



MADHYA PRADESH BHOJ(OPEN) UNIVERSITY BHOPAL

FOUNDATION OF IT AND MS OFFICE

BSDS-105

**FOUNDATION OF IT AND
MS OFFICE 2010**



मध्यप्रदेश भोज (मुक्त) विश्वविद्यालय – भोपाल
MADHYA PRADESH BHOJ (OPEN) UNIVERSITY - BHOPAL

Reviewer Committee

1. Dr. Sharad Gangele
Professor
RKDF University, Bhopal (M.P.)
2. Dr. Romsha Sharma
Professor
Sri Sathya Sai College for Women,
Bhopal (M.P.)
3. Dr. K. Mani Kandan Nair
Department of Computer Science
Makhanlal Chaturvedi National University of
Journalism & Communication, Bhopal (M.P.)

Advisory Committee

1. Dr. Jayant Sonwalkar
Hon'ble Vice Chancellor
Madhya Pradesh Bhoj (Open) University,
Bhopal (M.P.)
2. Dr. L.S. Solanki
Registrar
Madhya Pradesh Bhoj (Open) University,
Bhopal (M.P.)
3. Dr. Kishor John
Director
Madhya Pradesh Bhoj (Open) University,
Bhopal (M.P.)
4. Dr. Sharad Gangele
Professor
RKDF University, Bhopal (M.P.)
5. Dr. Romsha Sharma
Professor
Sri Sathya Sai College for Women,
Bhopal (M.P.)
6. Dr. K. Mani Kandan Nair
Department of Computer Science
Makhanlal Chaturvedi National University of
Journalism & Communication, Bhopal (M.P.)

COURSE WRITERS

Sanjay Saxena, Renowned author of over 130 books on Information Technology, New Delhi

Units (1.0-1.3, 1.5-1.11, 2.3.2-2.3.3, 3.0-3.2.1, 3.4)

Rajneesh Agrawal, Senior Scientist, Department of I.T., Government of India

Units (2.0-2.2.2, 4.3-4.3.3, 4.3.5, 4.5.6)

V.K. Govindan, Professor, Computer Engineering, Dept. of Computer Science and Engineering, N.I.T., Calicut, India

Unit (2.4-2.4.1, 2.4.3)

Manas Ghosh, Lecturer, R.C.C. Institute of Information Technology, Kolkata

Sudipta Pathak, System Consultant, Cognizant

Unit (4.0-4.2.7)

Preeti Khatri, Assistant Professor, Computer Science, S.O.I.T., I.M.S., Noida

Units (1.4, 2.3, 2.2.3, 2.4.2, 2.4.4-2.9, 3.3-3.3.4, 3.5-3.9, 4.3.4, 4.4-4.4.2, 4.5.1-4.5.3, 4.5.4, 4.5.5, 5)

Copyright © Reserved, Madhya Pradesh Bhoj (Open) University, Bhopal

All rights reserved. No part of this publication which is material protected by this copyright notice may be reproduced or transmitted or utilized or stored in any form or by any means now known or hereinafter invented, electronic, digital or mechanical, including photocopying, scanning, recording or by any information storage or retrieval system, without prior written permission from the Registrar, Madhya Pradesh Bhoj (Open) University, Bhopal.

Information contained in this book has been published by VIKAS® Publishing House Pvt. Ltd. and has been obtained by its Authors from sources believed to be reliable and are correct to the best of their knowledge. However, the Madhya Pradesh Bhoj (Open) University, Bhopal, Publisher and its Authors shall in no event be liable for any errors, omissions or damages arising out of use of this information and specifically disclaim any implied warranties or merchantability or fitness for any particular use.

Published by Registrar, MP Bhoj (Open) University, Bhopal in 2020



VIKAS® is the registered trademark of Vikas® Publishing House Pvt. Ltd.

VIKAS® PUBLISHING HOUSE PVT. LTD.

E-28, Sector-8, Noida - 201301 (UP)

Phone: 0120-4078900 • Fax: 0120-4078999

Regd. Office: A-27, 2nd Floor, Mohan Co-operative Industrial Estate, New Delhi 1100 44

• Website: www.vikaspublishing.com • Email: helpline@vikaspublishing.com

SYLLABI-BOOK MAPPING TABLE

Foundation of IT and MS Office 2010

Syllabi	Mapping in Book
<p>Unit – I Introduction to Computer Concepts: Elements of a Computer, History of Computers, Early Electronic Computers, Computer Generations, Classification of Computers. Basic Computer Architecture: The Central Processing Unit, Basic Computer Architecture. Input Output Devices: Keyed Input, Pointing Devices, touch-Sensitive Screens, Data Scanning devices, Voice Recognition Devices, Hard Copy Devices, Output Devices. Storage Media: Traditional Storage-Magnetic Tape, Today’s Preferred Storage-Magnetic disk.</p>	<p>Unit-1: Introduction to Computer (Pages 3-68)</p>
<p>Unit – II Data Representation: Decimal system, Binary System, Conversions, Fraction, Binary System to Other Number Systems, Octal Number System to Other Systems, Hexadecimal System to Other Systems. Software Concepts: Types of Software, software qualities, Programming Languages, Classification of Programming Languages, Classification of high level Languages, Object oriented languages. Operating Systems: History and Evolution, Function of an Operating System, Types of Operating Systems, Operating System as Resource Manager, BIOS.</p>	<p>Unit-2: Concepts of Data Representation and Software (Pages 69-128)</p>
<p>Unit – III DOS: Directories, Files, Wild cards, DOS Commands, Redirection option in DOS. Features of Unix Operating System: Multitasking Capability, Multi-user Capability, The Unix Connections, Components of The Unix Operating System, Unix Tools And Applications, The Unix Scheduler, The Unix File System, The Hierarchical File System of Unix, Basic Directory Commands, Basic File Commands, Shell Programming. Windows 95/98: Starting Windows 95/98, moving in Menus, Working with documents or files, Understanding the start Menus.</p>	<p>Unit-3: DOS, UNIX and Windows Operating System (Pages 129-188)</p>
<p>Unit – IV Database Management Systems: Data Processing, Database, Hierarchy of Database, why Database?, What is Database System?, DBMS Engine, Database Administrator, Database Users, Instances and Schemas, Data models. Communication Systems: The Basic Communication System, Data Transmission, Data Transmission Technique, Methods of Transmission, Error Detection, Communication Parameters, Modes of Transmission, Communication channels. Computer Virus: Spreading of Viruses, New Technology and Viruses, Methods to detect a Boot-sector Virus, Classification of Viruses, Effect of Viruses, Data Attacks, Session Hijacking, Viruses and Worms. Internet: Uniform Resource Locator (URI), Web Browsers, Domain Name Service (DNS), Internet Service Provider (ISP), Internet Requirements, Net Surfing, Internet Services, Internet Talking, Intranet.</p>	<p>Unit-4: DBMS and Communication Systems (Pages 189-278)</p>

Unit – V

Windows 2010: Introducing Windows 2010 Professional, New Ways to Communicate, File Systems.

