date and serial number of the Book ID.
4. Design and create a database for storing a Birthday List of your family and friends. Include fields such as name, birthday, favourite colour, and hobby. Insert at least six records in the database.

ANSWERS

Fill in the Blanks

1. Wizards
2. Microsoft
3. AutoNumber
4. .mdb
5. Navigation buttons

6. Queries
7. SQL
8. Graphical
9. Printing
10. Six

Multiple-choice Questions

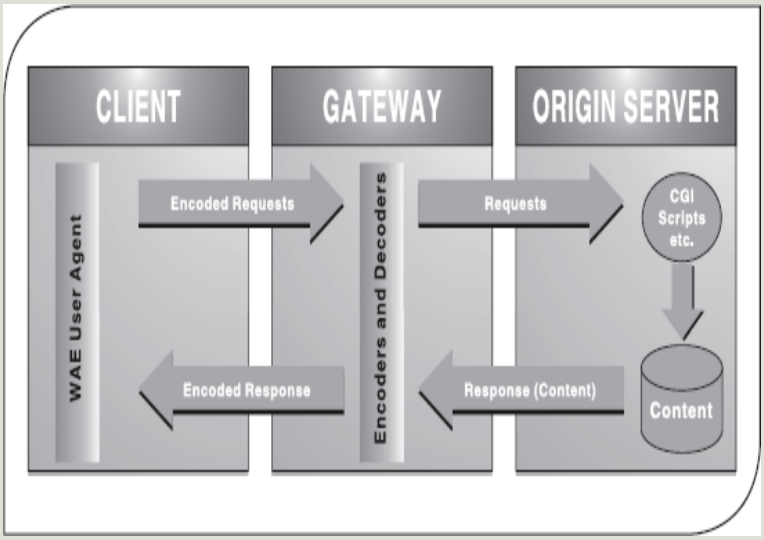
1. (c)
2. (a)
3. (d)
4. (b)
5. (d)
6. (d)
7. (c)
8. (d)
9. (b)
10. (a)

State True or False

1. False
2. False
3. False
4. True
5. True
6. False
7. False
8. True
9. False
10. True

Chapter 23

Current and Future Trends in IT



This chapter is penned to provide a foundation to harness the power of information technology in areas such as e-commerce, wireless applications and communication. The chapter begins with a discussion on e-commerce. Next, electronic data interchange and wireless application protocol (WAP) are studied. The later part of the chapter explains new technologies such as smart card, IPTV, RFID and BCI. The chapter concludes by providing a discussion on the technologies that will in future bring economies of scale to humankind.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- E-commerce, which is a way of carrying business communications and transactions using computers over Internet, or any other form of network
- EDI—a way of exchanging business data between two machines
- Wireless application protocol (WAP), which is the worldwide standard protocol stack specified for wireless communication networks by the WAP Forum
- Smart card—a small, portable pocket-size card having data storage ability
- IPTV, which is a system that provides the television services to the end users via packet switched network
- Blogging, a process of creating, maintaining, writing, or adding content to a blog
- RFID—an auto-identification technology, which is used to capture and transmit identification details (including location) of an object wirelessly
- Brain computer interface (BCI), which is a communication pathway between brain (of either human or animal) and some external device such as a computer system
- Imminent technologies including the concept of Pen PC, Napkin PC, Roll top computer, B-membrane PC, and USB memory stick

23.1 INTRODUCTION

In the early years, computer use was primarily confined to business and research. However, the 1980s and 1990s saw the development of cheaper and more versatile computers, which has led to an ever-growing number of users in the offices and in homes. Today, IT has interwoven itself into our socioeconomic fabric in such a manner that it is often said that we live in an information age. As we usher in the future, consumers are expected to

sign up in record numbers for information, education, and entertainment services provided over networked computers, or perhaps in altogether different computing environment. Technological revolutions have always had different effects on different segments of the economy. The trading activities have become much easier with the use of e-commerce and EDI. The development of services like WAP has enabled people to access global information anywhere and at anytime. Some new trends like Internet protocol television (IPTV) and radio frequency identification (RFID) are directed to make the people's lives more convenient. However, these technologies are not in full use yet. Besides, future has some more interesting PC concepts that if ever developed will not only add to the ease of computer usage but will also enhance its portability.

23.2 ELECTRONIC COMMERCE (E-COMMERCE)

Traditionally, commerce is seen as the exchange or buying and selling of goods and services, which involves exchange of money, and sometimes transportation of goods from one place to another. We see commercial activities all around us in many forms. For instance, at a grocery store, we can see two main activities that form the essence of commerce, the buying (people purchasing goods or service) and selling (people offering goods and services to buyers). With the passage of time and advancement in technology, commerce has added a new dimension in the form of electronic trade, a more effective and convenient manner than the traditional commerce which is revolutionizing the way business activities are carried out across the world. In today's business world, where time and money are considered as the most important factor for success, E-commerce plays an important role in bridging the gap between buyers and sellers (Figure 23.1).

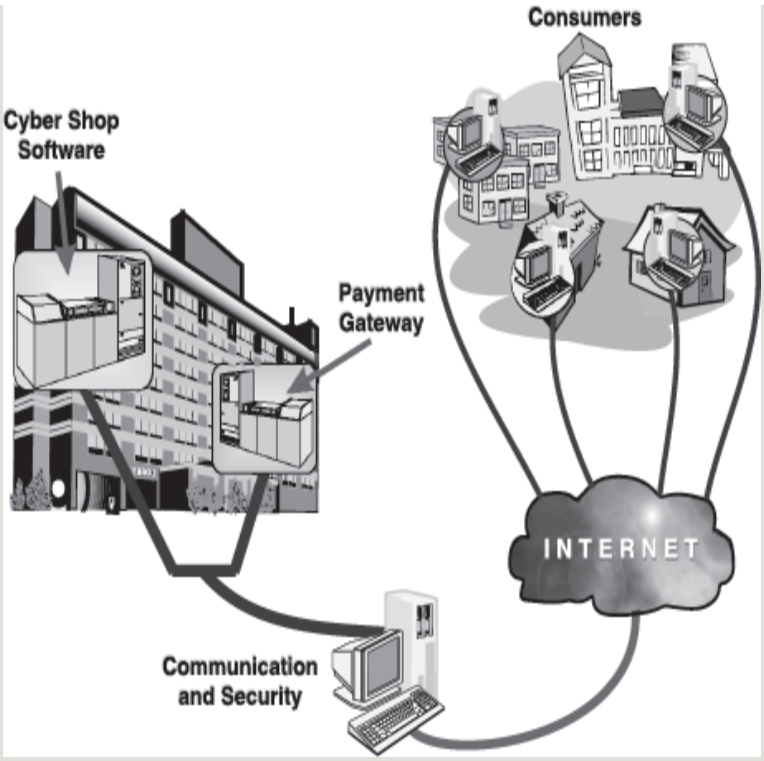


Figure 23.1 E-Commerce and Business Processes

Most people think E-commerce as buying and selling of products and services on the Internet. However, in reality, it is more than just selling and purchasing goods online. E-commerce is referred to as a way of carrying business communications and transactions using computers over Internet, or any other form of network. Here, transaction is referred to as exchanges taking place between two entities where one sells products or services to the other.

23.2.1 Types of E-Commerce

There exist four categories of e-commerce applications: business-to-onsumer (B2C), business-to-consumer (B2C), business to business (B2B), business-to-administration (B2A), and customer-to-customer (C2C).

Business-to-Consumer (B2C): B2C is a concept of online marketing and distribution of products and services over the Internet. It is a natural sequence of steps for many retailers who sell goods directly to the consumer. The general idea for achieving success in B2C e-commerce lies in reaching more customers, providing better services and increasing sales while spending less. For the consumer, it is relatively easier to appreciate the importance of e-commerce as it saves their precious time, which they spend moving around in crowded supermarkets, by enabling them to do the same from their home, by shopping online at any time in virtual Internet shopping malls, and having the goods delivered at their doorstep.

The supplier may be an existing retail outlet such as a grocery store; these businesses may have been slow in gearing-up for e-commerce compared to the innovative Internet companies, but they usally have a sound commercial structure as well as in-depth experience of running a business. This system can also be used to receive information such as share prices, insurance quotes, online newspapers, or weather forecasts.

Business-to-Business (B2B): B2B system is used for e-commerce in a pure business environment. It covers many ways where e-commerce can be used to support one business trading with the other. Much of the attention given to B2B system relates to developments such as e-marketplaces and e-exchanges, but most of the actual use of e-commerce is at a lower level in the e-commerce implementation cycle.

Typically, in the B2B environment, e-commerce can be used in procurement, order fulfilment, and managing trading–partner relationships. In recent times, e-commerce technologies have allowed the smallest of businesses to improve the processes of interacting with

customers. They are now developing services for individual clients rather than providing a standard service and evaluating all aspects of their business process from where the greatest return can be obtained. B2B e-commerce can also be used to:

- Develop, retain, and cultivate relationships with customers.
- Streamline the supply chain, manufacturing, procurement processes, and automate corporate processes to deliver the right products and services to customers quickly and cost-effectively.
- Capture, analyse, and share information about customers and company operations in order to make better decisions.

Business-to-Administration (B2A): In the B2A system, business community interacts electronically with public sector organizations. B2A application includes submission of planning applications, tax returns, or patent registration.

Customer-to-Customer (C2C): C2C system refers to e-commerce activities, which uses an auction style model. This model consists of person-to-person transaction that completely excludes businesses from the equation. C2C system enables customers to directly deal with each other. Examples of this type of system are *www.ebay.com* and *www.amazon.com*.

23.2.2 Processes in E-Commerce

All information in e-commerce is expressed in the form of binary values (bits), which makes the product more versatile. For example, a tourism catalogue data stored in a web server database can be presented electronically using the web, printed in a customized catalogue at specific market, or might be included in CD-ROM as a multimedia presentation. If any information is required to be published on the Internet, the hardware and software component required are web server, database, and multimedia authoring tools. Generally, there exist five important processes in an e-commerce lifecycle:

information sharing, ordering, online payment, fulfilment, and support and services.

Information Sharing: Sharing of information on the Internet is the foremost process in e-commerce activities. It makes the customer aware of all the product and services offered by the online vendor. This particular activity is carried out by means of advertising, marketing on the Internet, and customer information gathering process (see Figure 23.2). Apart from this, there are various networked communities (chat room, multi-party conference, and newsgroup) available on the Internet for the distribution of information.

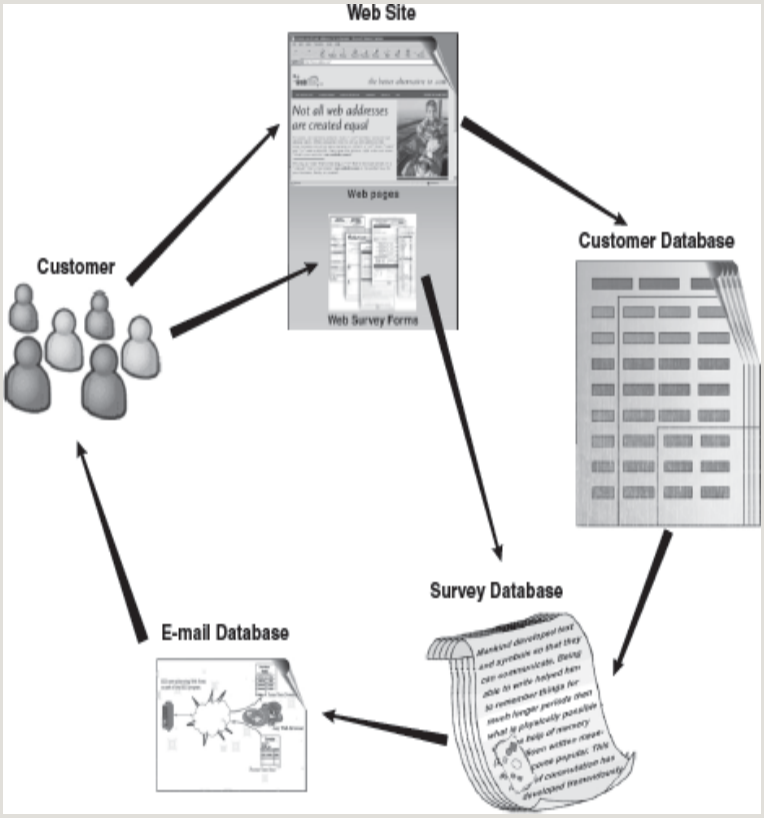


Figure 23.2 Information Gathering from Customer

World Wide Web provides one of the effective mediums for communication with the customers. Websites are designed to include product catalogues. One can obtain the data for products that are requested. At the same time, it also allows web visitors to provide information about them. Apart from this, timely notification, product upgradation, and new features are sent through e-mail.

On Internet, numerous support services are provided to the customers. For example, organization maintaining a web server can include questionnaire, which are available to the customer while using company's website. Repeated questions are often compiled into one set and distributed to the customer's e-mail. Large number of websites aim at providing information and data to the general public; however, there exist some significant sites on the Internet that cater to the business communities. These sites are called intermediaries, and their sole purpose lies in making buyers to interact, trade, and make purchases from the sellers.

Ordering: This process involves customers to electronically place order for the company's goods and services. Placing of order requires the use of client/server application where order of goods or services is placed on the electronic form placed on the company's website.

Online Payment: Online payment is a crucial process in e-commerce. The payment activity requires the use of high security to protect online money transactions from any unauthorized use. In most cases, data are transferred in encrypted form on a secured medium. Different mechanisms for payment are available on the Internet and the customers can use any of them.

Fulfilment: Modern businesses are highly dependent on the massive amount of information and precious data,

which is continuously being transferred over the computer network across the world. Many online trading companies make use of Internet for transferring funds, generating money, and providing information product (newsletter, articles, stock prices, and reports). However, when product takes the shape of physical entity, support of EDI is required to clear the sale of product to the buyer with up-to-date delivery notification issued to the buyer's e-mail.

Support and Services: Support and services are regarded as one of the most important processes in e-commerce activity. The continuing relationship of an organization with the customer depends on the kind of support and services provided to the customers. Relationship of an organization with the customer starts with the sale of the product. It might happen that customer will require assistance with product and services in future. Therefore, items such as technical notes, FAQs, and support documentation should be available on the company's website.

23.2.3 Types of Electronic Payment

The method of making payment on the Internet is essentially the electronic version of the traditional payment system, used in everyday life. The fundamental difference between the electronic payment system and the traditional one is that every transaction is carried in the form of digital string. Generally, there exist three major types of payment mode: *e-cash, e-cheque,* and *credit card*.

E-cash: E-cash is one of the systems that fit small-scale commercial transactions in realtime environment on the Internet. In e-cash, the currency is transformed into an electronic form. A bank issues electronic money and debits the customer's account by withdrawing the specified amount from customer account, which equals

the value of currency (token) issued to the customer. The bank then authenticates each token with the digital stamp before transmitting it to the customer computer. Now whenever the customer wants to spend e-cash, a token of roper amount to the merchant is transmitted who then relays them to the bank for verification and redemption. To ensure that no fraudulent activity occurs, bank records the serial number of each token that the customer spends. If the serial number of the token was already recorded in the bank's database, the bank would inform the merchant for its invalidity.

E-cheque: The payment mode for e-cheque is not much different from the paper cheque. A paper cheque is a kind of message given to the consumer's bank to transfer adequate fund from one customer's account to someone else's account. The message is sent to the intended receiver of the fund who presents the cheque to the bank in order to collect the money. In all aspects, an e-cheque also has the same features as the paper cheque. It sends a message to the receiver who endorses the cheque and presents it in his bank to obtain money. One advantage that e-cheque has over paper cheque is that it provides protection to the customer's account by encoding the account number with the bank's public key. This prevents the merchant from knowing the customer's account, thereby eliminating the chances of fraudulent transactions.

Credit Card: In credit card payment, customer gives a preliminary proof of the ability to pay by giving the credit card number to the vendor. The vendor then verifies it from the bank and issues a purchase slip for the consumer to endorse. The vendor then uses a copy of the purchase slip to collect money from the bank and in the next billing cycle, the consumer receives a statement of recorded transaction at the address specified by the customer. Using credit card to make purchases over the

Internet requires additional steps to make transaction secure for both the buyer and seller. Credit card information is sent as encrypted credit card number over the Internet.

23.2.4 Advantages of E-Commerce

Some of the advantages of e-commerce are as follows:

- **24/7 Access**: Online businesses never sleep as opposed to conventional businesses. Consumers can access goods and services anytime they want.
- **Reduced Prices:** The cost of product is reduced since stages along the value chain are decreased. For example, intermediaries can be eliminated by the company directly selling to the consumer instead of distributing through a retail store.
- **Global Marketplace**: Consumers can shop anywhere in the world. This also provides a wide selection of products and services to consumers.
- **Increased Potential Market Share:** Internet enables businesses to have access to international markets, thereby increasing their market share. Companies can also achieve greater scale of economies.
- **Lower Cost of Advertising**: Advertising on the Internet costs relatively less than advertising on print or television media. A company can spend a lot on advertising on the Internet if the company hires an external party to create their advertisement.
- **Low Start-Up Cost**: Anyone can start a company on the Internet. Start-up costs are lower for companies since there is less need for money for capital assets.
- **Secure Payment Systems**: Recent advancements in payment technologies allow encrypted, secure payment online.
- **Lower Cost of Practising Business**: Less number of employees, purchasing costs, order-processing costs associated with fax, telephone calls, data entry, and eliminating physical stores result in reduced transaction costs.