Microsoft Word 2010: Standard Toolbars, Page Setup, Editing of word Document, View menu, Insert Menu, Formatting a Document, Paragraphs Formatting.

Microsoft Excel 2010: How to Start Microsoft Excel 2010, Working with Toolbars, File Menu, Page Setup, Editing in Excel-2010, Working with View Menu, Inserting Elements in Microsoft Excel, Insert chart, Formatting, Tools of Microsoft Excel, Data Management.

Microsoft Access 2010: Features of Access, Creating a Table, Creating a Relationship, Types of Functions, Introduction to Queries, Forms, Reports.

Microsoft PowerPoint 2010: Starting PowerPoint 2010, Formatting, Inserting Text, Drawing Toolbar, Slide Show.

MS Outlook 2010: Managing multiple mail and news accounts, Formatting message text, Secure Messages, Reading Messages, Organizing mail messages.

Unit-5: Windows, MS Office
and Outlook 2010
(Pages 279-453)

CONTENTS

INTRODUCTION	1
UNIT 1 INTRODUCTION TO COMPUTER	3-68
1.0 Introduction	
1.1 Objectives	
1.2 Introduction to Computer Concepts	
1.2.1 Elements of a Computer	
1.2.2 History of Computers	
1.2.3 Generations of Computer	
1.2.4 Classification of Computers	
1.3 Central Processing Unit (CPU)	
1.4 Basic Computer Architecture	
1.5 Input and Output Devices	
1.5.1 Keyboard Devices	
1.5.2 Point and Draw Devices	
1.5.3 Touch Screen	
1.5.4 Scanning Devices	
1.5.5 Voice Recognition Devices	
1.5.6 Output Devices	
1.5.7 Hard Copy Devices	
1.6 Storage Media	
1.6.1 Traditional Storage-Magnetic Tapes	
1.6.2 Magnetic Disks	
1.7 Answers to ‘Check Your Progress’	
1.8 Summary	
1.9 Key Terms	
1.10 Self-Assessment Questions and Exercises	
1.11 Further Reading	
UNIT 2 CONCEPTS OF DATA REPRESENTATION AND SOFTWARE	69-128
2.0 Introduction	
2.1 Objectives	
2.2 Data Representation	
2.2.1 Decimal Number System	
2.3 Software Concepts	
2.3.1 Software Qualities	
2.3.2 Types of Software	
2.3.3 Programming Language	
2.3.4 Object Oriented Language	
2.4 History and Evolution of Operating Systems	
2.4.1 Functions of OS	
2.4.2 Operating System as Resource Manager	
2.4.3 Types of Operating System	
2.4.4 BIOS	
2.5 Answers to ‘Check Your Progress’	
2.6 Summary	
2.7 Key Terms	
2.8 Self-Assessment Questions and Exercises	
2.9 Further Reading	

UNIT 3 DOS, UNIX AND WINDOWS OPERATING SYSTEM

129-188

- 3.0 Introduction
- 3.1 Objectives
- 3.2 Disk Operating System (DOS)
 - 3.2.1 DOS Command Types—Internal and External
- 3.3 Unix Operating System
 - 3.3.1 Architecture of UNIX
 - 3.3.2 Features of UNIX
 - 3.3.3 Working with Commands
- 3.4 Windows 95/98
- 3.5 Answers to ‘Check Your Progress’
- 3.6 Summary
- 3.7 Key Terms
- 3.8 Self-Assessment Questions and Exercises
- 3.9 Further Reading

UNIT 4 DBMS AND COMMUNICATION SYSTEMS

189-278

- 4.0 Introduction
- 4.1 Objectives
- 4.2 Database Management Systems
 - 4.2.1 Data Processing
 - 4.2.2 Users of a DBMS
 - 4.2.3 Benefits and Limitations of the Database Approach
 - 4.2.4 Structure of a DBMS
 - 4.2.5 Database Administration
 - 4.2.6 DBMS Engine
 - 4.2.7 Data Models
- 4.3 Basic Communication System
 - 4.3.1 Channel Characteristics
 - 4.3.2 Modes of Transmission
 - 4.3.3 Data Transmission Systems/Technique
 - 4.3.4 Communication Parameters
 - 4.3.5 Error Detection
- 4.4 Computer Virus
 - 4.4.1 New Technology and Viruses
 - 4.4.2 Classification of Viruses
 - 4.4.3 Data Attacks
- 4.5 Internet
 - 4.5.1 Uniform Resource Locator (URL)
 - 4.5.2 Domain Name System (DNS)
 - 4.5.3 Web Browsers
 - 4.5.4 Internet Service Provider (ISP)
 - 4.5.5 Internet Services
 - 4.5.6 Internet Requirements
 - 4.5.7 Intranet
- 4.6 Answers to ‘Check Your Progress’
- 4.7 Summary
- 4.8 Key Terms
- 4.9 Self-Assessment Questions and Exercises
- 4.10 Further Reading

- 5.0 Introduction
- 5.1 Objectives
- 5.2 Windows 2010
 - 5.2.1 New Ways to Communicate (Online Services and Functionality)
- 5.3 Microsoft Word 2010
 - 5.3.1 Editing of Word Document
 - 5.3.2 Formatting Paragraph and Text in MS Word
 - 5.3.3 Working with Tables, Columns and Sections
 - 5.3.4 Bullets and Numbering
 - 5.3.5 Shortcut Keys for Various Tasks
- 5.4 Microsoft Excel 2010
 - 5.4.1 Creating, Opening and Working with Excel Documents
 - 5.4.2 Working with Cells
 - 5.4.3 Sorting and Filtering Information
 - 5.4.4 Working with Arithmetic Operators and Text Functions
 - 5.4.5 Working with Charts
 - 5.4.6 Modifying and Merging Rows and Columns
- 5.5 Microsoft Access 2010
 - 5.5.1 Features of MS Access 2010
 - 5.5.2 Creating a Table
 - 5.5.3 Form Design in MS Access 2010
 - 5.5.4 Query Design and Running in MS Access 2010
 - 5.5.5 Report Design in MS Access 2010
 - 5.5.6 Using MS Access 2010 for Business Decisions
- 5.6 Microsoft PowerPoint 2010
 - 5.6.1 Creating Presentation and Slides
 - 5.6.2 Applying Animation and Transition Effects
 - 5.6.3 Adding Sound and Video in Slides
 - 5.6.4 Working with Tables and Pictures
 - 5.6.5 Working with SmartArt
- 5.7 Microsoft Outlook 2010
 - 5.7.1 MS Outlook and E-mail Accounts
 - 5.7.2 Importing and Exporting Data
 - 5.7.3 Managing E-Mail in Outlook
 - 5.7.4 Managing E-mail Accounts
- 5.8 Answers to 'Check Your Progress'
- 5.9 Summary
- 5.10 Key Terms
- 5.11 Self-Assessment Questions and Exercises
- 5.12 Further Reading



INTRODUCTION

Computers have brought about major changes in all spheres of life. In fact, it is extremely difficult to imagine the world without computers. The fact that computers have made a big impact on many aspects of our lives can hardly be questioned. They have opened up an entire world of knowledge and information that is readily accessible. The computers need software to do specialized tasks. The software used in a computer system is grouped into applications software, system software and utility software. Today, we use computer systems in the organizations to automate the work with the help of specific software.

The information systems (IS) field has historically limited itself to the design, use and implementation of information technology (IT) in organizations. IT is the study, design, development, implementation, support or management of infotech. 'Information technology' is a general term that describes any technology that helps to produce, manipulate, store, communicate, and/or disseminate information. Today the core concerns of the IS field are not limited to the requirements of the IS department and the organization it is serving. IT is so widespread today that its advantages and threats are felt beyond organizational or disciplinary boundaries. IT is infused into the foundation of society. The impact of IT in one area is instantly

felt by other areas. An IT-literate individual not only does well at work, but also becomes helpful in IT-related tasks performed at the social and political fields. As is generally considered, IT and IS are not the same thing. IT is assumed to be the product (the artifact) created by human activity. IS, on the other hand, is the field that studies phenomena surrounding IT. Another important factor that distinguishes the IS field from other fields of study is its focus on IT. As a field, IS would not have emerged to its present state without IT.

The term 'Office Automation' refers to all tools and methods that are applied to the activities which make it possible to process written, visual and audio data in a computer aided manner. They process data, store information, solve complex mathematical problems, track inventory and even control temperature and lighting in office buildings. All this can be done with the help of various computer office application programs/software, such as Desktop Publishing, PageMaker, Microsoft Word, Microsoft Excel and Microsoft PowerPoint. Desktop Publishing, also termed as DTP, combines a personal computer and What You See IS What You Get (WYSIWYG) page layout software to create publication documents.

This book, *Foundation of IT and MS Office 2010*, follows the SIM format wherein each Unit begins with an Introduction to the topic followed by an outline of the 'Objectives'. The detailed content is then presented in a simple and an organized manner, interspersed with 'Check Your Progress' questions to test the understanding of the students. A 'Summary' along with a list of 'Key Terms' and a set of 'Self-Assessment Questions and Exercises' is also provided at the end of each unit for effective recapitulation.

NOTES



UNIT 1 INTRODUCTION TO COMPUTER

NOTES

Structure

- 1.0 Introduction
- 1.1 Objectives
- 1.2 Introduction to Computer Concepts
 - 1.2.1 Elements of a Computer
 - 1.2.2 History of Computers
 - 1.2.3 Generations of Computer
 - 1.2.4 Classification of Computers
- 1.3 Central Processing Unit (CPU)
- 1.4 Basic Computer Architecture
- 1.5 Input and Output Devices
 - 1.5.1 Keyboard Devices
 - 1.5.2 Point and Draw Devices
 - 1.5.3 Touch Screen
 - 1.5.4 Scanning Devices
 - 1.5.5 Voice Recognition Devices
 - 1.5.6 Output Devices
 - 1.5.7 Hard Copy Devices
- 1.6 Storage Media
 - 1.6.1 Traditional Storage-Magnetic Tapes
 - 1.6.2 Magnetic Disks
- 1.7 Answers to ‘Check Your Progress’
- 1.8 Summary
- 1.9 Key Terms
- 1.10 Self-Assessment Questions and Exercises
- 1.11 Further Reading

1.0 INTRODUCTION

A computer may be defined as a device that can operate upon information or data. The information is processed based on a set of instructions provided to generate the output. The increasing popularity of the computer has proved that it is a powerful and useful tool. The size, shape, cost and performance of computers have changed over the years, but the basic logical structure has not. A user of the computer can simultaneously play games while a Word document is being printed. History of computers and the various generations of its evolution and classification of computers such as PCs, laptops, PDAs, etc., and also get to know about micro, mini, mainframe and supercomputers.

Input devices are used in order to transfer data to the memory of the computer. The Arithmetic Logic Unit (ALU) is responsible for calculations, to which this data from the memory is stored. The different input as well as output devices used. Input devices can be of various types, such as keyboard, mouse, scanners, webcams, touch screen monitors, etc. It is significant for computers to connect with the external world, i.e., receive and send data and information. An

NOTES

electromechanical device that accepts data from the computer and translates it into a form that can be understood by the external world is known as an output device. The processed data, stored in the memory of the computer, is sent to an output unit, which then transforms the internal representation of data into a form that can be read by the users.

In this unit, you will study about the basic concepts of computer, elements of a computer, history of computers, early electronic computers, computer generations, classification of computers, central processing unit, basic computer architecture, input/output devices input/output devices and storage media.