23.2.5 Disadvantages of E-Commerce

Every innovation in technology brings certain limitations and drawbacks along with it. The same is the case with e-commerce. The following are its disadvantages:

- **Lack of Personal Examination of Products**: Buying products through the Internet do not allow physical examination of products. Images of the products may be available for viewing; however, there is a risk involved in the uncertainty of the quality of the product that the consumer is purchasing.

- **Hardware and Software**: There are specific hardware and software that are essential to start an e-commerce company, which are bulky and costly.
- **Distribution Problems**: Distribution must be very efficient especially when catering to a global market.
- **Maintenance of Website**: E-commerce websites must be maintained and updated regularly which leads to extra labour costs.
- **Extra Costs**: Even though the company may initially save money by cutting middlemen and not having to invest much on capital assets, other costs may be incurred that include start-up costs of the company in terms of hardware, software, training of employees, costs to maintain the website, and distribution costs.
- **Training and Maintenance**: It is important to have well-skilled and trained workers to create and maintain e-commerce facilities of a company.
- **Security Risk**: An e-commerce business exposes itself to security risks and may be susceptible to destruction and disclosure of confidential data, data transfer, and transaction risks (as in online payments) or virus attacks. The possibility of stealing credit card numbers is one of the major threats in e-commerce activity.

23.3 ELECTRONIC DATA INTERCHANGE (EDI)

Previously, we had discussed that E-commerce uses digital methods to exchange the variety of data needed to conduct business. Like every other business organization, online companies also exchange data related to ordering and paying for items. In fact, e-commerce covers a broad spectrum of all types of data exchange in such paperless environment. Every day, businesses generate and process a large volume of paper documents. The paper documents, ranging from purchase orders and invoices to product catalogues and sales reports, provide vital information flow which must precede, accompany, or follow the physical goods in a commercial transaction. Any interruption in the physical or information flow halts the smooth operation of the supply chain, leading to significant cost overruns.

Today, a majority of the data in commercial paper documents are generated from computer applications. These paper documents are printed and copied before the information they contain is finally communicated to the respected parties. The business partner, in turn, inputs all this information into another computer

application for further processing. An increasing number of companies have found the above process extremely slow, costly, and unreliable. The need for a faster, cheaper, and more accurate solution for exchanging commercial data has become a significant priority for many companies and organizations. This gave rise to EDI. It is a subset of e-commerce and is a new way of managing information. At its most basic, EDI is the transmission of business data between two machines. More specifically, EDI is defined as a collection of public standard message formats and a data element dictionary that allows trading partners to exchange data in a simple way using any electronic messaging service. Trading partners are all the entities involved in an exchange such as a vendor and its customers. These standard message formats provide an application-neutral format (that is, a format independent of any application) for direct computer-to-computer exchange of information.

23.3.1 How EDI Works?

A two-way working process of EDI-based transactions between vendor and supplier has been depicted in Figure 23.3. After business agreements are reached between the two EDI-enabled partners, no human touch is required. The ultimate EDI implementation is a totally automated interaction between the business applications. A step-wise execution of processes involved in EDI transaction for purchase, shipment, and payment is as follows:

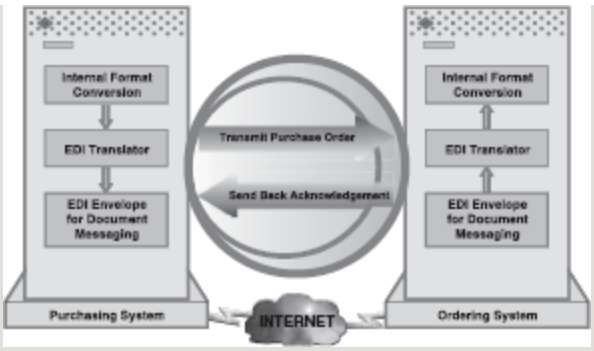


Figure 23.3 Working of EDI

- The sender's EDI translation software translates the document into a standard format and establishes the communications connection by dialling the telephone number of the network.
- The message from the sender is transferred to an electronic mailbox on the network.
- From the electronic mailbox, the trading partner's software retrieves the file, interprets the message, checks that it complies with the EDI standards, and finally, stores it. It also sends a functional acknowledgement to inform the sender if the message was received and complied with the EDI standard.
- To read and process the interpreted message, the receiver has two options. Either use the EDI translation software to produce a hard copy of the message or restructure (map) the message into a format required by the existing computer applications before further processing.

23.3.2 EDI Benefits

The goal of EDI is to provide a standard format and data element dictionary to exchange data in an electronic environment. EDI solves the problem of sharing data between dissimilar operating systems, applications, and computers using some neutral format. EDI is designed to:

- Eliminate any manual data entry into one or more software applications for analysis and reporting which dramatically reduces errors, data entry costs, and time delays.
- Improve data consistency because the format and data element dictionary provide a standard set of parameters for all contracts while accommodating unique data content.
- Eliminate proprietary software formats, interfaces, and programming effort because the standard format is application neutral.
- Provide the means for all contractors to use a standard format for all their government customers eliminating contract-specific data format

and data delivery costs.

- Provide program managers with a common set of commercial tools and processes they can rely on and which are transparent to them.

23.3.3 EDI Limitations

EDI is not a technology but a new way of doing business. Though identifying the costs related to EDI is relatively easy, identifying the benefits of EDI, especially when a company is not using EDI, is hard, since most of the real benefits are strategic and intangible in nature. As such, EDI will require a different approach to cost-benefit analysis and it is essential for the success of a corporate EDI program that management has appropriate expectations on the return on investment (ROI). Still a cost-benefit analysis is not so essential as to justify the investment in EDI in order to prioritize applications and the allocation of information technology resources.

23.4 WIRELESS APPLICATION PROTOCOL (WAP)

Hardly a decade ago, when Internet just started, it was mere a thought to access the global information from our own room. In the late 1990s as the Internet grew, this thought turned into reality. From the last few years, people are more inclined to mobility, which implies a growing need of accessing global information from anywhere and at all the time. Today, the access to Internet is no more restricted to desktop computers only; people can access Internet on their mobile phones or PDAs also. One can even connect mobile phone with a computer to access Internet on the computer. The technology that has made this wireless Internet feasible is *WAP*.

WAP is the worldwide standard protocol stack specified for wireless communication networks by the WAP Forum. It enables developing wireless applications independent of the wireless access technology being used and accessing Internet using mobile devices. WAP uses

wireless markup language (WML), which is a form of eXtensible Markup Language (XML), and provides a way of adding information to the extent that tells the device receiving the content as to what action to perform.

There were certain issues in bringing wireless Internet on the mobile devices. Most cell phones and WAP-enabled PDAs have data transfer rate of 14.4 kbps or less as compared to a typical 56 kbps modem. The relatively small size of the LCD on a cell phone or PDA presented another challenge. Most web pages were designed for a resolution of 640 × 480 pixels, which is acceptable for a desktop computer. However, the pages do not fit on a mobile device's screen, which have a resolution of around 150 × 150 pixels. WAP took each of these limitations into account and provided a way to work with a typical wireless device.

Note: *Like WWW, the user interface to WAP is through a mini browser in the WAP-enabled devices.*

The WAP Model: WAP model is based on the existing WWW model, with the significant difference of having a gateway between the client and web server. The role of gateway is to translate the WAP-formatted messages received from WAP-enabled device into HTTP messages that can be sent to any web server on the Internet and vice versa. The following steps describe the process of retrieving information from a website using WAP-enabled device.

- The device is turned on and a mini browser is opened.
- The device then sends out a radio signal, searching for the service.
- A connection is made with the service provider.
- After that, a desired URL is provided.
- A request is then sent to a gateway server using WAP.
- The gateway server retrieves information using HTTP from the website.
- The gateway server encodes the HTTP data as WML.
- The WML-encoded data is then sent to the WAP-enabled device.

- The requested web page is displayed on the device (Figure 23.4).

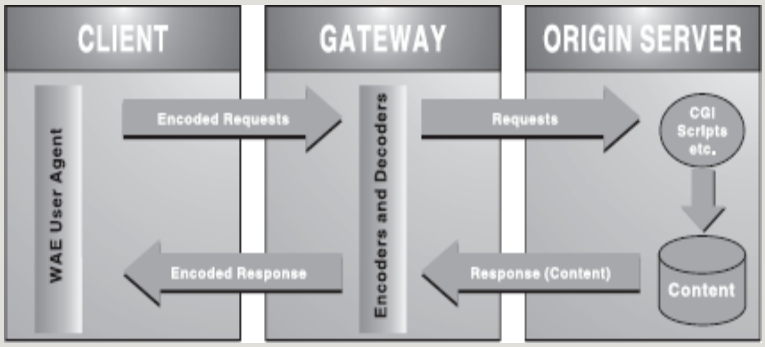


Figure 23.4 WAP Model

WAP Applications: One of the most significant advantages of mobile Internet over personal computer is the ability to instantly identify user's geographic location. Some other interesting applications of WAP are as follows:

- WAP applications are used by the children for playing games, on the cellular phones.
- Music companies offer to download the latest pop hits to customers in a daily offering.
- WAP-enabled devices are used by small-to medium-sized companies, which allows a field officer to access customer order information and stock availability details using WAP handset.
- Many cellular companies use WAP applications in customer care and provisioning, message notification and call management, e-mail, mapping and location services, weather and traffic alerts, sports, and financial services, address book and directory services, and corporate intranet applications.

Earlier, the WAP services were limited to simple information services, but with the availability of higher bandwidth and speeds, several new service types have emerged. Some of these services are:

- **Infotainment** includes weather forecasts, stock quotes, horoscopes, and news.
- **Messaging** includes e-mail, voicemail, and unified messaging.

- **Personal information management** includes call management and personal directories, which enable the modification of personal information.
- **Financial services** include mobile banking and mobile e-commerce services.
- **Location-based services** that are dependent on location include mapping and vehicle location information.

23.5 SMART CARD

Smart cards are small, portable pocket-size cards having data storage ability. Some people might confuse smart cards with conventional credit/debit cards that have a plastic base with magnetic stripe containing some information. But smart cards are quite different from magnetic stripe cards. Most smart cards incorporate an Integrated Circuit Chip (ICC) inside them having some computational functionality and I/O support. They come in various sizes ranging from as small as mobile phone prepaid cards to as big as credit/debit cards.

According to the applications for which smart cards are used, they are classified as *microprocessor cards and memory cards*. A microprocessor card is just like a complete computer in itself. It is embedded with a microprocessor and memory (RAM, ROM, and EEPROM) on the chip. The microprocessor helps in processing of information stored on the card and this processing is carried out in RAM. The data on the card are stored in the form of different files that are arranged in a file system in EEPROM. In ROM, the operating system resides that is responsible for managing the file system in EEPROM, executing the required functions in RAM, and for security and access control to data. The inclusion of operating system makes these cards multifunctional. The data on microprocessor cards can be added, deleted, or manipulated. Examples of these cards are stored-value cards (cards that hold some monetary value) and the cards that allow set-top boxes on televisions to remain secure from piracy. A memory card, on the other hand, is embedded with memory

(EEPROM) only. Thus, data can only be read from such cards but cannot be added or manipulated. The memory cards are designed to undertake pre-defined operations. The prepaid mobile phone cards are the most widely used memory cards.

Note: *The present-day microprocessor cards are embedded with 32-bit RISC processor.*

Depending on the way the smart cards are used, they are categorized as *contact* and *contactless* cards. The contact smart cards have the same size as that of conventional credit/ debit cards. These cards have contact pads attached to the chip that needs to be in the direct contact with the smart card reader whenever these cards are to be accessed. On the other hand, the contactless smart cards need not come in direct contact with the reader. They have an embedded antenna attached to the chip that communicates with the reader to read or write data on the card. Some contactless smart cards can be accessed accurately up to 1.5 m away from the reader while some need to be only few millimetres away.

Smart Card Applications: Smart cards have greatly improved the convenience and security of transaction. They provide tamper-proof storage of user and account identity. They also provide vital components of system security for the exchange of data throughout any type of network. They protect against a full range of security threats, from careless storage of user passwords to sophisticated system hacks. Multifunction cards can also serve as network system access and store value and other data. People worldwide are using smart cards for a wide variety of daily tasks, which include:

- **Personal Finance**: As banks enter in competition in newly opened markets, such as investment brokerages, smart card is used for making secure transactions at an increased rate. This improves the customer satisfaction as customers can now use it for electronic transfer of fund over the Internet.

- **Health Care**: In health care institutions, the smart cards are used to store information about patients' history and their verification details. The health professionals can instantly access this information whenever required as well as update the information.
- **Mobile Communications**: Smart cards are used to identify the GSM digital mobile phones. All the information that is required for authenticating and billing the user is stored on the card so that the user can use any of the phone terminals.
- **Electronic Purse**: A smart card can be used as a stored-value card that holds some monetary value. While making payments, the smart card reader can access the current amount stored on the card and subtract from it the amount for purchases made. Thus, one need not carry cash with himself/herself
- **Campus Identification and Access**: Businesses need simple identity cards for all its employees and similarly, universities for its employees as well as for its students. Most of these people are also granted access to certain data, equipments, and departments according to their status. Microprocessor-based smart cards incorporate identity with access privileges and store value for use in various locations, such as cafeterias and stores.

23.6 INTERNET PROOTCOL TELEVISION (IPTV)

IPTV is a system that provides the television services to the end users via packet switched network. This network uses Internet protocol suite to transmit audio, video, and control signals, thereby providing quality of service, security, and reliability.

IPTV sends data to the end user(s) in two ways, namely, *unicast and multicast* (or *broadcast*). In unicasting, data are sent to a single end user. Video on demand (VoD) is the best example of unicast in which only the subscriber who has requested for the video can view the transmitted video. On the other hand, in multicasting (or broadcasting), the same data are sent to multiple users simultaneously. For example, in Live TV, a particular video is delivered to many viewers (Figure 23.5).

Figure 23.5 IPTV

The core of any IPTV system is the *operator's central distribution centre* where the video stream is encoded and split into packets, which are transmitted over the IP network. The packets travelling through the network are received by the *local exchanges*—the IPTV middleware that routes the packets to the individual subscribers. The local exchanges are responsible for dealing with user authentication, VoD requests, billing, or channel change requests from the subscribers. In addition, here, some local content like advertisements, VoD, etc., can also be added. From the local exchanges, the packets are sent to the subscribers via local network. At the subscribers' end, there is a device known as *set-top box*, which receives the incoming data, reassembles the packets into coherent video stream, and decodes it to be viewed on the television or some other form of screen.

Note: *To use the bandwidth efficiently, the IPTV video is compressed using compression techniques, MPEG-2 and MPEG-4.*

IPTV offers many advantages, some of which are as follows:

- Unlike the traditional television which requires wires for its working, IPTV works on various Internet services that transmit the video more efficiently and save space consumed by tangling wires.
- IPTV supports VoD that allows an individual user to select and watch the desired videos at anytime. In addition, the viewer can play, pause, or stop the running video.

- By using electronic programming guides, IPTV provides better interactivity between the service provider and the end user. For instance, while watching a match, viewers can check the player's status.
- IPTV also facilitates to capture real-time broadcast television and record it on the local server, connected through a network. This helps the end user to view the recorded program of his own choice.

23.7 BLOGGING

The term *blog* is the shortened form of *web log* which was discovered by Jorn Barger in the year 1997. It is basically a type or part of website maintained by an individual with regular entries of news on some particular topic (like education, entertainment, or politics), description of events, personal thoughts, and so on. Primarily, blogs include information in textual form, but they may also include hypertext, images, videos, and links to other blogs or web pages. An important property of blog is that it allows the readers to leave their comments and share their thoughts with the blog owner or other readers. The entries posted on blog are usually arranged and displayed in reverse chronological order and can be viewed in an HTML browser.

A blog is an innovative way to organize your thoughts and present articles. Moreover, it helps to keep up with the latest information posted on various sites making it a popular personal publishing tool. The process of creating, maintaining, writing, or adding content to a blog is known as *blogging* and the person who posts entries onto the blogs is known as *blogger*.

Types of Blog: There are different types of blog depending upon the purpose for which blog has been created, the content a blog contains, or the way the content has been added to a blog. Some common types of blog are listed below:

- **Personal Blogs**: Personal blog is the most common and traditional form of blog. It can be considered as an online diary maintained by an individual on which he/she records daily events, commentary, or articles. These blogs help the bloggers to share their feelings and thoughts with other members or friends.