1.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the basic concepts of computers
- Explain the elements of computer
- Discuss the history of computers
- Explain the generations of computers
- Understand the classification of computers
- Define central processing unit
- Understand the basic computer architecture
- Explain the input/output devices
- Discuss about the keyboard input and pointing devices
- Describe the touch-sensitive screens and data scanning devices
- Elaborate on the voice recognition and hard copy devices
- Understand the traditional storage-magnetic tape
- Explain the preferred storage-magnetic disk

1.2 INTRODUCTION TO COMPUTER CONCEPTS

Computers have undergone great transformation over the past decade; however, the basic logical structure remains the same. A computer primarily constitutes of three integral components, viz. input devices, central processing unit (CPU) and output devices. The CPU constitutes of the main memory, the arithmetic logic unit and the control unit.

Apart from these three basic components, computers have secondary storage devices known as auxiliary storage or backing storage that store data and instructions on a long-term basis.

The following are the primary functions of a computer:

- **Inputting:** The process in which the user puts in a set of commands to process data into the computer system.

- **Storing:** The process of recording data and information so that it can be retrieved for use whenever required.
- **Processing:** This process implies performing arithmetic or logical operations on data to convert them into useful information. Arithmetic operations include addition, subtraction, multiplication and division, and logical operations include comparisons such as equal to, less than and greater than, etc.
- **Outputting:** This is the process of providing results to the user. These can be in the form of visual display and/or printed reports.
- **Controlling:** This refers to directing the sequence and the manner in which all the previous functions are carried out.

A detailed description of the components that perform these tasks is as follows.

1.2.1 Elements of a Computer

1. Input Unit

Programs and data are required to be present in a computer system before any operation can be performed. A program denotes the set of instructions which the computer has to carry out, and data is the information on which these instructions are to operate. If the task is to rearrange a list of telephone subscribers in alphabetical order, the sequence of instructions that will guide the computer through this operation is the program, while the list of names to be sorted is the data.

The Input Unit is responsible for transferring data and instructions from the external environment into the computer system. Instructions and data enter the input unit through the particular input device used (keyboard, scanner, card reader, etc.). These instructions and data are then converted into binary codes (computer-acceptable form) and sent to the computer system for further processing.

2. Central Processing Unit

The central processing unit is known as the brain of the computer. It is a blend of the control unit, the ALU and the primary memory that are described as follows:

- **Main memory (primary storage):** The main memory or the primary storage of the computer system is responsible for storing all the instructions and data. The data is then transferred to the Arithmetic Logical Unit (ALU) for processing. After this, the final output is again stored back in the primary storage, until it is further sent to the output device.

The primary storage also temporarily stores any intermediate result generated by the ALU. So data and instructions move frequently between the ALU and the primary storage before the processing is complete. It should be noted that no processing occurs within the primary storage.

- **Arithmetic logic unit:** In addition to the basic four arithmetic operations, viz. addition, subtraction, multiplication and division, the ALU also performs logic comparison operations including equal to, lesser than or greater than.

NOTES

NOTES

- **Control unit:** The function of the control unit is to ensure that according to the stored instructions, the right operation is done on the right data at the right time. The control unit receives instructions and commands from the programs in the primary memory, processes them and ensures that the commands are executed in the desired order by all the other units of the computer system. In effect, the control unit is comparable to the central nervous system of the human body.

3. Output Unit

Computers understand, process data and return the output in a binary form. The basic function of the output unit is to convert these results into a human-readable form before providing the output through various output devices, such as terminals and printers.

The storage capacity of the primary memory of the computer is limited. Often, it is necessary to store large amounts of data. So, additional memory, called secondary storage or auxiliary memory, is used in most computer systems.

Secondary storage is storage other than the primary storage. These are peripheral devices connected to and controlled by the computer to allow permanent storage of data and programs. Usually, hardware devices like magnetic tapes and magnetic disks fall in this category.

Advantages and Disadvantages of Computers

The increasing popularity of the computer has proved that it is a powerful and useful tool. Its usefulness is due to its following features:

- **Speed:** Computers are very fast. They can process millions of instructions every second. The speed is related to the amount of data it processes and the time it takes to complete the processing task.
- **Storage:** Computers can store vast amounts of information in the form of files, which can be recalled at any time. These files help in easy and speedy retrieval of information. This type of storage is known as electronic storage system.
- **Accuracy:** In addition to being fast, computers are also accurate. The degree of accuracy for a particular computer depends upon its design. Most errors in computers are not of a technical nature and are human. Usually, programmers are responsible for these errors.
- **Diligence:** Computers can perform any complicated task accurately without making any error. Computers do not suffer from carelessness, boredom or tiredness. Moreover, their efficiency does not decrease with age.
- **Versatility:** Computers perform various tasks depending upon the instructions given to them and their hardware characteristics. They are capable of performing any task, provided the task is reduced to a series of logical steps. A computer can be used to prepare a Word document and in between called to search for another document that is stored in its memory. It can perform both tasks simultaneously.

- **No IQ:** Computers do not have their own intelligence and their IQ (Intelligence Quotient) is zero. Hence, the user can and has to decide what tasks a computer should perform.
- **No feelings:** Computers have no feelings because they are machines. They cannot make judgements as they process on the basis of a set of instructions, called programs, provided by the users.

Though computers can do better than human beings in terms of accuracy, speed and memory, there are certain disadvantages of computer systems as they depend on human beings for their operations and functions. The following are some of the disadvantages of computers:

- They depend on human beings who program them for efficient, accurate and fast functioning.
- Computers do not have their own intelligence and thus cannot think intelligently or work independently like human beings.
- They follow instructions given by programs or by users.
- They can neither take decisions nor can correct wrong instructions.
- Programmers or users maintain and update them.
- As with many other modern appliances, computers also need electric power to run.

1.2.2 History of Computers

The first mechanical adding machine was invented by Blaise Pascal in 1642. Later, in 1671, Baron Gottfried Wilhelm von Leibniz of Germany invented the first calculator. Around this time, Herman Hollerith developed the concept of punched cards, which were extensively used as an input medium in mechanical adding machines.

Charles Babbage, a 19th century professor at Cambridge University, is considered the father of the modern digital computer. During this period, mathematical and statistical tables were prepared by a group of clerks. However, utmost care and precaution could not eliminate human errors.