- **Corporate Blogs**: Corporate blog, also known as *business blog*, is created and maintained for business purposes. Such blogs are mainly used for marketing or advertising the company's products to the outside world. Also, they can be used within the organization to promote communication and culture among the employees as well as to keep them up-to-date with the latest information about products, services, or activities.
- **Media Blogs**: Media blogs constitute blogs containing media-type content. For example, blog containing videos known as *vblog*, containing photographs known as *photoblog*, or containing sketches known as *sketchblog*.

Apart from these types, blogs can also be categorized on the basis of device type used to add content to the blogs. For example, blogs added by using mobile devices like cell phone are known as *moblog*.

Creating and Using a Blog: Many websites offer to create a blog account free of cost. Some of these websites include *www.blogger.com*, *www.weebly.com*, *www.blogspot.com*, *www.thoughts.com*, and many more. Here, we will explain the steps to create a blog on *www.blogger.com* website.

1. Type *www.blogger.com* in the address bar of the web browser and press **Enter**. The homepage of the website opens (see Figure 23.6).

Figure 23.6 Home Page of Website

2. Click on **CREATE A BLOG**. This displays Create a Google Account webpage (see Figure 23.7).

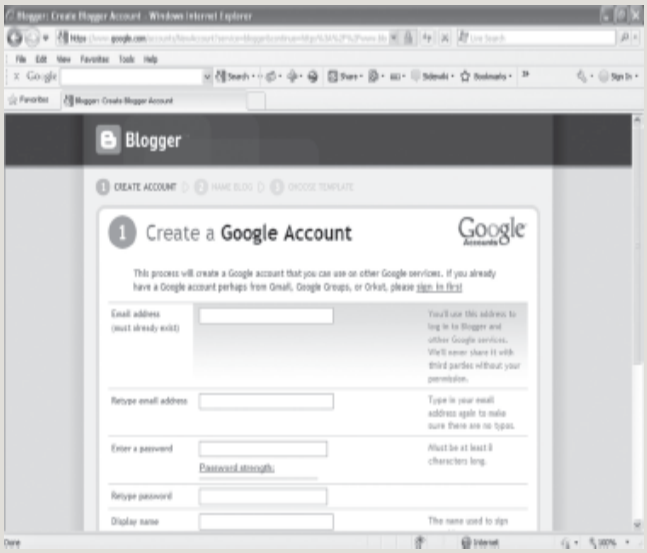


Figure 23.7 Create a Google Account Webpage

3. Enter details like e-mail address, password, display name, birth date, and verification word in the webpage.

4. After entering the details, click **CONTINUE** on the bottom right of the webpage. This displays the Name your blog webpage (see Figure 23.8).

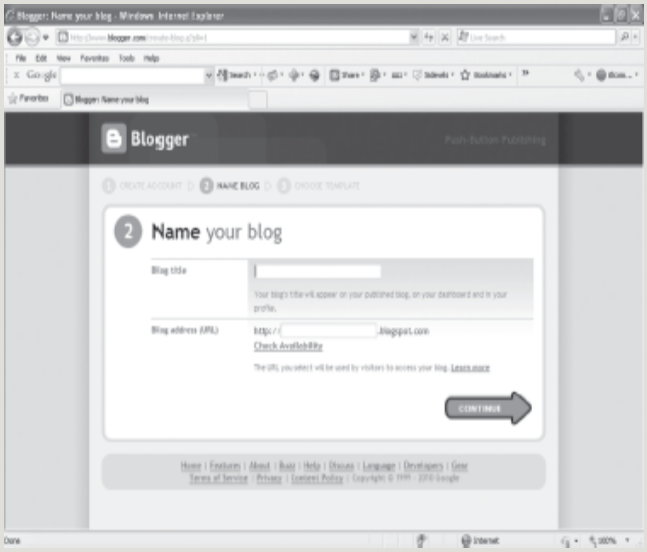


Figure 23.8 Name Your Blog Webpage

5. Enter the desired **Blog title** and **Blog address (URL)**.
6. Click **CONTINUE** to display the Choose a starter template webpage (Figure 23.9).

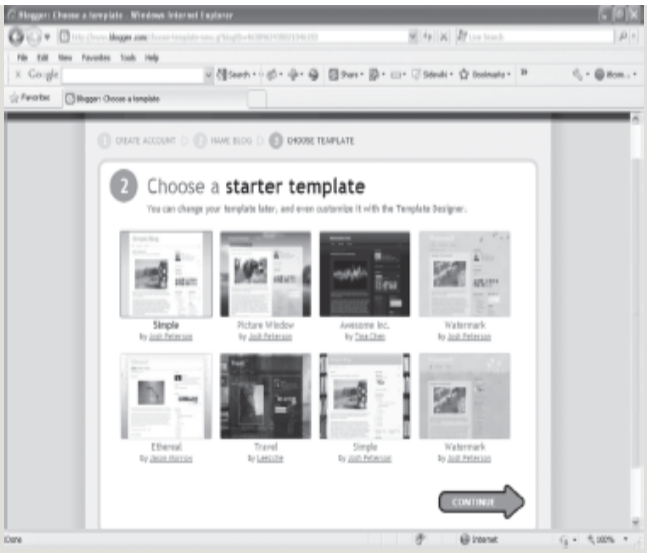


Figure 23.9 Choose a Starter Template Webpage

7. Select a design of your choice from the given templates. Whatever design you choose, your blog will look like that.
8. Click **CONTINUE** to complete the blog creation process. A webpage appears that informs that your blog has been created (see Figure 23.10) and now, you can add posts to it or customize it to change its appearance.

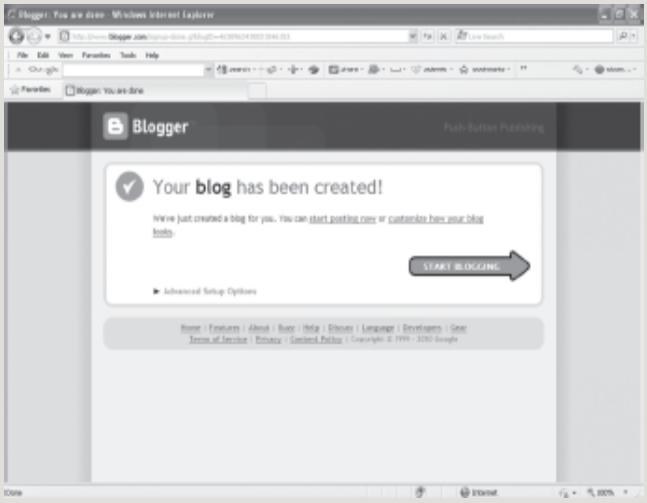


Figure 23.10 Blog Created

9. Click **START BLOGGING** to start using your blog. A webpage showing the display name (as specified by you) at the top appears. By default, the **Postings** tab is activated on this webpage (Figure 23.11).
10. Write the text in the text area provided to add a new post. You can also format the text using the toolbar appearing above the text area.

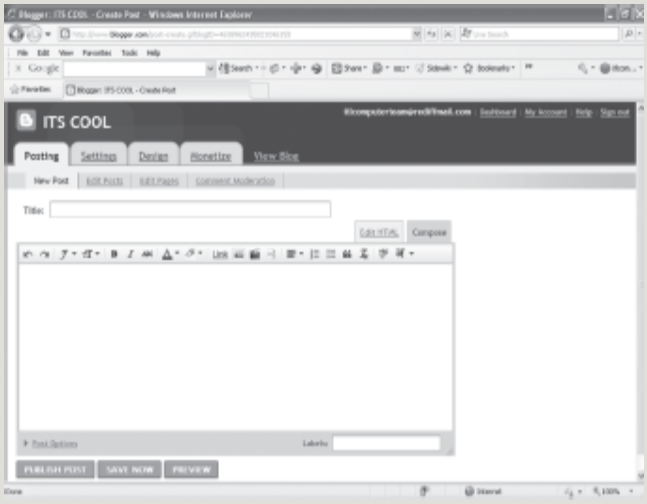


Figure 23.11 Webpage with Posting as the Active Tab

11. To add an image or video to the post, click **Insert image** button ( ) or **Insert a video** button ( ) on the toolbar, respectively (Figure 23.12).

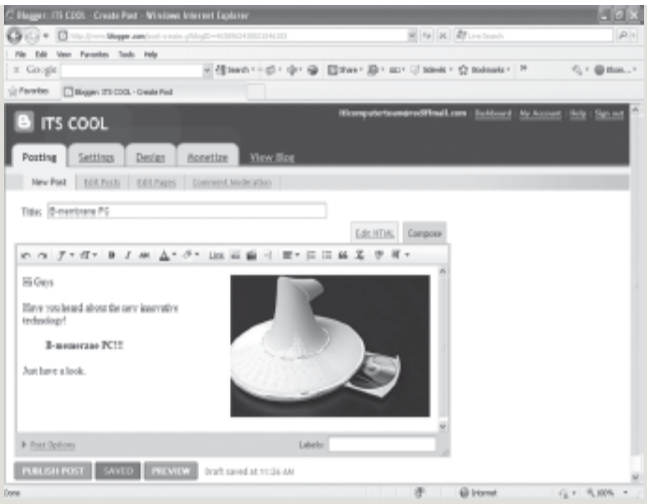


Figure 23.12 Text and Image Added to the Blog Post

12. Click the **PUBLISH POST** button at the bottom left corner to publish the blog post. A webpage appears that indicates your blog post has been published. Also this page provides links that enable you to view the post, edit the post, or create a new post (Figure 23.13).

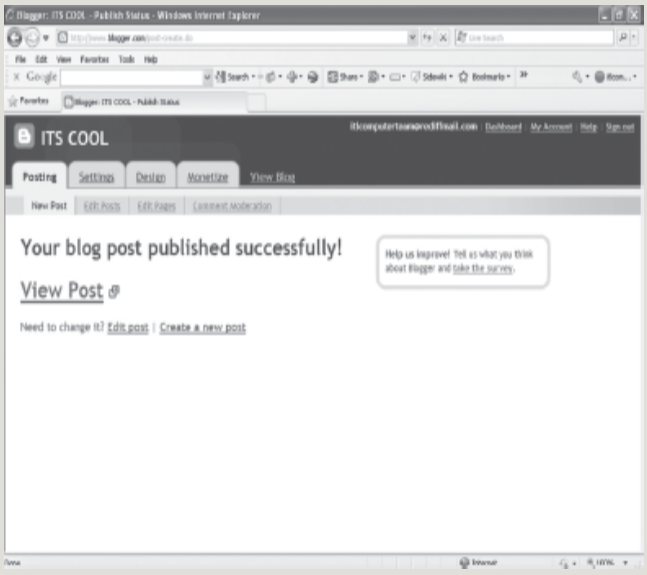


Figure 23.13 Publishing Blog Post

13. Click on **View Post** link to view your post. A webpage similar to as shown in Figure 23.14 appears that shows your blog posts.

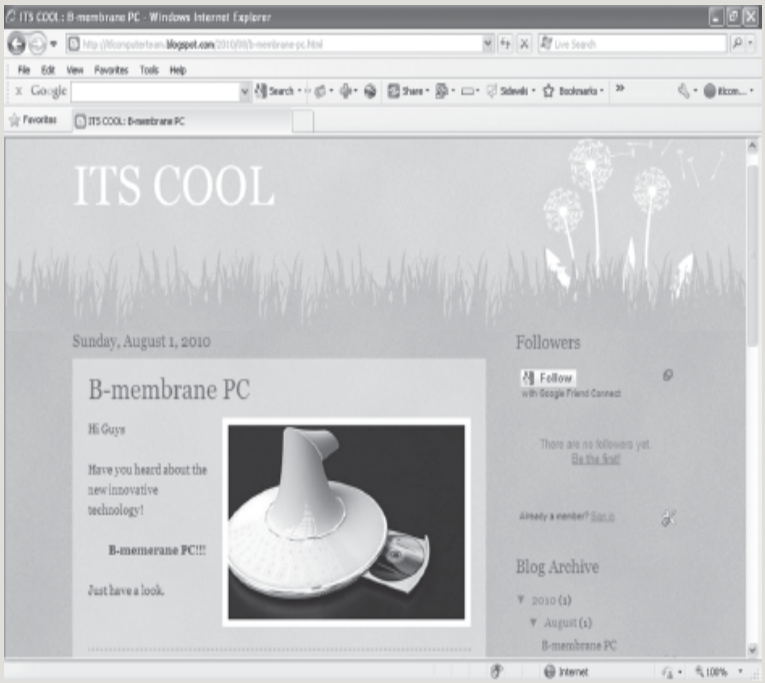


Figure 23.14 Viewing the Blog Posts

23.8 RADIO FREQUENCY IDENTIFICATION (RFID)

RFID is an auto-identification technology, which is used to capture and transmit identification details (including location) of an object wirelessly. RFID system consists of the three basic components, namely, *RFID tag* (*or transponder*), *reader*, and *processing device* (such as computer). An RFID tag contains the identification number and other data of the object on the silicon chip in encoded form. It is placed on or embedded in the object to be tracked. The stored data on RFID tag is accessed (i.e. read/write) by the reader. A reader is a device that activates the RFID tag (when it comes in the response range of the reader) by emitting radio frequency signal through antenna. Upon receiving the signal from the

reader, the tag responds by transmitting the data to the reader. This data are then sent by the reader to the computer connected with it for further processing (Figure 23.15).

Figure 23.15 An RFID System

FACT FILE

RFID versus Bar Code

Bar code technology is also an auto-identification technology; however, RFID is an improvement over barcode in many ways.

- RFID does not require any wire or line of sight between the reader and an RFID tag. This helps in identifying an object within a particular range without even seeing (or touching) it.
- RFID has read/write capabilities. The information stored on RFID tag can be changed whenever required.
- RFID can detect moving objects within long range (say 100 feet) and can work in harsh environment.
- More than one RFID tag can be accessed simultaneously.

Various applications of RFID technology are as follows:

- One of the most common uses of RFID is in asset tracking. By placing the RFID tags in (or on) the products, companies can keep a track of

their product, which minimizes the risk of misplacing, stealing, or underutilizing the product.

- Companies also use RFID technology in the supply chain management to trace their product during shipment.
- RFID is finding its way in payment systems. Using RFID, a person can pay his road toll without stopping his car.
- RFID is used in security and access control systems. Government has also started using RFID technology in making e-passport for security purpose.

23.9 BRAIN COMPUTER INTERFACE (BCI)

BCI, as the name implies, is a communication pathway between the brain (of either human or animal) and some external device like computer system. BCI devices are used to transmit or capture the brain signals, which can further be used to restore function or movement to sensory organs or to direct implanted prosthetic device as natural limbs. BCI is an important achievement in medical technology as it enables the persons suffering from some kind of physical disability to carry out their tasks like walking, picking up a glass of water, etc., just by thinking. As for now, BCI devices have succeeded in restoring damaged sight, movement, hearing, and in curing neurological problems such as depression. However, in the near future, it will be possible to intensify the human capacity and to influence someone's thinking (Figure 23.16).

Figure 23.16 A Human Interacting through BCI

The BCI works just because of the way the brain functions. Our brain is full of neurons, which fire whenever we think or perform some action. As the neurons fre, the electric signals are generated. These signals move at a very fast speed from neuron to neuron through a complex web of connections. Though the path that the electric signals follow is insulated by a white fatty substance in the brain called myelin, some electric signals get away. These signals are captured by the researchers and used for controlling some external device, activity, or for helping people to overcome their severe physical disabilities.

Types of BCI: BCI can be of the following types.

- **Invasive BCI**: In this type of interface, the BCI device is implanted directly into the brain of the patient. Therefore, the signals captured from the brain are of high quality. These high-quality signals can be used to control the robotic arms or legs implanted in the disabled people, thereby providing them functionality. These signals can also be used for restoring the vision by connecting external camera. But since these devices are implanted inside the brain, scar tissue can form over the device, which results in blockage of signals, reduced efficiency, and increased risk to the patient.
- **Partially Invasive BCI**: In this type of interface, the BCI device is planted outside the brain but inside the skull of the patient. It has lower signal strength as compared to invasive BCI, but the risk of scar formation over the device is less.
- **Non-Invasive BCI**: In this type of interface, the BCI device is planted onto the skull of the patient. It is the safest interface but has the least signal quality as compared to other interfaces. The non-invasive BCI device can enable the patient to move muscular implants and restore partial movement. Although it is easy to use, cheap, and portable, intensive training is required to carry out different functions.

Techniques of BCI: There are various means of reading the brain activity, some of which are as follows:

- **Direct Neural Contact**: It is an invasive technique in which a chip containing several pins (having width less than a human hair) is embedded into the brain such that the pins are in direct contact with the neurons. As the brain starts movement, the neurons fire and with the help of chip, scientists are able to read the firings of neurons in the brain. The language generated by the neuron firing is forwarded to a computer translator that decodes the neural language into computer language by using some special algorithms. The decoded information is transferred to another computer which in turn instructs the machine for the action to be performed. With the current technologies, this

technique is highly impracticable. However, the speed with which nanotechnology is advancing, someday it might become reality.