In 1842, Babbage came up with a new idea of the Analytical Engine, which was meant to be completely automatic. This machine was capable of performing basic arithmetic functions. However, these machines were difficult to manufacture because the precision required to manufacture them was not available at that time.

The following is a brief description of the various generations of computers.

- **Mark I Computer (1937–44):** This was the first fully automatic calculating machine designed by Howard A. Aiken, the design of which was based on the technique of punching card machinery. In this technique, both mechanical and electronic components were used.
- **Atanasoff-Berry Computer (1939–42):** This computer was developed by Dr John Atanasoff to solve certain mathematical equations. It used forty-five vacuum tubes for internal logic and capacitors for storage.

NOTES

NOTES

- **ENIAC (1943–46):** The Electronic Numerical Integrator and Computer (ENIAC) was the first electronic computer developed for military requirements and was used for many years to solve ballistic problems.
- **EDVAC (1946–52):** One of the drawbacks of ENIAC was that its programs were wired on boards, which made it difficult to change them. To overcome the drawbacks of ENIAC, the Electronic Discrete Variable Automatic Computer (EDVAC) was designed. The basic idea behind this concept was that sequences of instructions could be stored in the memory of the computer for automatically directing the flow of operations.
- **EDSAC (1947–49):** Professor Maurice Wilkes developed the Electronic Delay Storage Automatic Calculator (EDSAC), by which addition and multiplication operations could be accomplished.
- **UNIVAC I (1951):** The UNIVersal Automatic Computer (UNIVAC) was the first digital computer to be installed in the Census Bureau in 1951 and was used continuously for 10 years. In 1952, International Business Machines (IBM) introduced the 701 commercial computers. These computers could be used for scientific and business purposes.

1.2.3 Generations of Computer

The history of computer development can be divided into different phases which are often referred to as generations of computing devices. ‘Generation’ in computer terminology is a ‘step’ in technology. Each generation of computers is characterized by a major technological development that fundamentally changes the way computers operate, resulting in increasingly smaller, cheaper, more powerful, efficient and reliable devices with decreasing energy consumption and lesser generation of heat.

Originally, the term ‘generation’ was used to distinguish between varying hardware technologies, but nowadays, it includes both hardware and software.

The following are the characteristics of each generation of computers:

1. First-generation (1940–1956): Vacuum Tubes

The first computers used vacuum tubes in their electronic circuits and magnetic drums for memory. A vacuum tube was a delicate glass device that used filaments as a source of electrons and could control and amplify electronic signals. Figure 1.1 displays a vacuum tube.



Fig. 1.1 A Vacuum Tube

These computers could perform computations in milliseconds but were enormous in size, occupying almost an entire room. They were very expensive to operate and in addition to using a great deal of electricity, generated a lot of heat resulting in malfunctioning.

First-generation computers relied on machine language (binary-coded programs) to perform operations and could solve only one problem at a time. Input was based on punch cards and paper tape, and output was displayed on printouts.

Early computers like ENIAC, EDVAC, UNIVAC I all can be classified as first-generation computers.

2. Second-generation (1956–1963): Transistors

Transistors developed in 1947 replaced vacuum tubes in the second-generation computers. The transistor was far superior compared to vacuum tube, making computers smaller, faster, cheaper, more energy-efficient and more reliable than their first-generation predecessors. Although transistors also generated a great deal of heat that could damage the computer, it was a great improvement over the vacuum tube. Second-generation computers still relied on punched cards for input and printouts for output.

The cryptic binary machine language was followed by the symbolic or assembly language that allowed programmers to specify instructions in words. High-level programming languages like COBOL and FORTRAN were also being developed at this time.

These were also the first computers that stored their instructions in the memory, which advanced from magnetic drum to magnetic core technology. The first computers of this generation were specifically developed for the atomic energy industry.

3. Third-generation (1964–1971): Integrated Circuits

Transistors were clearly an improvement over the vacuum tube but still generated a lot of heat resulting in computer damage.

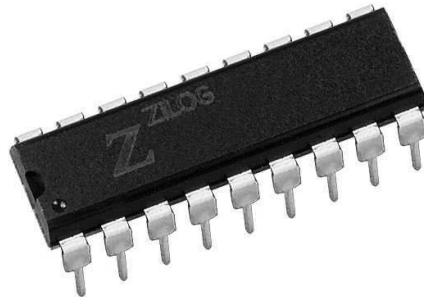


Fig. 1.2 An IC Chip

The development of integrated circuit (see Figure 1.2) by Jack Kilby in 1958, an engineer with Texas Instruments, was the greatest achievement of the third-generation of computers.

Instead of punched cards and printouts, users interacted with third-generation computers through devices like keyboards and monitors. They also interfaced with an operating system that allowed the device to run many different applications at one time with a central program that monitored the memory.

Now, the computers became accessible to the masses because they were substantially smaller and cheaper than their predecessors.

NOTES

4. Fourth-generation (1971–Present): Microprocessors

NOTES

Large Scale Integration (LSI) were developed which could fit hundreds of components onto a single chip. By 1980s, Very Large Scale Integration (VLSI) squeezed thousands of components onto a single chip. Ultra Large Scale Integration (ULSI) increased that number to millions.

The ability to fit so much processing capability in an area so small, helped to reduce the size and price of the computers. It also increased its power, efficiency and reliability.

Initially, the IC technology was used only for constructing the processor, but it was soon discovered that the same technology could also be used for the construction of memory. The first memory chip was constructed in 1970 and could hold 256 bits. Figure 1.2 displays an IC chip.

As more and more components were fabricated on a single chip, fewer and fewer chips were needed to construct the processor. The Intel 4004 chip, developed in 1971, located all the components of the computer — from Central Processing Unit and Memory to Input/Output controls — on a single chip. This was the first microprocessor. Figure 1.3 displays the intel pentium microprocessor chip.