- **Electroencephalography (EEG)**: It is the first and most common non-invasive technique that was discovered by Hans Berger. In this technique, the electrodes, which are placed against the scalp of the patient, are used to capture the brain signals. Nowadays, this technique is widely used due to its low cost, ease of use, and high temporal resolution. But it is less accurate than direct neural contact and can capture low-frequency signals.
- **Magnetoencephalography (MEG)**: It is a newer non-invasive technique which is much more accurate than EEG. It communicates with the brain by recording magnetic fields produced by it. The main drawback of this technique is that it requires more equipments than any other technique. For using MEG, a room filled with super-conducting magnets and giant super-cooling helium tanks surrounded by shielded walls is required.
- **Functional Magnetic Resonance Imaging (fMRI)**: It measures the hemodynamic (the flow of blood and level of oxygen in the blood) related to the neural activity in the brain using uses MRI (Magnetic Resonance Imaging) technology. As MRI requires special equipments, this technique is costly.

23.10 IMMINENT TECHNOLOGIES

Evolution of technology in the current world shall prevail, as long as there remains gift of knowledge and wisdom in humanity. Striving for success has made humankind to make the best use of technology in making things easier and faster in everyday life. Some futuristic technologies are discussed below.

23.10.1 Pen Size Computer

Pen size computer is one of the major breakthroughs of this era. It is a revolutionary miniature computer that comes in the shape and size of a pen. The conceptual prototype of this computer was developed by a Japanese company, NEC Corp and NEC Design Ltd., way back in 2003, and was named as *P-ISM (Pen-style Personal Networking Gadget Package)* (Figure 23.17).

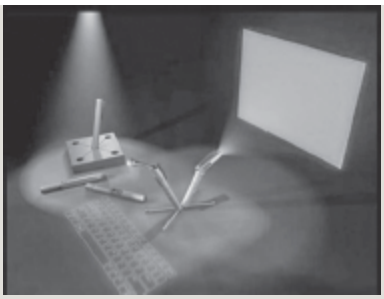


Figure 23.17 Pen Size Computer

The main idea underlying the concept of pen size computer is to break a computer into parts. Pen size computer is a set of five different components (pens) that altogether make up a computer and each component has a specific function to perform. All the pens are connected to each other wirelessly and communicate using Bluetooth technology. The description of different components of pen size computer is as follows (Figure 23.18):

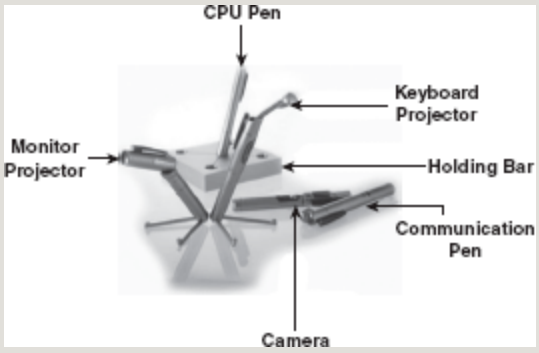


Figure 23.18 Components of Pen Size Computer

- **CPU Pen**: This pen serves as a CPU that administrates everything in the computer.

- **Keyboard Projector**: This pen projects a virtual keyboard on the desk or any other flat surface. It emits the laser beam on the surface, which generates a full-size, perfectly operating virtual keyboard. You can work on this virtual keyboard in the same manner as you do on the normal keyboard. Also, when you press the keys of virtual keyboard, you can listen the same tapping sound that is produced in case of normal keyboard.
- **Monitor Projector**: This pen projects an A4 size screen on a flat surfaceand displays the output produced by the computer on it. The resolution of the projector is 1024 × 768 approximately.
- **Camera**: This pen helps in video conferencing, video recording, etc, just like a webcam. It provides a three-dimensional visual communication.
- **Communication Pen**: This pen serves as a cellular phone and is used to connect the pen size computer to Internet through Tri-wireless modes (Bluetooth, 802.11B/G and cellular). It can serve as a pointing device (to enter the input) as well as contains earpiece.

All the pens of pen size computer rest on a *holding block* which keeps the mass storage together as well as recharges the batteries.

23.10.2 Napkin PC

Have you ever thought that your PC can resemble the design of a napkin holder? This revolutionary idea of PC prototype named as *Napkin PC* is given by Avery Holleman. He has also won Microsoft's Next-Gen PC Design Competition's first prize for this innovative idea.

The Napkin PC consists of a napkin holder having a base station PC, number of e-paper napkins, and a set of coloured pens. The e-paper napkin looks like a paper and works as a storage and display medium. It is reusable—can be rewritten thousands of times. The user can write or draw on the napkin (i.e. can give commands) using one of the pens. Anything that the user writes on the napkin is transmitted by the pen to the napkin interface using short-range RF technology. Not only this, the pen also keeps track of who is drawing what. Note that both the pen and the napkin interface interact with the base station using long-range RF (Figure 23.19).

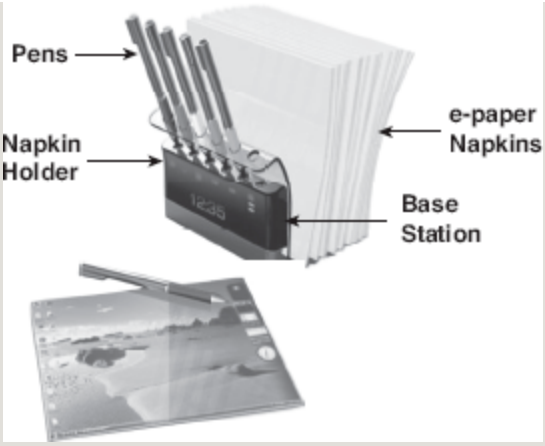


Figure 23.19 Napkin PC

The best part of Napkin PC is that it is very easy to work with because of the use of pen and paper. Besides this, it offers many other advantages, some of which are as follows:

- Multitasking will become easy. Unlike other computers where you have to shuffle through windows on a single interface, you can work on multiple napkin interfaces in parallel with much ease.
- E-paper can retain the images intact for indefinite period of time. Thus, the need for print is eliminated. Moreover, if the data on the e-paper napkin are no longer required, it can be simply returned to the base station as a fresh napkin.
- Napkin PC requires very little power, even battery is not required. It uses a single-layer circuit board for inductive power.
- In business meeting, entrepreneurs can work together to share their personal ideas by passing the napkins to each other and thus can work in collaboration.
- Artists can pin-up the e-paper napkins containing the pictures in an art gallery for exhibition while the engineers and architects can place them (where each napkin contains a small sketch) together to make a giant digital sketch for large-scale blue prints.

23.10.3 Roll Top Computer

Roll top computer is a portable and stylized computer concept unveiled by Orkin Design of Germany. The basic idea behind this concept is to enhance the portability of

conventional book-form laptops by turning them into unfurling and convolving portable computers. Due to enhanced portability of roll top computer, you can roll it like a newspaper (when not in use) and carry it around wherever you want.

The conceptual prototype of roll top computer comprises two parts: an energy-efficient and flexible *OLED (organic light-emitting diode) display* with multi-touch screen capabilities and a *tower* on which the display can be wrapped around (see Figure 23.20 (a)). The tower incorporates a power adapter, USB ports, webcam, and speaker within itself. A multimedia integrated pen is also incorporated in the device.

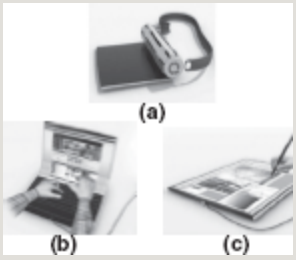


Figure 23.20 Roll Top Computer

The roll top computer can be used as a laptop, tablet PC, or a Television. As you unroll it (but not fully), it can be used as a 13-inch laptop/notebook (see Figure 23.20 (b)) with a portion of display being used as a keyboard. If you flatten it out, it can be used as a graphics tablet with an enclosed stylus pen (see Figure 23.20 (c)). When you fully unroll it, you get a 17-inch flat screen display to view the videos. This conceptual device also contains a hanging belt using which it can be easily carried over the shoulders. In short, the roll top computer is an all-in-one

gadget that removes the burden of carrying heavy laptop bags.

23.10.4 B-membrane PC

B-membrane is an innovative and revolutionary PC prototype designed by a Korean designer Won-Seok-Lee. On the first look, it does not seem like a conventional computer at all; rather it resembles a space ship. It contains all the computer components within itself. The features that make the B-membrane computer concept interesting include:

- A *built-in projector* that is capable of projecting the computer screen on almost any surface at which the projector is pointed to. The interesting thing is that the projector is *omnidirectional*, which means you can point it in any direction on the surface. The use of projector eliminates the need of bulky monitors completely.
- A membrane like *touch screen keyboard* that appears only when you need to use it.
- An *optical drive* incorporated into the design using which you can play CDs or DVDs on the computer (Figure 23.21).

Figure 23.21 B-membrance PC

Apart from being used as a PC, B-membrane computer can also be used as a source of ambient light. When you turn off this computer, its projector transforms it to an ambient light effect system that can display lighting on any surface. One aspect that seems unfavourable to B-

membrane computer is its shape and size, which makes it difficult to carry it around.

23.10.5 USB Memory Stick

A Japanese designer Mac Funamizu devised the idea of USB memory stick that combines the concept of disco lights with flash drive. The USB memory stick looks similar to pen/flash drive and has a transparent (glass) body that illuminates to present its storage capacity as well as the data stored in it. The glass shows different colours of light to indicate different type of data stored in the flash drive. For example, it shows pink light for images, blue for documents, and green for movies. However, if the glass shows no light, then it indicates no data are stored in it. On the other hand, if the glass is filled with lights then it means no storage space is available in the flash drive (Figure 23.22).

Figure 23.22 USb Memory Stick

LET US SUMMARIZE

1. *E-commerce* is a way of carrying business communications and transactions using computers over Internet. However, most people think E-commerce to be just buying and selling of products and services online on the Internet.
2. There exist four distinct categories of e-commerce applications: *B2C, B2B, B2A,* and C2C.
3. There exist five important processes in an e-commerce lifecycle; these are *information sharing, ordering, payment, fulfilment*, and *support and services*.
4. The method of making payment on the Internet is essentially the electronic version of the traditional payment system, used in everyday life. Generally, there exist three major types of payment mode: *e-cash, e-cheque*, and *credit card*.
5. *EDI* is defined as a collection of public standard message formats and a data element dictionary that allows trading

partners to exchange data in a simple way using any electronic messaging service. It is a transmission between businesses of information in standard, computer-readable format which includes electronic order placement, electronic shipping notification, electronic invoicing, and many other business transactions which computers can actually perform better than people do on the computer network.

6. WAP is the worldwide standard protocol stack specified for wireless communication networks by the WAP Forum. It enables developing wireless applications independent of the wireless access technology being used and accessing Internet using mobile devices.
7. Smart cards are small, portable pocket-size cards having data storage ability. Most smart cards incorporate an ICC inside them having some computational functionality and I/O support. They come in various sizes ranging from as small as mobile phone prepaid cards to as big as credit/debit cards.
8. *IPTV* is a system that provides the television services to the end users via packet switched network. This network uses Internet protocol suite to transmit audio, video, and control signals, thereby providing quality of service, security, and reliability.
9. The term *blog* is the shortened form of *web log* which was discovered by Jorn Barger in the year 1997. It is basically a type or part of website maintained by an individual with regular entries of news on some particular topic (like education, entertainment, or politics), description of events, personal thoughts, and so on.
10. There are different types of blog depending upon the purpose for which blog has been created, the content a blog contains, or the way the content has been added to a blog. Some common types of blog include *personal blogs, corporate blogs, and media blogs*.
11. *RFID* is an auto-identification technology, which is used to capture and transmit identification details (including location) of an object wirelessly. RFID system consists of the three basic components, namely, *RFID tag (or transponder), reader*, and *processing device* (such as computer).
12. *BCI*, as the name implies, is a communication pathway between the brain (of either human or animal) and some external device like computer system. BCI devices are used to transmit or capture the brain signals, which can further be used to restore function or movement to the sensory organs or to direct implanted prosthetic device as natural limbs.
13. BCI can be of three types: *invasive, partially invasive*, and *non-invasive*. Direct neural contact technique is an example of invasive BCI while EEG and MEG are the most commonly used non-invasive BCI techniques.
14. *Pen size computer* is one of the major breakthroughs of this era. It is a revolutionary miniature computer that comes in the shape and size of a pen. It is a set of five different components (pens) that altogether make up a computer, and each component has a specific function to perform. These components include *CPU pen, keyboard projector, monitor projector, camera, and communication pen*.
15. The *Napkin PC* consists of a napkin holder having a base station PC, number of e-paper napkins, and a set of coloured pens.
16. *Roll top computer* is a portable and stylized computer concept that comprises two parts: an energy-efficient and flexible *OLED display* with multi-touch screen capabilities and a tower on

which the display can be wrapped around. The tower incorporates a power adapter, USB ports, webcam, and speaker within itself.

17. *B-membrane* is an innovative and revolutionary PC prototype that resembles a space ship. It comprises a built-in projector, membrane like touch screen keyboard, and an optical drive. It can also be used as a source of ambient light when it is turned off.
18. *USB memory stick* looks similar to pen/flash drive and has a transparent (glass) body that illuminates to present its storage capacity as well as the data stored in the flash drive.

EXERCISES

Fill in the Blanks

1. ..... is referred to as a way of carrying business communications and transactions using computer over Internet, or any other form of network.
2. ..... In B2B environment, e-commerce can be used in ....., ..., and .....relationship.
3. The fundamental difference between the electronic payment system and the traditional one is that every transaction is carried in the form of .....
4. The goal of EDI is to provide a ..... and ..... to exchange data in an electronic environment.
5. All information in e-commerce is expressed in the form of ....., which makes the product more versatile.
6. In ....., the BCI device is planted outside the brain but inside the skull of the patient.
7. In Napkin PC, ... orks as a storage and display medium.
8. IPTV sends data to the end user(s) in two ways, namely,..... and .....
9. A ..... is an innovative way to organize your thoughts and present articles.
10. The basic idea behind ..... concept is to enhance the portability of conventional book-form laptops by turning them into unfurling and convolving portable computers.

Multiple-choice Questions

1. Which of these is a type of electronic payment?
    1. E-cash
    2. E-cheque
    3. Credit Card
    4. All of these
2. Which of these is not an e-commerce activity?
    1. B2B
    2. C2B
    3. B2A
    4. C2C
3. Which of the following language is used by WAP?
    1. HTML
    2. WML
    3. XML
    4. Both (a) and (b)

4. Which of the following is a process in e-commerce life cycle?
    1. Service
    2. Payments
    3. Ordering
    4. All of these
5. Which component of Pen PC is used to connect the computer to Internet?
    1. Communication pen
    2. Camera pen
    3. CPU pen
    4. None of these
6. Which of the following is not a component of RFID system?
    1. Transponder
    2. Reader
    3. Set-top box
    4. Computer
7. Which of the following is not an advantage provided by IPTV?
    1. Online bill payment
    2. No wiring
    3. Video on demand
    4. Better interactivity
8. Which of the following is an invasive BCI technique?
    1. EEG
    2. Direct neural contact
    3. MEG
    4. MRI
9. Which of the following designers has devised the concept of Napkin PC?
    1. Mac Funamizu
    2. Orkin Design
    3. Won-Seok-Lee
    4. Avery Hollman
10. Which of the following is not included in a B-membrane PC?
    1. Projector
    2. Touch screen keyboard
    3. Multimedia pen
    4. Optical drive

State True or False

1. E-commerce refers to the physical transfer of goods on the Internet.
2. All processes in e-commerce are expressed in terms of numeric values.
3. Most cell phones and WAP-enabled PDAs have data transfer rate of 14.4 kbps or less as compared to a typical 56 kbps modem.
4. HTML can be used instead of WML in WAP.
5. Most web pages are designed for a resolution of 640 × 480 pixels, which is acceptable for a desktop computer.
6. In Pen PC, the pens communicate with each other via RF technology.
7. The USB memory stick fashes blue light for images.
8. Won-Seok-Lee has devised the concept of B-membrane PC.
9. E-paper can retain the images intact for an indefinite period of time.

10. In IPTV, the local exchanges receive the incoming data, reassemble the packets into coherent video stream, and decode it to be viewed on the television or some other form of screen.

Descriptive Questions

1. What do you mean by e-commerce? Discuss its advantages and disadvantages.
2. Explain the concept of EDI. List some benefits of EDI.
3. What is brain computer interface? Explain its types and techniques.
4. What is the purpose of blogs? List the steps for creating a blog.
5. Discuss RFID along with its applications.
6. Write short notes on:
    1. Smart card and its categories
    2. Applications of WAP
    3. Internet protocol television
    4. USB memory stick
7. Explain the futuristic computers concept in brief.

ANSWERS

Fill in the Blanks

1. E-commerce
2. Procurement, Order fulfilment, Managing trading-partner
3. Digital string
4. Standard format, Data element dictionary
5. Binary values
6. Partially invasive BCI
7. E-paper napkin
8. Unicast, Multicast
9. Blog
10. Roll top computer

Multiple-choice Questions

1. (d)
2. (b)
3. (b)
4. (d)
5. (a)
6. (c)
7. (a)
8. (b)
9. (d)
10. (c)

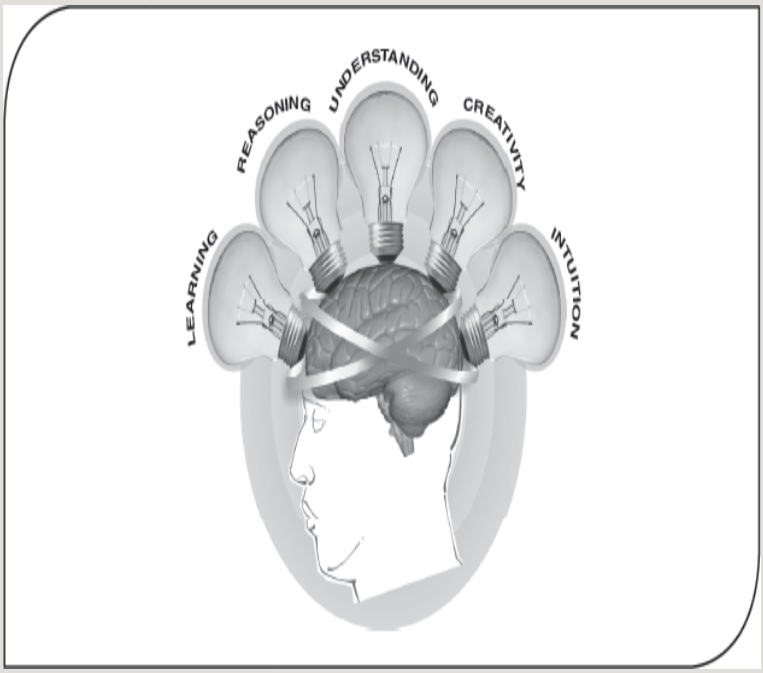
State True or False

1. True
2. False
3. True
4. True

5. True
6. True
7. True
8. False
9. True
10. True

Chapter 24

Artificial Intelligence (AL)



This chapter introduces the reader to the fundamentals of Artificial Intelligence (AI). Artificial Intelligence is a science, which focuses on creating machines that engages on the behaviours performed by humans. The chapter begins by giving an overview of AI. The discussion continues with knowledge-based systems and various searching algorithms used in the development of AI system. The later part of the chapter explains with the natural language processing and importance of expert systems in future implications. The chapter concludes by providing a brief discussion on neural networks and AI languages.

CHAPTER OBJECTIVES

After reading this chapter, you will be able to understand:

- The concept of Artificial Intelligence, which is considered as a science and engineering of making intelligent machines, especially intelligent computer programs
- Knowledge-based systems, which make extensive use of knowledge to perform complex tasks.
- Various searching techniques used to find an optimal state out of available states.
- Natural language processing, which refers to using language for expression of thoughts and feelings.
- Expert systems that are used to find solutions to a large number of complex problem domains.
- Neural networks, which focus on pattern recognition and pattern generation, embedded within an overall focus on developments of systems

24.1 INTRODUCTION

For a very long time, even before the advent of digital computers, many scientists were convinced that machines could be made to exhibit intelligent behaviour. Hence, it was natural that as soon as computers appeared, researchers began to program them to do things, considered to be possible for the human mind only such as solving non-numerical problems, understanding the English sentences, and playing chess. In the last few decades, computers have shrunk in size and also in terms of cost. The memory inside the computer system has increased so much that it is now equivalent to a substantial portion of human brain's storage capacity. With the availability of new hardware and software, computers are being specifically developed for performing some very complex tasks to ease the pressure on human mind. For example, nowadays

computers are used to forecast weather conditions and simulate extraordinary galactic events like the birth of a star. Assessing these complex problems requires a lot of computational work, which puts tremendous strain on human mind. Scientists realized that human mind cannot be pushed beyond certain limits and thus they began working on the development of systems, which have certain level of intelligence, similar to that of human brain. This gave birth to the concept of *Artificial intelligence* (AI).

John McCarthy coined the term "Artificial Intelligence" in 1956 at the Massachusetts Institute of Technology (MIT). According to McCarthy, AI is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable. In simple terms, AI is the branch of computer science devoted to developing programs that enable computers to display behaviour that can broadly be characterized as intelligent. This chapter deals with the basic understanding of AI and some of its most important aspects such as expert system, natural language processing (NLP), neural networks, and robotics.

24.1.1 What Is Intelligence?

Previously, we defined AI as a branch of computer science that deals with the study and creation of computer systems that exhibit some form of intelligence. Now the question arises what is intelligence. We can define intelligence as the ability to acquire, retrieve, and use knowledge in a meaningful way. It includes both raw and refined knowledge and the ability to memorize, recall facts, and express emotions. Mainstream thinking in psychology regards human intelligence not as a single ability or cognitive process but as an array of separate

components. Research in AI has focussed chiefly on the following components of intelligence:

- **Learning**: It is the process of acquiring knowledge, skills, experience, or values by study, experience, or training.
- **Reasoning**: It refers to the ability of drawing conclusions that are appropriate to the situation in hand.
- **Understanding**: It refers to the identification of the significance, interpretation, or explanation for certain data or information. Simply put, it is the ability to employ knowledge.
- **Creativity**: It is the ability to generate new ideas or to conceive new perspectives on existing ideas. The creativity process involves producing ideas, which are original and potentially useful.
- **Intuition**: It is the inner knowledge, without rational processes and without being aware of how we know. Essentially, intuition is an uncanny sixth sense that tells people that whether they are right (Figure 24.1).

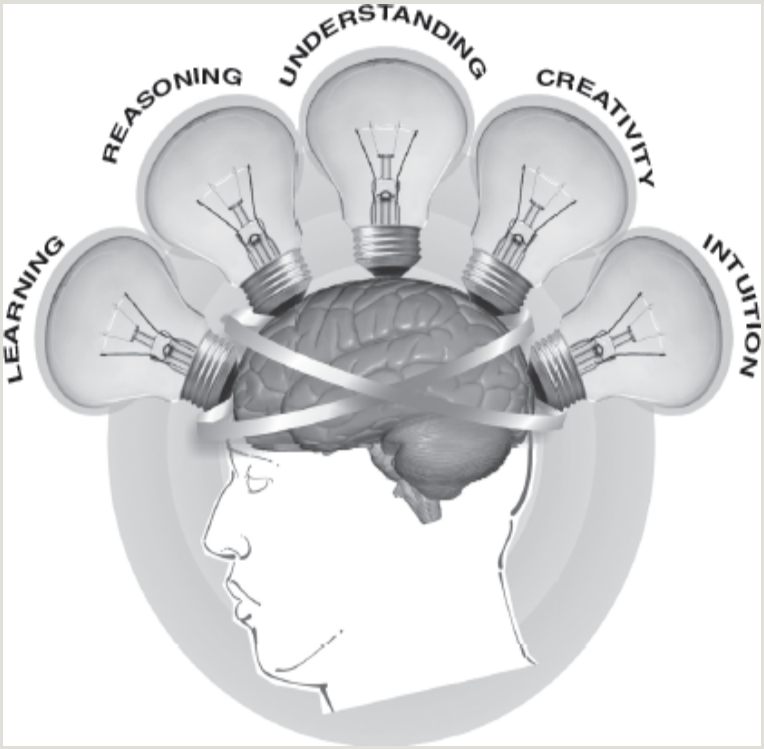


Figure 24.1 Intelligence

Scientists around the world have been trying to develop computer systems exhibiting the above features and they succeeded partially too. Systems have been developed to solve complex mathematical problems, plan strategies for military purposes, diagnose new medical diseases, and many such applications.

24.1.2 Turing Test

In 1950, Alan Turing proposed a test (known as Turing test) to prove the intelligence of a machine. This test provides the basis of what came out to be known as AI. There are three participants involved in this test: two humans and a machine. The idea was that a human and a machine, both outside the view, would give answers to a human interrogator, preferably in the form of a printed response. Based on the answers received, the interrogator would decide which respondent was the machine. The task of the machine is to deceive the interrogator into believing that it is a human being. The task of the other participant is to convince the interrogator that he/she is human. If the interrogator cannot reliably distinguish the human from the machine, then the machine does possess (artificial) intelligence (Figure 24.2).

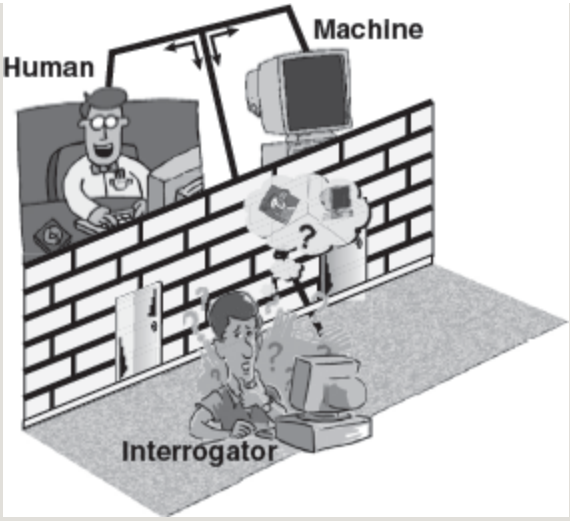


Figure 24.2 Turing Test

24.2 ARTIFICIAL INTELLIGENCE: DEFINITION

Researchers around the world have defined AI in varied perspectives. Some of the well-known definitions of AI are:

"The automation of activities that we associate with human thinking, activities such as decision-making, problem-solving, learning" —Bellman, 1978	"The study of how to make computers do things at which, at the moment, people are better" —Rich and Knight, 1991
"The study of mental facilities through the use of computational models" —Charniak and McDermott, 1985	"The branch of computer science that is concerned with the automation of intelligent behaviour" —Luger and Stubblefield, 1993

As evident from these definitions, AI falls into four categories as shown in Figure 24.3.

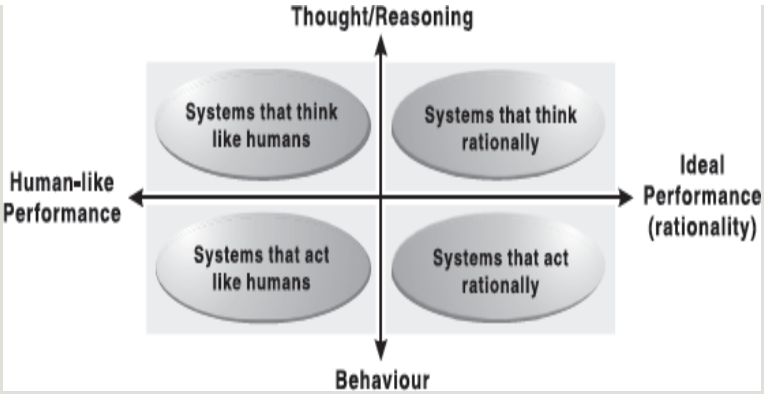


Figure 24.3 Categories of AI

Each of these categories represents a unique field in its own right but at the same time all four categories share a common ground.

- **Thinking Humanly:** Thinking comes naturally to human beings and we term it as common sense. The discipline that deals with the basic physics of human intelligence is known as cognitive science. Cognitive scientists try to construct theories of how the human mind works. In this process, a model of intelligent human behaviour is created to simulate on a computer to determine if it exhibits the same intelligent behaviour as humans.
- **Acting Humanly:** The system should pass the Turing test, which distinguishes intelligent entities from unintelligent ones. Acting humanly also includes physical interactions with environment such as speech recognition and robotics.
- **Thinking Rationally:** Rational behaviour means doing the right thing, which is expected to maximize goal achievement in available information. Thinking rationally involves mathematical logic as a tool (notation plus derivation rules). The problems and knowledge must be translated into formal descriptions. The system should use an abstract reasoning mechanism to derive a solution.
- **Acting Rationally:** It may not necessarily involve thinking but thinking should be in the service of rational action. Hence, acting rationally and thinking humanly are complementary. Most AI researchers in Computer Science go for acting rationally.

24.2.1 Domain Areas of AI

Various domain areas where AI is currently being used include:

- **Game Playing:** The greatest advancements in AI have been achieved in the field of game playing. For example, computer chess programs are now capable of beating humans. In May 1997, an IBM super computer called Deep Blue defeated the world chess champion Garry Kasparov.
- **Speech Recognition:** Computer speech recognition has reached a practical level for limited purposes. For example, airline organizations have replaced keyboard for flight information by a system using speech recognition of flight numbers and city names.
- **Understanding Natural Language:** Just getting a sequence of words into a computer is not enough, and even parsing sentences is not enough. The computer has to be provided with an understanding of the context of the text, and this is presently possible only for very limited systems.
- **Computer Vision:** The world is composed of three-dimensional objects, but the inputs to the human eye and computer's cameras are two dimensional. Some useful programs can work solely in two dimensions, but full computer vision requires partial three-dimensional information. At present, there are only limited ways of representing three-dimensional information directly, and they are not as good as what humans evidently use.
- **Expert Systems:** Computers are programmed to make decisions in real-life situations. For example, some expert systems help doctors diagnose diseases based on patient's symptoms. One of the first expert systems was MYCIN, which diagnosed bacterial infections of the blood and suggested treatments. This system performed better than medical students or practicing doctors.
- **Heuristic Classification:** A heuristic is a way of trying to discover something or an idea embedded in a program. An example is advising whether to accept a proposed credit card purchase. Information is available about the owner of the credit card, his record of payment, and also about the item he is buying and about the establishment from which he is buying it (for example, whether there have been previous credit card frauds at this establishment).
- **Neural Networks:** Computers are programmed to simulate intelligence by attempting to reproduce the types of physical connections that occur in human brains. Neural computing systems mimic the brain through a network of highly interconnected, processing elements, which give them learning capabilities and enable them to recognize and to understand subtle or complex patterns.
- **Robotics:** It deals with programming computers to see, hear, and react to other sensory stimuli. In the area of robotics, computers are now widely used in assembly plants, but they are capable of only limited tasks. Robots have great difficulty identifying objects based on appearance or feel, and they still move and handle objects clumsily.

THINGS TO REMEMBER

Applications of AI

The potential applications of AI are numerous. They extend from the military for autonomous control and target identification, to the entertainment industry for computer games and robotic pets. It can also be applied in big establishments dealing with huge amounts of information such as hospitals, banks, and insurances, which can use AI to predict customer behaviour and detect trends.

Researchers have still not been able to produce systems, which can match the natural abilities possessed by an average human being. These abilities include learning new sounds as well as recognizing them later, adapting to new conditions, recognizing and remembering numerous objects through eyes, and many such things. These are some of the problems faced by researchers pioneering in the field of AI, and all these problems require breakthroughs before one can expect a system to behave and react like humans and exhibit 100 per cent intelligence. Currently, AI focuses on lucrative domain-specific applications, which do not necessarily require the full extent of AI capabilities. There is little doubt among researchers that artificial machines will be capable of intelligent thought in the near future. It is just a question of “how and when.”

24.2.2 Brief History of AI

Contrary to popular belief, AI did not start with the first computer. The first computer was developed in early 1940s. The discovery of the computer allowed AI to become possible. However, the link between human intelligence and machines was not made until the early

1950s when Norbert Wiener, an American computer scientist, observed that all intelligent behaviour was the result of feedback mechanisms, which could also be simulated by machines. Soon after, Allen Newell and Herbert Simon developed “The Logic Theorist,” considered by many to be the first AI program.

In 1956, John McCarthy organized a conference in Dartmouth on artificial Intelligence, thus coining the term. The conference laid the groundwork for the research on AI. In 1958, McCarthy developed the LISP (list processing) language, which is still used in AI programming even today. The progress in this field continued and it seems, for now, that we can expect to see many more advances. Some of the main highlights in the life of AI are listed in Table 24.1.

Table 24.1 Milestones in AI Development

Year	Milestones
1964	Danny Bobrow showed that computers were able to understand natural language enough to solve algebra word programs
1970	Colmcrauer created PROLOG
1974	First computer-controlled robot was developed
1979	The Stanford Cart, the first computer-controlled vehicle, was built The Journal of American Medical Association said that MYCIN was as good as medical experts
1981	Kazuhiro Fuchi announced Japanese Fifth Generation Project
1994	Mid-1990s brought major advances in all areas of AI including games, translations, reasoning, and virtual reality
1997	IBM computer Deep Blue beat chess world champion Garry Kasparov
1999	AI-based information extraction programs like web crawlers became an essential web tool
2000	Interactive robot pets became commercially available, the most famous being Sony's AIBO. MIT showed off Kismet, a robot able to express emotions. Carnegie Mellon robot Nomad explores remote areas of the Antarctica and locates meteorites
2002	Researchers at the MIT artificial Intelligence Lab, introduced Roomba, a vacuum cleaning robot
2005	Stanford robot won the DARPA Grand Challenge by driving autonomously for 131 miles along a desert trail. Also, Honda's ASIMO robot, a humanoid robot, is able to walk as fast as a human, delivering trays to customers in restaurant settings
2007	A team of researchers at the University of Alberta solved the game of Checkers
2009	The Blue Brain Project had successfully simulated parts of a rat's brain

AI is not just hype, as many real working applications use AI techniques. However, AI is not yet a reality either, as AI researchers are yet to develop a system that spans the full spectrum of intelligent behaviour.