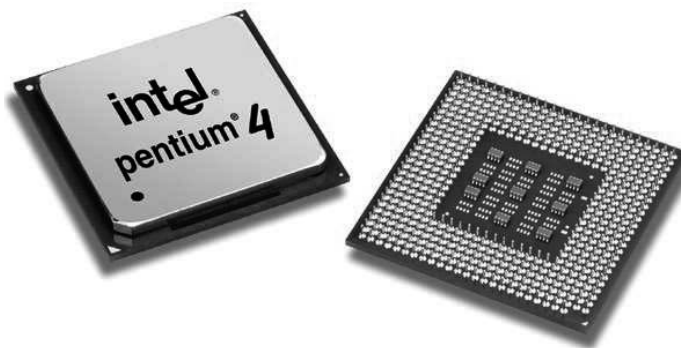


Fig. 1.3 The Intel Pentium Microprocessor Chip

IBM introduced its first computer in 1981 for the home users, and in 1984 Apple introduced the Macintosh. Microprocessors also advanced from the realm of desktop computers to advanced technologies and many areas of life as more and more everyday devices began to use microprocessors.

As computers increased in computing power, it was possible to connect them together to form networks, which eventually led to the development of the Internet. Fourth-generation computers also marked the development of GUIs, the mouse and various handheld devices.

5. Fifth-generation (Present and Beyond): Artificial Intelligence

The fifth-generation computers are being developed using the technology of artificial intelligence; for instance, voice recognition systems. Parallel processing and supercomputers have lead to the further development of artificial intelligence. In the future, quantum computation and molecular technology will tremendously

transform computers. The fifth-generation aims at creating devices that respond to input in natural language and are capable of learning and self organization.

Table 1.1 provides a list of various computer generations.

Table 1.1 Generation of Computers

Generation	Time	Hardware	Software	Features	Examples
I	1942-1955	Vacuum Tubes	Machine Language (Binary Language)	High-speed electronic switching device; memory type was electromagnetic; bulky in size; generated a large amount of heat; frequent technical faults; required constant maintenance; used for scientific purposes; air-conditioning required	ENIAC, EDVAC, EDSAC, UNIVAC I
II	1955-1964	Transistors	High-level languages FORTRAN, COBOL, ALGOL, SNOBOL	Better electronic switching devices than vacuum tubes; made of germanium semiconductors; memory type was magnetic cores; powerful and more reliable; easy to handle; much smaller than vacuum tubes; generated less heat as compared to vacuum tubes; used for business and industries for commercial data processing; air-conditioning required	Livermore Atomic Research Computer (LARC), IBM
III	1964-1975	Integrated Circuits (ICs) made up of transistors, resistors and capacitors fixed on single silicon chip	High-level languages PL/1, PASCAL, BASIC, VISUAL BASIC, C, C++, C#, Java	ICs were smaller than transistors; consumed less power; dissipated less heat as compared to transistors; more reliable and faster than earlier generations; capable of performing about 1 million instructions per second; large storage capacity; used for both scientific and commercial purposes; air-conditioning required	Mainframe, Minicomputers
IV	1975-1989	Microprocessor made up of Large Scale Integration Circuits (LSI) and Very Large Scale Integration Circuits (VLSI)	Advanced Java (J2EE, JDO, JavaBeans), PHP, HTML, XML, SQL	Microprocessor had control on logical instructions and memory; semiconductor memories; personal computers were assembled; used in LAN and WAN to connect multiple computers at a time; used graphical user interface; smaller, more reliable and cheaper than third-generation computers; larger primary and secondary storage memories; had Computer Supported Cooperative Working (CSCW); air-conditioning not required	Personal Computers (PCs), LAN, WAN, CSCW
V	1989-Present	Ultra Scale Large Integration (USLI), Optical Disks	Artificial Intelligence, PROLOG, OPS5, Mercury	PCs were assembled – portable and non-portable, powerful desktop PCs and workstations; less prone to hardware failure; user-friendly features – Internet, e-mailing; air-conditioning not required	Portable PCs, Palmtop Computers, Laptop

NOTES

1.2.4 Classification of Computers

Computers can be classified on the basis of their size, processing speed and cost. The various types of computers are:

- Personal computers
- Workstations
- Notebook/laptop computers
- Tablet PC
- PDA
- Mainframe computers
- Supercomputers

NOTES

Analog Computers

These types of computers are involved in industrial process controls and measure physical quantities, like pressure, temperature, etc. These computers do not use binary digits but use electrical signals to provide output with electrical resistance, voltage, etc. The memory of these computers is not much and they can be used only for specific calculations but their speed is more than digital computers.

These electrical properties allow calculations to be performed in real time or even faster at the speed of light. The main mathematical operations it applies include summation, inversion, exponentiation, logarithm, integration, differentiation, multiplication and division.

Digital Computers

These types of computers are primarily involved in data processing and problem-solving for specific programs. In digital computers, data is stored as digits (numbers) and processes. Letters, words, symbols and complete texts are digitally represented, i.e., using only two digits 0 and 1. Digital computers have a lot of memory for storing data.

Digital computers constitute input-output devices, main memory, control unit and arithmetic logic unit. Data is processed with logical circuits, also known as digital circuits. All the circuits processing data inside a computer function in an extremely synchronized mode; which is further controlled using a steady oscillator acting as the computer's 'clock'. Hence, the digital computers operate on very high speed and are able to perform trillions of logical or arithmetic operations per second to provide quick solution to problems, which is not possible for a human being to do manually.

Hybrid Computers

Hybrid computers are a mixture of digital and analog computers. A hybrid computer uses the best characteristics of digital and analog computers. It helps the user to process both continuous and discrete data. Hybrid computers are generally used for weather forecasting and industrial process control.

The digital component basically functions as a controller to provide logical operations, whereas the analog component provides solutions of differential equations. Remember that the hybrid computers are different from hybrid systems. The hybrid system is a digital computer equipped with an analog-to-digital converter for input and a digital-to-analog converter for output. The term 'hybrid computer' represents a combination of different digital technologies to process specific applications with the help of various specific processor technologies.