24.3 KNOWLEDGE-BASED SYSTEMS

Knowledge is defined as the remembering of previously learned material. It is the psychological result of perception and learning and reasoning. It is the body of

truth, information, and principles acquired by humankind. In humans, knowledge is stored as a collection of interconnected neurons (brain cells) in the brain, which contains approximately 1012 neurons. These neuron interconnections provide approximately 1014 bits of potential storage capacity. In computers, knowledge is stored in the form of voltage states. The gap between human and computer storage capacities has come down with computer now being able to store hundreds of gigabytes of data. However, there still exists a wide gap between humans and computers in terms of efficiency and the way data are represented. Knowledge plays an important role in building intelligent systems. Earlier, researchers developed general-purpose problem-solvers as an approach to prove theorems in geometry, perform complicated operations, and guide robots to perform the required tasks. All the systems made for these purposes failed to be of great help, as they required much of human interference for both identifying and finding solution to problems. These systems became effective only after the steps to achieve solution were incorporated with specific rules and facts. These systems were very general in nature, i.e. these were developed to solve different kinds of problems and were not specific to one problem. This realization led to the design of systems commonly known as knowledge-based systems. Many knowledge-based systems have been developed for different fields, for example, for diagnosing diseases, for configuring computer systems, in electrical circuits and financial markets, for scheduling airplanes and buses, and so on.

THINGS TO REMEMBER

MYCIN

Knowledge-based systems make extensive use of knowledge to perform complex tasks. This, in turn, led researchers to develop knowledge representation schemes and techniques to manipulate knowledge. In these systems, knowledge is stored in such a way that new knowledge can easily be integrated into the existing knowledge. This feature greatly simplified the construction and maintenance of knowledge-based systems. One of the earliest knowledge-based systems was MYCIN, which was developed to diagnose infectious blood diseases.

Knowledge plays an important role in building intelligent systems. Earlier, researchers developed general-purpose problem-solvers as an approach to prove theorems in geometry, perform complicated operations, and guide robots to perform the required tasks. All the systems made for these purposes failed to be of great help, as they required much of human interference for both identifying and finding solution to problems. These systems became effective only after the steps to achieve solution were incorporated with specific rules and facts. These systems were very general in nature, i.e. these were developed to solve different kinds of problems and were not specific to one problem. This realization led to the design of systems commonly known as knowledge-based systems. Many knowledge-based systems have been developed for different fields, for example, for diagnosing diseases, for configuring computer systems, in electrical circuits and financial markets, for scheduling airplanes and buses, and so on.

24.3.1 Knowledge Acquisition

A knowledge-based system depends on quantity as well as on quality of knowledge. For this, knowledge needs to be acquired from different sources such as specified procedures, rules, and facts. This acquisition of knowledge is one of the biggest bottlenecks in the construction of a knowledge-based system. Converting information gathered from various sources into knowledge and then using it is a big challenge. This may involve making raw knowledge explicit, identifying gaps in the knowledge already stored in the memory, acquiring and integrating new knowledge, and so on.

Knowledge acquisition is the foremost process in building knowledge-based systems. It is subject to in-depth research devoted towards the development of methods and software tools to provide knowledge content for knowledge-based systems. Many tools and techniques are now available for acquiring, analysing, and modelling knowledge. For example, the *matrix-based* technique involves the construction of grids to indicate problems encountered under some specific conditions. Some methodologies have been introduced to provide framework and standard knowledge to help knowledge acquisition activities and ensure that the development of each knowledge-based system is performed in an efficient manner.

24.3.2 Knowledge Organization

Knowledge-based systems usually require thousands of facts and rules to be incorporated in a system's memory. Therefore, it is essential that knowledge is stored in such a way that retrieval of information can be easy. Otherwise, a lot of time will be wasted in searching and testing large numbers of items (knowledge) stored in memory. One way of organizing knowledge in the memory is by using a method known as indexing. In this

method, knowledge is grouped in such a way that the "key word points" are used to access the required group. By using this method, only a fraction of knowledge is referred to instead of referring the whole memory. Thus, it saves time and leads to improved efficiency. Further, if appropriate representation scheme is used then the complexity involved in organizing knowledge can be avoided.

24.3.3 Knowledge Manipulation

In knowledge-based systems, decisions and actions are based on the manipulation of knowledge in specified ways. Typically, some form of inputs (usually from keyboard) initiates a search. This process requires that the known facts be compared, and possibly be altered, in some way. The manipulation process is continued (may require further inputs) until a final solution is found. Manipulation of knowledge is the computational equivalent of reasoning, which requires a form of inference or deduction, using the knowledge and inferring roles.

24.3.4 Knowledge Representation

Knowledge can be acquired, organized, and manipulated but it is meaningless unless represented in a proper manner. Some AI researchers thought that the best way to represent knowledge would be the way it is represented in human mind. Unfortunately, the representation of knowledge in the human mind is not yet known. For this reason, various AI languages and notations have been proposed for representing knowledge in systems exhibiting intelligence. Generally, these are based on logic and mathematics, which includes easily parsed grammar to ease processing in a system. Knowledge can be represented in many forms, as spoken or written words, in the form of graphical representation, as strings, and so on. The choice of

representation depends on the problem to be solved (Figure 24.4).

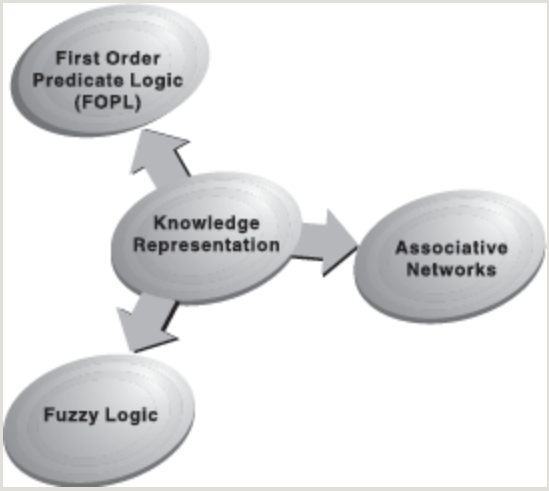


Figure 24.4 Different Knowledge Representation Schemes

THINGS TO REMEMBER

Knowledge Representation

There are various prerequisites for choosing a knowledge representation scheme. Generally, a suitable knowledge representation scheme should:

- Have sufficient expressive power for encoding the particular knowledge-based system.
- Possess a clean semantic basis, such that the meaning of the knowledge present in the knowledge base is easy to grasp, especially by the user.
- Permit efficient algorithmic interpretation.

- Allow explanation and justification of the solutions to be obtained.

AI researchers, for developing intelligent systems, have used several representation schemes. One of the most successful and widely used representation schemes is first-order predicate logic (FOPL). It has gained popularity because it is a well-developed theory, which uses valid forms of reasoning, and has good expressive power. However, this scheme cannot be used in areas involving common sense reasoning. Another representation scheme is fuzzy logic, which is a process of reaching conclusions based on information and facts that are not 100 per cent certain. Other type of representation scheme, which is commonly used, is associative networks. It is a graphical representation scheme and uses concepts and nodes where each node represents a concept and arcs are used to define the relations between the concepts.

First-Order Predicate Logic (FOPL): *FOPL* is one of the most widely used representation schemes developed by researchers for formal reasoning, primarily in the area of mathematics. This scheme offers a formal approach to reasoning and has a very strong base, which is essential when new systems are to be developed. It is a very descriptive representation scheme with a well-founded method of deriving new knowledge from a database. FOPL comprises logic, which contains symbolic representation to represent concepts. Like any logical theory, FOPL consists of specific design to construct syntactically correct statements (the well-formed formulas). Statements written in natural languages, for example, English, can be translated into a well-designed symbolic form. This symbolic form consists of the following symbols and rules.

- **Predicate:** This symbol returns a value as true or false. Capital letters and capital words like A, B, NOT, EQUAL are used to represent predicates.

- **Functions:** These are used to indicate relations and are denoted by small letters. For example, p, q, sister-of are used to represent functions.
- **Variables:** These are terms having different values and are represented by small letters, for example, x, y, and so on.
- **Constants:** These terms have fixed values, which cannot be changed and are denoted by small letters or numbers such as a, b, 3.
- **Logical Quantifiers:** There are two Quantifiers: existential and universal. Existential quantification (3x) means that there exists an x while universal quantifer (Vx) means for all x.
- **Connectors:** Symbols like v (disjunction), ~ (negation), a (conjunction), → (implication), and ↔ (equivalence) are used for connecting statements.

Example:

Let us discuss a simple example. Consider the following two statements.

1. All humans are mortal.
2. Maria is human.

Since all humans are mortal and Maria is a human being, it may be concluded that Maria is also mortal.

Hence, the following sentence is also correct.

Maria is mortal.

In FOPL, we represent the above discussion as:

For all objects h:

Human(h) → Mortal(h)

Thus, substituting Maria for h:

Human(Maria) → Mortal(Maria)

“for all” is written as ∀, so

((∀h)(Human(h) → Mortal(h)) ^ Human(Maria)) → Mortal(Maria)

Fuzzy Logic: Fuzzy logic is a type of knowledge representation that resembles human thinking. Like humans, it can measure imprecise or vague entities. Logic based on the two truth values, true and false, is sometimes inadequate when describing human reasoning. Fuzzy logic uses the whole interval between 0 (false) and 1 (true) to describe human reasoning. In two-valued logic, a proposition is either true or false, but not both. The truth or falsity, which is assigned to a statement, is its truth value. In fuzzy logic, statements are not necessarily only true or false, but also can be very unlikely, or more, or less certain.

Lotfi Zadeh, a professor at the University of California at Berkley, developed the concept of fuzzy logic in 1965. Fuzzy logic allows computers to emulate the human reasoning process, which makes decisions based on vague or incomplete data, by assigning values of degree to all the elements of a set. Some of the characteristics of fuzzy logic are:

- It is useful in manipulating information that is incomplete, imprecise, or unreliable.
- It deals with ambiguities, rather than only yes/no options.
- Fuzzy sets may overlap one another such as something that is both sweet and sour.
- It uses the operations AND, OR, and NOT.
- It can be implemented in hardware or software or in both.

The approach followed by fuzzy logic is different from the earlier schemes. It uses a simple rule “IF A AND B THEN C” approach for solving a control problem rather than modelling a system mathematically (as done in FOPL). It is empirically based and depends on a user's experience rather than on technical knowledge of a system. Let us consider a simple example to understand the execution of fuzzy logic rule. Suppose, we have an object which is made up of glass and it emits light. Then, the fuzzy logic notation will be as follows:

IF (object is made up of glass) AND (light is yellow)
THEN (object is a bulb) or

IF (object is made up of glass) AND (light is not yellow)
THEN (object may or may not be a bulb) (Figure 24.5)

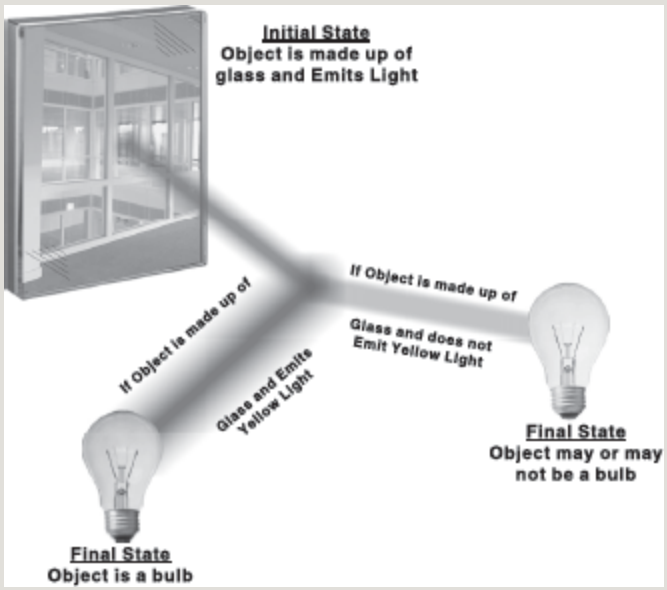


Figure 24.5 Fuzzy Logic

Fuzzy logic was originally conceived as a method for handling and sorting data, but researchers have used it successfully to control many system applications. An advantage of fuzzy logic is that it can be easily integrated into all types of products ranging from handheld to a large computerized processing system. Some applications where fuzzy logic has been very successful are consumer electronic goods such as washing machines, air conditioners, digital cameras, and video games. It is also used in automobiles for braking

purpose, for cruise control, and has been successfully incorporated into some microprocessors.

Automatic washing machines, which became popular in the 1960s, use fuzzy logic to control the washing system. The fuzzy logic automates the washing process by controlling the water and soap intake, draining, and rotation of the drum in a proper sequence. The fuzzy logic present in it senses the load size (number of clothes) and detergent concentration and accordingly adjusts the timing for wash cycles. Different types of clothes can be handled using different programmed cycles. For example, a wool wash needs a low temperature and less agitation as compared to cotton wash. Most automatic washing machines control the sequence using an electromechanical timer which has inbuilt fuzzy logic.

Associative Networks: An associative network, also known as semantic network, is commonly used as a form of knowledge representation. It is a directed graph consisting of labelled nodes and arcs, where the nodes represent concepts and arcs represent relationship between concepts. Computer implementations of semantic networks were first developed for AI and machine translation (MT), but they have long been used in philosophy, psychology, and linguistics. The most common form of semantic networks uses the relationship between nodes to represent IS-A and HAS relationships between nodes.

Figure 24.6 shows that a car IS-A vehicle; a vehicle HAS wheels. This kind of a relationship establishes an inheritance hierarchy in the network, with nodes lower down in the network inheriting properties from the nodes higher up. There are a number of commonly used semantic networks, which include definitional networks, assertional networks, implicational networks, executable

networks, learning networks, and hybrid networks. All these semantic networks have a declarative graphic representation, which can be used either to represent knowledge or to support automated systems for reasoning about knowledge.

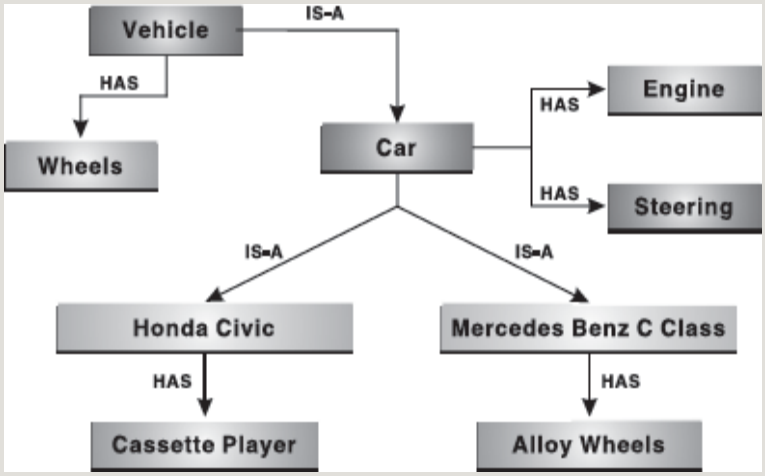


Figure 24.6 Associative Networks with IS-A and HAS Relationship

24.4 SEARCHING TECHNIQUES

In 1950s, researchers were engaged in developing systems in the field of theorem-proving and problem-solving. In both fields, the developed programs were characterized by complex algorithms that had a general solving capability independent of the problem domain at hand. Proving theorem was concerned with a given set of logics by a computer. Researchers in the field of problem-solving focused on the development of computer systems with a general capability for solving different types of problems. One of the early systems for problem-solving is GPS (general problem solver), developed by Alan Newell, Herb Simon, and Jack Shaw in 1960. It was developed as a research tool, which was responsible for separating task knowledge from the

problem-solving part. The aim involved was to write a single computer program, which could solve any problem. In GPS, a given problem is represented in terms of an initial state, a final state, and a set of steps to transform existing states into new states. However, GPS was not very successful. The main disadvantages of using this were:

- Representing a non-trivial problem in terms of GPS proved to be a daunting task.
- GPS turned out to be inefficient because it was a general problem-solver and was not designed for handling specific problems.

Although the success of GPS as a problem-solver was limited, it shifted the attention of AI researchers towards more specialized systems. This shift in attention from general problem-solvers to specialized systems led to the development of certain searching techniques to solve specific problems.

Problems are defined in terms of different states and then searching is used to find an optimal state out of available states. For every problem, there are a number of alternatives available to find a solution. For example, a program written to understand natural language should be able to find exact matching words rather than finding individual letters. A search algorithm is represented as a directed graph or in terms of tree formation. A search tree is a graph in which each node (state) has only one parent, except the starting node, also known as the root node. The nodes that branch out of the parent node are known as children nodes. The nodes at which no further Classification of node (expansion) is specified are known as terminal nodes. The starting node is the initial state and the goal node is the final state (where optimal solution is achieved). While performing a search, a solution to the problem is considered as a sequence of operations that helps in achieving goal state from the given initial state. The best solution is achieved when a

technique uses least operations or costs as compared to other states.

Using tree as a technique for finding a path, search is carried out by first identifying the child nodes from a parent node. Once this is accomplished, child nodes are generated from the parent node and this process is known as expansion of node. Search problems can be categorized into two groups:

- When only information relating to goal node and starting node is given, then search is performed using a uninformed or blind search.
- When some additional information is present other than the starting and goal node, then informed or heuristic search is used (Figure 24.7).

Figure 24.7 Searching Techniques

24.4.1 Uninformed Search

An uninformed search proceeds in a systematic way by exploring nodes in some predetermined order. This form of search can only distinguish between goal nodes from

other nodes in the search tree and only knows the length of the path or the cost of the path from the starting node to each node. The important search strategies used in this form include breadth-first search and depth-first search.

Breadth-First Search: In this searching technique, all the child nodes are placed at the end of the queue in any order. Search is performed by expanding all nodes at one level before proceeding to the next level (see Figure 24.8). Hence, all nodes at level i in the search tree are expanded before proceeding to level i +1. This strategy has the advantage of finding a minimal length solution whenever there exists one. However, many nodes are required to be expanded before reaching the goal node, which increases the time involved.

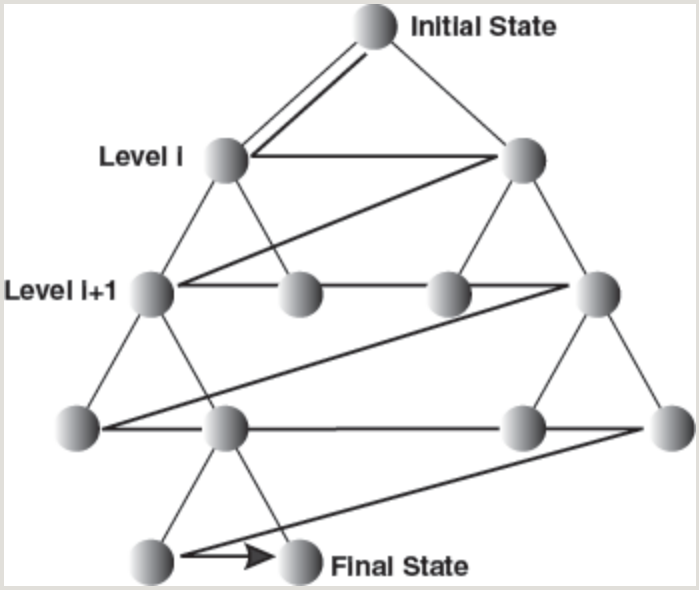


Figure 24.8 Breadth-First Search

Depth-First Search: Choosing the node at the deepest level of the tree performs the depth-first search.

This is accomplished by first generating a child node, say at level i, from the most recently expanded node at level i – 1. This process is continued until the goal node is found or when the deepest level of the tree is reached. If the goal node is not found or when the deepest level is reached, search hits a dead-end and returns back to expand nodes at a preceding level. This process continues until a goal node is found or a failure occurs (Figure 24.9).

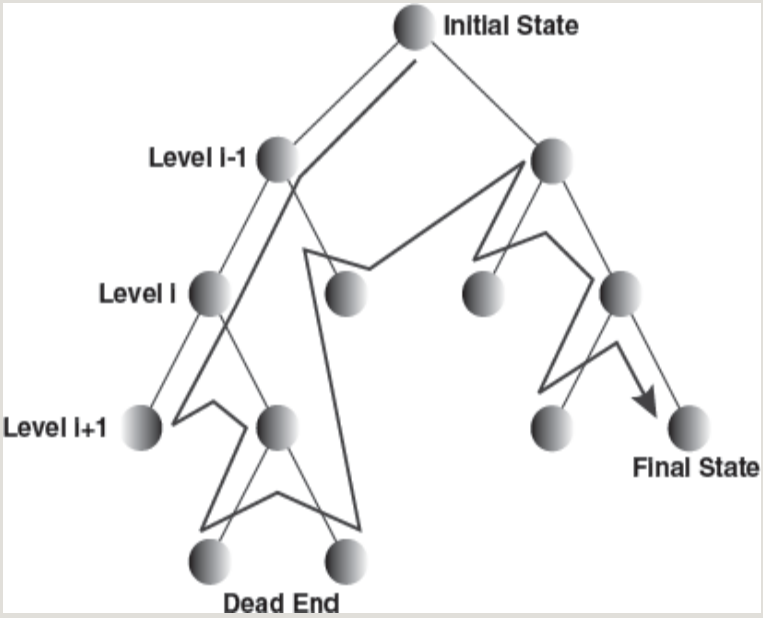


Figure 24.9 Depth-First Search

In cases where a problem has many solutions, depth-first search is a better option than breadth-first search, since there is a good chance that a goal node will be achieved after exploring only a small portion of the tree in depth-first search. However, depth-first search may not perform properly even when a solution exists at a relatively shallow level. Therefore, depth-first search is

not complete and not optimal and is avoided for problems with deep or infinite paths.

24.4.2 Informed Search

Uninformed methods of search are capable of systematically exploring the search tree to find goal state. However, such methods have proved inefficient in most cases. Therefore, informed search strategies are used to find optimal solution (goal state). When information other than the starting node and goal node is given, a more efficient search method, known as informed search, is used. In informed search, a priority-ordered queue is used. This ordering is determined by an evaluation function, which for each node on the queue returns a number that signifies the value of that node. Functions that calculate such estimates are called heuristic functions. Heuristic means a common sense rule (or set of rules) that is intended to increase the probability of finding a solution of a given problem. In simple terms, heuristic information is used in searching techniques to guess which nodes to expand next, thereby obtaining a solution in less time. Solutions based on heuristics do not provide any guarantee in terms of achieving success because heuristics may be inaccurate or it may not find the best answer. Various search techniques using informed search include best-first search and hill-climbing search.

Best-First Search: This technique uses heuristic information for performing search and finding the most promising (the least valued node) path to achieve the goal node. This technique always finds an optimal solution, even when local variance is present between the nodes. It also makes use of all estimates that were computed for previously generated nodes and makes a best selection among them. In this searching technique, the first element in the queue is removed, queue is expanded, and child nodes are placed in queue at either

end and then arranged in ascending order (Figure 24.10).

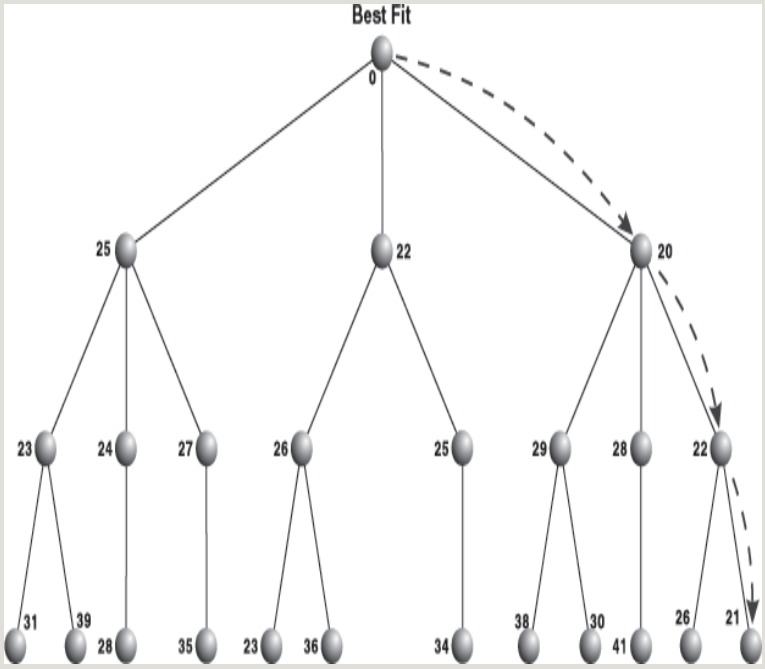


Figure 24.10 Best-First Search

Hill-Climbing Search: In this method, hill is referred as goal state of any given problem. At each point in the search path, a successor node that appears to lead quickly to the top of the hill (goal state) is selected for exploration. When all the child nodes have been generated, alternative solutions are evaluated using heuristic function. The path that appears most promising (optimal) is chosen and no further reference to the parent node or children node is retained (Figure 24.11).

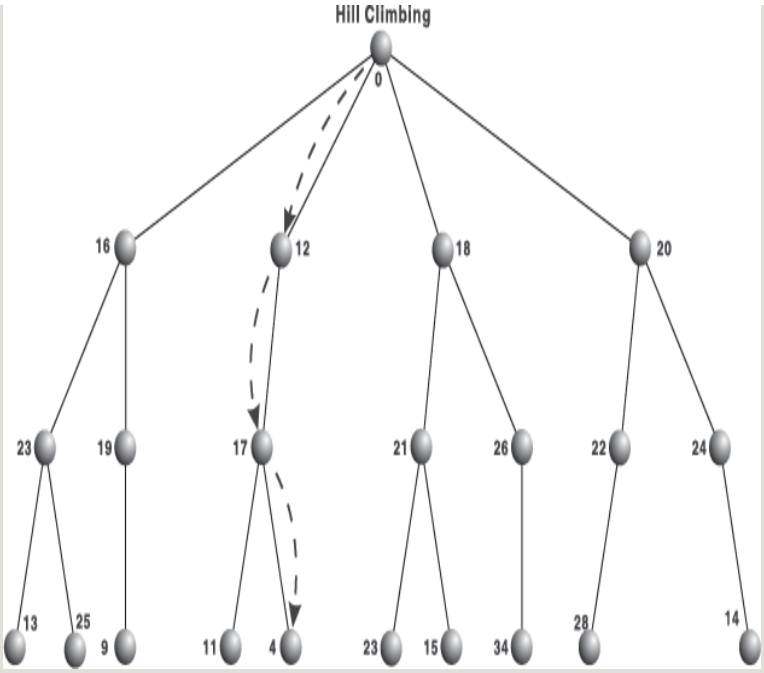


Figure 24.11 Hill-Climbing Search

24.5 NATURAL LANGUAGE PROCESSING (NLP)

Natural language refers to a medium (like English) for expressing thoughts and feelings. It requires both linguistic knowledge of the particular language being used and worldly knowledge related to the topic being discussed. Understanding natural languages is difficult. Language processing is divided into two tasks.

- Processing written text by using lexical, syntactic, and semantic knowledge of the language.
- Processing spoken language by using all the information mentioned plus additional knowledge about phonology.

NLP involves the interpretation, manipulation, or generation of human language by computer so that one can communicate with a computer as one communicates with another person. It studies the problems inherent in the processing and manipulation of natural language. Achieving this goal is not easy, as it requires proper understanding of language, knowing concepts related

with a word or a phrase. It is ironic that even though humans develop the computer, it still cannot understand the language spoken by humans. Humans perceive and communicate using five basic organs. These organs are eyes for visualizing, ears for hearing, skin for feeling, nose for smell, and tongue for taste. Among all these, visualization and hearing are very complex to comprehend. These are the major tasks facing researchers engaged in developing intelligent systems. This is because computers use mathematical processes to carry out tasks and understand these processes much more quickly than most humans. However, computers are not able to process human or natural language as compared to human beings. Thus, developing programs, which can relate to natural language, is very difficult. NLP involves many aspects of linguistics. These require analysis at low levels as well as at higher levels of syntax and discourse. The steps involved in the process of natural language comprehension are:

Morphological Analysis: Individual words are analysed into components, and non-word tokens (like punctuation) are separated from the words. For example, in the phrase “Jack's office,” the proper noun Jack is separated from the possessive suffix “'s.”

Syntactic Analysis: Linear sequences of words are transformed into structures to show relationship between one word and another. This step converts the list of words of the sentence into a structure that defines the units represented by that list. Constraints imposed include word order (“Adam 10 dollars” is an illegal constituent in the sentence “I gave Adam 10 dollars”), number agreement, and case agreement.

Semantic Analysis: In this step, the structures created by the syntactic analyser are assigned meanings. This step maps individual words into appropriate objects

in the knowledge base and creates correct structures corresponding to the individual words. For example, the sentence “ice-cream ate John” will be termed as a meaningless sentence.

Discourse Integration: This step involves introduction of entities in the sentence either defined explicitly or introduced earlier in a logical manner. Here meaning of an individual sentence depends on the sentences that precede it and influence the sentences, which is yet to come. For example, the word “it” in the sentence: “Tom needed it” depends on the preceding context, while the word “Tom” could influence the meaning of later sentences such as “He always laughs.”

Pragmatic Analysis: This is a high-level knowledge, which relates to the use of sentences in different context and how this context affects the meaning of the sentences.

THINGS TO REMEMBER

Natural Language Processing

One important use of natural language processing (NLP) is to employ everyday language to search databases. Most common database search engines presently use Boolean searching technique (NLP search engines), which signifies the relationship between the subjects and leads to more accurate searching. The search engines also break the language barrier with its inbuilt translation feature by using machine translation technology.

24.5.1 Uses of NLP

NLP is mainly used in the following areas:

- **Knowledge Acquisition:** This includes programs that can read books and manuals or the newspaper. Therefore, there is no need to explicitly encode all the knowledge needed to solve problems.
- **Information Retrieval:** This includes searching articles about a given topic. The programs are developed in such a way that they determine whether the article matches a given query.
- **Translation:** This requires machines to translate from one language to another automatically.
- **Helper Tools:** It is used in the areas of spell and grammar checkers, accessibility aids, screen readers (such as JAWS), computer-assisted language learning (CALL), lexicography, and MT.

24.6 EXPERT SYSTEMS

A recent application of AI is expert systems, which began to emerge during the 1970s. These systems have proven effective for finding solutions to a large number of problem domains, which earlier required the assistance of human mind.

"An expert system can be defined as embodiment within a computer of knowledge-based component from an expert skill in such a form that the machine can offer intelligent advice or take an intelligent decision about a processing function." It is an AI application that uses a knowledge base of human expertise to aid in solving problems. The degree of problem-solving is based on the quality of the data and rules obtained from the human expert. This system simulates the judgment and behaviour of a human or an organization that has expert knowledge and experience in a particular field. Typically, such a system contains a knowledge base containing accumulated experience and a set of rules for applying the knowledge base to each particular situation.

The primary goal of expert systems research is to make solutions available to decision-makers and technicians involved in an expert system. A system is designed in expert systems by manipulating and understanding knowledge, which is acquired through experience and extensive learning—rather than using simple information. Problem-solving is accomplished by

applying specific knowledge rather than a specific technique. This reflects the belief that human experts do not process their knowledge differently from each other, but they do possess different knowledge. The principal distinction between expert systems and traditional problem-solving programs is the way in which the problem-related query is programmed. In traditional applications, problem is encoded in both program and data structures. While in expert system approach, the problem is encoded in data structures only. The general architecture of an expert system consists of two components:

- A problem-dependent set of data declarations called the knowledge base or rule base, and
- A problem-independent program called the inference engine. Some expert systems are designed to act or perform like human experts, while others are designed to aid human expert.

24.6.1 Characteristics of Expert System

An expert system must exhibit the following set of characteristics:

- It should solve difficult programs in a domain, as good as or better than human experts.
- It should possess vast qualities of domain-specific knowledge to the minute details.
- It should permit the use of heuristics search process.
- It should be capable of accepting advice, modify, update, and expand.
- It should deal with uncertain and irrelevant data.
- It should communicate in their natural languages.
- It should possess the capacity to cater the individual's desire.

24.6.2 Applications Areas of Expert System

Expert System originally focused its activities on specific products such as spelling correctors for both general purposes and/or specialized use for governments as well as for specific clients. This initial experience led to improvement in technical knowledge in text processing and development of various types of different

applications. The main areas where expert systems are used include chemistry, biology, aerospace, finance, banking, and military.

Expert systems are currently being used in applications such as medical diagnosis, equipment repair, financial, and insurance planning, route scheduling for delivery vehicles, production control and training. For example, there are expert systems that can diagnose human illnesses, make financial forecasts, and so on. These expert-based applications of AI have greatly enhanced the productivity.

The early knowledge-based systems were mostly concerned with the field of medical diagnosis only. The best-known knowledge-based system in medicine, developed in the 1970s, was MYCIN. The development of this knowledge-based system took place at Stanford University. This system assists medical specialists in internal medicine to the diagnosis and the treatment of a number of infectious diseases, particularly meningitis and bacterial septicaemia. When a patient shows signs of such an infectious disease, usually a sample of blood and urine is taken in order to determine the bacterium species that causes the infection. MYCIN gives an interim indication of the organisms that are most likely to be the cause of the infection. Given this indication, MYCIN advises the administration of a number of drugs that control the disease by suppressing the indicated organisms.