General Purpose Computers

Workstations are high-end, general-purpose computers designed to meet the computing needs of engineers, architects and other professionals who need computers with greater processing power, larger storage and better graphic display facilities. These are commonly used for Computer-Aided Design (CAD) and for

multimedia applications such as creating special audio-visual effects for television programmes and movies. A workstation looks like a PC and can be used by only one person at a time. The characteristics of a workstation, which are often used to differentiate it from a PC, are as follows:

- **Display facility:** Most workstations have a large-screen monitor (21 inches or more) capable of displaying high-resolution graphics as compared to PCs, which have a small-screen monitor (19 inches or less).
- **Storage capacity:** Workstations have a larger main memory than PCs, which have only a few hundred MB of main memory. The hard disk capacity of workstations is also more than that of PCs.
- **Processing power:** The processing power of workstations is several times greater than that of PCs.
- **Operating system:** PCs can run any of the five major operating systems—MS-DOS, MS-Windows, Windows-NT, Linux and Unix—but all workstations generally run the Unix operating system or a variation of it such as AIX (used in IBM workstations), Solaris (used in SUN workstations) and HPUX (used in HP workstations).
- **Processor design:** PCs normally use CPUs based on the Complex Instruction Set Computer (CISC) technology, whereas workstation CPUs are based on the Reduced Instruction Set Computer (RISC) technology.

NOTES

Special Purpose Computer

A special purpose computer is a digital or an analog computer specifically designed to perform desired specific task. These are high-performance computing systems with special hardware architecture, which is dedicated to solve a specific problem. This is performed with the help of specially programmed FPGA chips or custom VLSI chips. They are used for special applications, for example, astrophysics computations, GRAPE-6 (for astrophysics and molecular dynamics), Hydra (for playing chess), MDGRAPE-3 (for protein structure computations), etc.

Micro, Mini, Mainframe and Supercomputers

These are as follows:

(i) Microcomputers

Microcomputers are developed from advanced computer technology. They are commonly used at home, classroom and in the workplace. Microcomputers are called home computers, personal computers, laptops, personal digital assistants, etc. They are powerful and easy to operate. In recent years, computers were made portable and affordable. The major characteristics of a microcomputer are as follows:

- Microcomputers are capable of performing data processing jobs and solving numerical programs. Microcomputers work rapidly like minicomputers.
- Microcomputers have reasonable memory capacity which can be measured in megabytes.

NOTES

- Microcomputers are reasonably priced. Varieties of microcomputers are available in the market which can be as per the requirement of smaller business companies and educational institutions.
- Processing speed of microcomputers is measured in megahertz. A microcomputer running at 90MHz works approximately at 90 MIPS.
- Microcomputers have drives for floppy disk, compact disk and hard disks.
- Only one user can operate a microcomputer at a time.
- Microcomputers are usually dedicated to one job. Millions of people use microcomputers to increase their personal productivity.
- Useful accessory tools, such as clock, calendar, calculator, daily schedule reminders, scratch pads, etc., are available in a microcomputer.
- Laptop computers, also called notebook computers, are microcomputers. They use the battery power source. Laptop computers have a keyboard, mouse, floppy disc drive, CD drive, hard disk drive and monitor. Laptop computers are expensive in comparison to personal computers.

(ii) Minicomputers

Minicomputers are a cheaper version of mainframe computers. The processing power and cost of a minicomputer are less than that of the mainframe. The minicomputers have big memory sizes and faster processing speed compared to the microcomputer. Minicomputers are also called workgroup systems because they are well suited to the requirements of the minor workgroups within an organization. The major characteristics of a minicomputer are as follows:

- Minicomputers have great problem solving capabilities.
- Minicomputers have reasonable memory capacity which can be measured in megabytes or gigabytes.
- Minicomputers have quick processing speeds and operating systems facilitated with multitasking and network capabilities.
- Minicomputers have drives for floppy disk, magnetic tape, compact disk, hard disks, etc.
- Minicomputers can serve as network servers.
- Minicomputers are used as a substitute of one mainframe by big organizations.

(iii) Mainframe computers

Mainframe computers are generally used for handling the needs of information processing of organizations like banks, insurance companies, hospitals and railways. This type of system is placed in a central location with several user terminals connected to it. The user terminals act as access stations and may be located in the same building (refer to Figure 1.4).



Fig. 1.4 Mainframe Computer

Mainframe computers are bigger and more expensive than workstations. They look like a row of large file cabinets and need a large room with closely monitored humidity and temperature levels. A mainframe system of lower configuration is often referred to as a minicomputer system. The various components of a mainframe computer are as follows:

- **Host, front-end and back-end computers:** A mainframe system consists of several computers, such as a host computer that carries out most of the computations and has direct control over all other computers. The front-end portion is used for handling communications to and from all the user terminals connected to the mainframe computer. The back-end portion is used to handle data input/output operations. The host computer and other computers are located in the systems room, to which entry is restricted to system administrators and maintenance staff only.
- **Consoles:** Console terminals are directly connected to the host computer and are mainly used by the system administrator to perform certain administrative tasks like installing new software on the system, taking system backups and changing the configuration of the system.
- **Storage devices:** A mainframe computer has several magnetic disk drives directly connected to the back-end computer. The host computer, via the back-end computer gets all data from these magnetic disks. In addition, a mainframe computer also has a few tape drives and a magnetic tape library (located in the systems room) for restoration and backup of data. The tape drives are present in the users' room, so that users' tapes can be used for input and output.
- **User terminals:** User terminals are used to access the required stations, which may be present at different locations. Since mainframe computers support multiprogramming with time-sharing, they can run different operating systems for multiple users at the same time.
- **Output devices:** A mainframe computer has several output devices like printers and plotters, connected to the back-end computer, so that these devices can be used for taking outputs by the users.

NOTES

NOTES

(iv) Supercomputers

Supercomputers are the most powerful and expensive computers available today. They are mainly used for processing scientific applications that involve tasks with highly complex calculations and for solving problems with mechanical physics, such as weather forecasting and climate research systems, nuclear weapon simulation and simulation of automated aircrafts. Supercomputers are mainly used by military organizations, major research and development centres, universities and chemical laboratories.

Supercomputers use multiprocessing and parallel processing technologies to solve complex problems quickly. They use multiprocessors, which help the user to divide a complex problem into smaller problems. A parallel program is written in a manner that can break up the original problem into smaller modules. Supercomputers also support multiprogramming, which allows simultaneous access to the computer by multiple users. Some of the manufacturers of supercomputers are IBM, Silicon Graphics, Fujitsu and Intel.