With advances in the last decade, today's expert system clients can choose from a variety of commercial software packages with easy-to-use interfaces.

24.6.3 Users in Expert Systems

Generally, three individuals are involved in expert systems. These are:

- **End-User:** The end-user is an individual who uses the system for its problem-solving assistance
- **Problem-Solving Expert:** This expert builds the knowledge base for a system and is responsible for transferring the entire knowledge about the specific domain to the system.
- **Knowledge Engineer**: The knowledge engineer is concerned with assisting the representation chosen for the expert system. This engineer defines the inference engine used to process knowledge, which is required to obtain useful problem-solving activity (Figure 24.12).

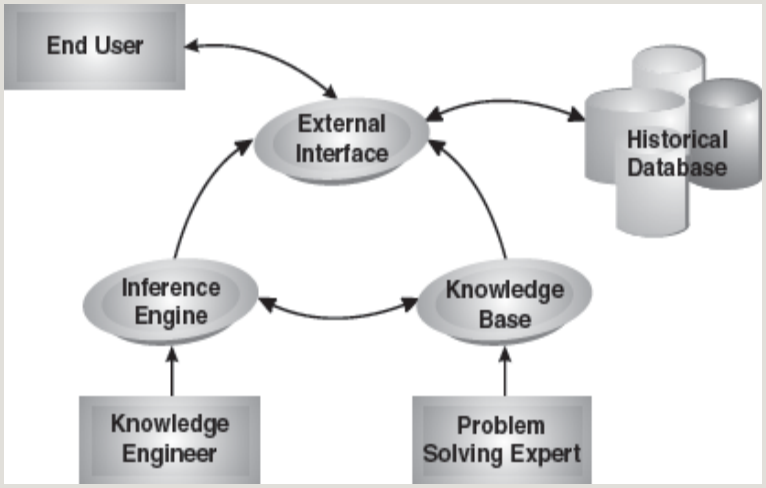


Figure 24.12 Expret System Users

24.6.4 Benefits of Expert Systems

The following are the benefits of expert system:

- Expert systems work faster than humans, which means that less number of workers are needed. This reduces the cost of the product and increases the output.
- Expert systems provide consistent advice, thus reducing the error rate.
- Expert systems make knowledge accessible to people easily, which query the system for some specific advice.
- Expert system provides a great amount of flexibility to the system.
- Expert systems are useful in interdisciplinary domains where multiple experts are required.

24.6.5 Limitations of Expert Systems

The following are limitations of expert systems:

- It is very specific in nature. The major problems lie in extracting the knowledge, building, and maintaining large knowledge bases.
- Lack of proper knowledge representation hampers the progress in expert system development.
- Program description is language dependent.
- Expert system development is quite expensive and time consuming.
- Present-day expert system lack human reasoning to some extent.

24.7 NEURAL NETWORKS

Neural networks have emerged as a specialized field within AI and engineering disciplines with collective efforts of engineers, physicists, mathematicians, and scientists (both computer and neuroscientists). Here, the area of research focuses on pattern recognition and pattern generation, embedded within an overall focus on development of systems.

Neural networks are computing systems, which imitate human brain through a network of highly, interconnected processing elements. These processes give these networks learning capabilities and enable them to recognize and to understand complex patterns. The key element of neural network is the presence of the information processing system. This system is composed of a large number of highly interconnected processing elements (neurons) working together to solve specific problems. Neural networks just like humans learn by example, that is, they are configured for a specific application such as pattern recognition or data Classification, through a learning process (Figure 24.13).

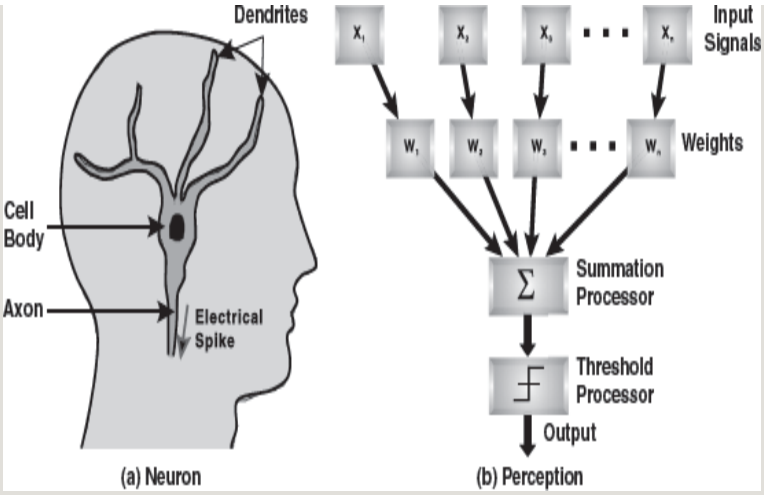


Figure 24.13 Neuron and Perceptron

A perceptron is one of the earliest neural network model built by Rosenblatt in 1962. In this model, a neuron takes weighted sum of inputs and sends the output if sum is greater than any adjustable threshold value. The input in a perceptron are $x_1, x_2,..., x_n$ and connection weights are $w_1, w_2,..., w_n$. If the presence of some feature xi tends to cause the perceptron to fre, the weight wi will be positive and if the feature xi inhibits the perceptron, the weight wi will be negative.

Neural networks take a different approach to solve a problem as compared to conventional computers. Conventional computer follows a set of instructions in order to solve a problem and requires specific steps to solve a problem. This restricts the problem-solving capability of conventional computers to problems whose solutions are already known. Neural network's ability to learn by example makes it a very flexible and powerful concept. Furthermore, there is no need to devise an algorithm in order to perform a specific task, that is,

there is no need to understand the internal mechanisms of a problem.

Traditional AI and neural networks are generally considered for solving different types of problems. These two approaches appear to be different but currently the strengths of both the concepts are merged together to develop a system. The neural networks have been applied successfully in the AI field to speech recognition, image analysis and adaptive control, and to construct software agents (like video games) or autonomous robots.

24.7.1 Benefits of Neural Networks

Either humans or other computer techniques can use neural networks, with their ability to derive sense from complicated or imprecise data, to extract patterns and detect trends that are too complex to be noticed. Some of the important benefits of neural networks are:

- **Ease of Use:** Neural networks learn by example. The neural network user gathers representative data and then invokes training algorithms to automatically learn the structure of the data, thus reducing complexity.
- **Adaptive Learning:** An ability to learn how to do tasks based on the data given for training provides these networks an edge over the present systems.
- **Self-Organization:** A neural network can create its own organization or representation of the information it receives during learning time.
- **Real-Time Operation:** Computations are carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.

24.7.2 Applications of Neural Networks

Neural networks are applicable in areas where a relationship between the predictor variables (independents, inputs) and predicted variables (dependents, outputs) exists. Neural networks have been used in a wide range of applications, the majority of which include areas associated with problems in pattern recognition and control theory. A few examples of

problems to which neural network analysis have been applied successfully are:

- **Medicine:** A variety of health-related indices (like respiration rate) can be monitored. Neural networks have been used to recognize the predictive pattern so that the appropriate treatment can be prescribed.
- **Stock Market:** Neural networks are used by many technical analysts to make predictions about stock prices based upon a number of factors such as past performance of stocks and various economic indicators that are available.
- **Monitoring Machinery:** Neural networks can be instrumental in minimizing costs by bringing down costs of maintaining machines.
- **Engine Management:** The neural network controls various parameters within which the engine functions, in order to achieve a particular goal such as minimizing fuel consumption.

The prospect for neural networks is excellent. With the increase in sophistication of the underlying theory and range of applicability of the techniques, the research in this field is certain to persist and bring new strength into related fields.

24.8 AI LANGUAGES

AI programs are written in specialized languages designed specially for AI. The two AI languages commonly used are PROLOG (PROgramming LOGic) and LISP (LISt Processing). Both languages have features, which make them suitable for AI programming. These include support for list processing, pattern matching, and examining programming. The basic differences between AI and conventional programming are illustrated in Table 24.2.

Table 24.2 Comparing AI languages with Conventional Languages

Feature	AL Languages	Conventional Languages
Processing Type	Symbolic	Numeric
Technique Used	Heuristics search	Algorithmic search
Solution Steps	Implicit	Explicit and precise
Answer Sought	Satisfactory	Optimal
Involves	Large knowledge base	Large database
Knowledge	Imprecise	Precise
Process	Inferential	Repetitive

LISP was developed in the 1950s and was regarded as the base of functional programming. It is based on function definitions and remains a useful, powerful, and widely used programming language till today. LISP uses the list as its fundamental representation for function definitions and provides a wide range of inbuilt functions for manipulating lists. The list allows complex symbol structures (representing AI knowledge) to be easily manipulated. LISP is still a highly popular language used for the development of knowledge-based systems, although it has obtained significant competition from Java and C++, particularly in the development of systems that involve some numerical computation.

PROLOG is another language used for AI programming. It is presented with a set of facts, rules, and goals, and it attempts to prove that the goals become true when the set of rules are applied to facts. PROLOG is considered as a programming language derived from logic and theorem-proving. In particular, it is based on first-order predicate calculus. Writing simple PROLOG programs is similar to writing statements in predicate calculus, and “running” them involves setting queries for PROLOG.

LET US SUMMARIZE

1. John McCarthy coined AI in 1956 at the Massachusetts Institute of Technology (MIT). According to him, AI is the science and engineering of making intelligent machines, especially intelligent computer programs.
2. Intelligence is the ability to acquire, retrieve, and use knowledge in a meaningful way. It includes both raw and refined knowledge and the ability to memorize, recall facts, and express emotions.
3. Intelligence is exhibited through learning, reasoning, understanding, creativity, and intuition.
4. Alan Turing proposed a Turing test to prove the intelligence of a machine. The test involved three participants: two humans and a machine.
5. AI is currently being used in various domain areas such as game playing, speech recognition, understanding natural language, computer vision, expert systems, heuristic Classification, neural networks, and robotics.
6. Knowledge is defined as the ability to remember earlier learned material, which involves the recalling of a wide range of materials; all that is required is bringing the appropriate information to mind.
7. Knowledge-based systems make extensive use of knowledge to perform complex tasks. This, in turn, led researchers to develop knowledge representation schemes and techniques to manipulate knowledge.
8. *Knowledge acquisition* involves conversion of information gathered from various sources into knowledge.
9. *Knowledge organization* involves storage of knowledge in such a way that retrieval of information becomes easy. This is done using indexing.
10. *Knowledge manipulation* is the computational equivalent of reasoning. This requires a form of inference or deduction, using knowledge and inferring roles.
11. *Knowledge representation* is based on logic and mathematics and has an easily parsed grammar to ease processing in a system. Some knowledge representation schemes are FOPL, fuzzy logic, and associative network.
12. Searching techniques involves the use of an optimal state out of available states. For every problem, there are a number of alternatives available to find a solution. It can be categorized into two groups: *uninformed search and informed search*.
13. Uninformed search proceeds in a systematic way by exploring nodes in some predetermined order. The important search strategies used in this form include breadth-first search and depth-first search.
14. Informed search strategies are used to find optimal solution. It involves evaluation function, called heuristic functions. It includes *best-first search* and *hill-climbing search*.
15. NLP involves the interpretation, manipulation, or generation of human language by computer so that one can communicate with a computer as one communicates with another person. It involves morphological analysis, syntactic analysis, semantic analysis, discourse integration, and pragmatic analysis.
16. Expert system is an AI application that uses a knowledge base of human expertise to aid in solving problems. The degree of problem-solving is based on the quality of the data and rules

obtained from the human expert. This system simulates the judgment and behaviour of a human or an organization that has expert knowledge and experience in a particular field.

17. Expert systems are currently being used in applications such as medical diagnosis, equipment repair, financial and insurance planning, route scheduling for delivery vehicles, production control, and training.
18. Neural networks are computing systems, which imitate human brain through a network of highly, interconnected processing elements. These processes give these networks learning capabilities and enable them to recognize and to understand complex patterns.
19. Neural networks are applicable in areas where a relationship between the predictor variables and predicted variables exists such as medicine, stock market, monitoring machinery, and engine management.
20. AI languages are specialized languages designed specially for AI applications. These are PROLOG and LISP.

EXERCISES

Fill in the Blanks

1. Associative network is a direct graph consisting of.....and.....
2. The important search strategies used in informed search are.....and.....
3. One of the earliest expert system developed for blood diseases was.....
4. Intelligence includes both.....and.....knowledge.
5. Knowledge plays an important role in building.....systems.
6. Fuzzy logic is another type of.....scheme.
7. Neural network's ability to learn by example makes it very.....and.....
8. Technique used in AI language is.....search.
9. .....language uses the list as its fundamental representation for function definition.
10. ..... PROLOG is presented with a set of.....,....., and.....

Multiple-choice Questions

1. AI is concerned with making computers behave like.....
    1. Humans
    2. Robots
    3. Home appliances
    4. None of these
2. In Turing test the number of participants are.....
    1. One
    2. Three
    3. Four
    4. None of these
3. Knowledge-based systems have been developed for fields such as.....
    1. Financial markets

2. Diagnosing diseases
3. Scheduling airplanes
4. All of these

4. One way of organizing knowledge in memory is by using a method known as.....

    1. Chaining
    2. Matrix-based technique
    3. Indexing
    4. None of these

5. FOPL consists of some symbols and rules which are.....

    1. Predicate
    2. Functions
    3. Variables
    4. All of these

6. Fuzzy logic has been very successful in.....applications.

    1. Washing machines
    2. Air conditioners
    3. Dishwasher
    4. All of these

7. Which of the following search is not an uninformed search?

    1. Best-first search
    2. Depth-first search
    3. Breadth-first search
    4. All of these

8. NLP is mainly used in.....

    1. Knowledge acquisition
    2. Information retrieval
    3. Translation
    4. All of these

9. .....individuals are not involved in expert system.

    1. Mechanical engineer
    2. End-user.....
    3. Problem-solving expert
    4. Knowledge engineer.....

10. ..... Which of the following is an AI language?.....

    1. COBOL
    2. LISP.....
    3. HTML
    4. XML

State True or False

1. Intelligence is defined as the ability to acquire, retrieve, and use knowledge in a meaningful way.
2. In 1950, Alan Turing proposed a test to prove the intelligence of a human.
3. Thinking humanly, acting humanly, thinking rationally, acting rationally are four categories of AI.
4. “The Logic Theorist” developed by Allen Newell and Herbert Simon is considered as the first AI program.
5. Knowledge-based systems make extensive use of knowledge to perform complex tasks.
6. Matrix-based technique is one of the techniques available for knowledge acquisition.
7. FOPL is one of the most widely used representation scheme.
8. Search problems can be categorized into three groups.

9. NLP involves the interpretation, manipulation, or generation of human language by computer.
10. The general architecture of an expert system consists of two components: rule-base and inference engine.

Descriptive Questions

1. Describe various components of AI.
2. Is Turing test sufficient to judge whether a machine is intelligent? Explain.
3. Name the domain areas of AI and explain any four of them.
4. Describe associative networks by giving an example.
5. How NLP is used in developing AI system?
6. What is an expert system? Give its characteristics and limitations.
7. Explain neural networks along with its advantages.
8. Write short notes on:
    1. Breadth-first search
    2. Depth-first search
    3. Hill-climbing search
    4. Knowledge representation

ANSWERS

Fill in the Blanks

1. Nodes, Arcs
2. Hill climbing, Best-first
3. MYCIN
4. Raw, Refined
5. Intelligent
6. Representation
7. Flexible, Powerful
8. Heuristic
9. LISP
10. Facts, Rules, Goals

Multiple-choice Questions

1. (a)
2. (b)
3. (d)
4. (c)
5. (d)
6. (d)
7. (a)
8. (d)
9. (a)
10. (b)

State True or False

1. True
2. False

3. True
4. True
5. True
6. True
7. True
8. False
9. True
10. True

Acknowledgements

- Our publisher Pearson Education, their editorial team and panel reviewers for their valuable contributions toward content enrichment.
- Our technical and editorial consultants for devoting their precious time to improve the quality of the book.
- Our entire research and development team who have put in their sincere efforts to bring out a high-quality book.

Rohit Khurana
Founder and CEO
ITL ESL

Copyright © 2012 Dorling Kindersley (India) Pvt. Ltd.
Licensees of Pearson Education in South Asia

No part of this eBook may be used or reproduced in any manner whatsoever without the publisher's prior written consent.

This eBook may or may not include all assets that were part of the print version. The publisher reserves the right to remove any material in this eBook at any time.

ISBN 9788131760291

ePub ISBN XXXXXXXXXXXXX

Head Office: A-8(A), Sector 62, Knowledge Boulevard, 7th Floor, NOIDA 201 309, India

Registered Office: 11 Local Shopping Centre, Panchsheel Park, New Delhi 110 017, India

For instructor's resources
click http://pearsoned.co.in/ITLEducationSolutionsLimited/