Personal Computers

A PC is a small single-user microprocessor-based computer that sits on your desktop and is generally used at homes, offices, and schools. As the name implies, PCs were mainly designed to meet the personal computing needs of individuals. Personal computers are used for preparing normal text documents, spreadsheets with predefined calculations and business analysis charts, database management systems, accounting systems and also for designing office stationary, banners, bills and handouts.

The configuration varies from one PC to another depending on its usage. However, it consists of a CPU or system unit, a monitor, a keyboard and a mouse. It has a main circuit board or motherboard (consisting of the CPU and the memory), hard disk storage, floppy disk drive, CD-ROM drive and some special add-on cards (like Network Interface Card) and ports for connecting peripheral devices like printers.

PCs are available in two models—desktop and tower. In the desktop model, the monitor is positioned on top of the system unit, whereas in the tower model the system unit is designed to stand by the side of the monitor or even on the floor to save desktop space. Due to this feature, the tower model is more popular.

Some popular operating systems for PCs are MS-DOS, MS-Windows, Windows-NT, Linux and Unix. Most of these operating systems can perform many functions at the same time which eases operation and saves time when a user has to switch between two or more applications while performing a job. Some leading PC manufacturers are IBM, Apple, Compaq, Dell, Toshiba and Siemens.

Types of personal computers

Different types of personal computers are as follows:

1. Notebook/laptop Computers

Notebook computers are battery-operated personal computers. Smaller than the size of a briefcase, these are portable computers and can be used in places like

libraries, in meetings or even while travelling. Popularly known as laptop computers, or simply laptops. Notebook computers are usually more expensive as compared to desktop computers though they have almost the same functions, but since they are sleeker and portable they have a complex design and are more difficult to manufacture. These computers have large storage space and other peripherals such as serial port, PC card, modem or network interface card, CD-ROM drive and printer. They can also be connected to a network to download data from other computers or to the Internet. A notebook computer has a keyboard, a flat screen with liquid crystal color (LCD) display (see Figure 1.5), and can also have a trackball and a pointing stick.



Fig. 1.5 A Laptop Computer

A notebook computer uses the MS-DOS or WINDOWS operating system. The data processing capability of a notebook computer is as good as an ordinary PC because both use the same type of processor, such as an Intel Pentium processor. However, a notebook computer generally has lesser hard disk storage than a PC.

2. Tablet PC

Tablet PC is a mobile computer that looks like a notebook or a small writing slate but uses a stylus pen or your finger tip to write on the touch screen. It saves whatever you scribble on the screen with the pen, in the same way as you have written it. The same picture can then be converted to text with the help of a HR (hand recognition) software.

3. PDA

A Personal Digital Assistant (PDA) is a small, palm sized, hand-held computer which has a small color touch screen with audio and video features. They are nowadays used as smart phones, web enabled palmtop computers, portable media players or gaming devices.

Most PDAs today typically have a touch screen for data entry, a data storage/memory card, bluetooth, Wi-Fi or an infrared connectivity and can be used to access the Internet and other networks.

NOTES

1.3 CENTRAL PROCESSING UNIT (CPU)

NOTES

The central processing unit (CPU) is the most important component of a digital computer that interprets the instructions, and processes the data contained in computer programs. CPU works as the brain of the computer and performs most of the calculations. It is also referred as processor and is the most important component of a computer. For large computers, a CPU may require one or more printed circuit boards (PCBs), but in the case of PCs it comes in the form of a single chip called a microprocessor. PCB is a board that contains the circuitry used to connect the components of a PC. Figure 1.6 shows the block diagram of a CPU.

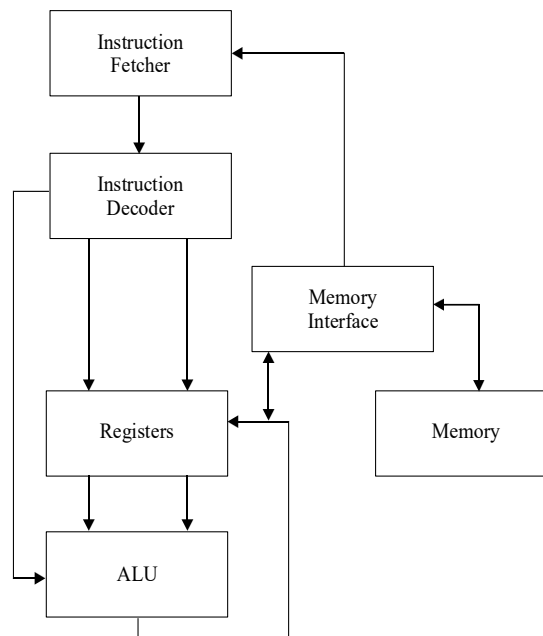


Fig. 1.6 Block Diagram of a CPU

The two main components of a CPU are:

- The Arithmetic Logic Unit (ALU), which performs arithmetic and logical operations.
- The Control Unit, which extracts instructions from the memory and converts them into a form that the computer can understand and executes them. In this process, it takes the help of the ALU whenever necessary.

Components of the Central Processing Unit

Execution of programs is the main function of the computer. The program to be executed is a set of instructions that is stored in the computer's memory. The central processing unit (CPU) executes the instructions of the program to complete a task. Also, all the major calculations and comparisons are carried out inside the CPU. Additionally, the CPU is responsible for activating and controlling the operations of various units of the computer system. It activates the peripherals to perform input or output function.

The CPU is made up of three major components (as seen in Figure 1.7): the register set (associated with the main memory) that stores the intermediate data during the execution of instructions, the arithmetic logic unit (ALU) which performs the required micro-operations for executing the instructions, and the control unit that supervises the transfer of information among the registers and instructs the ALU as to which operation to perform.

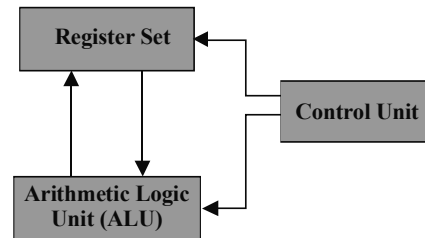


Fig. 1.7 Major Components of the CPU

Control Unit

The control unit plays an important role in the functioning of the CPU itself and transfer of data/information from a device to the CPU or vice versa. It does not perform the actual processing of the data, but manages and coordinates the entire computer system including the input and output devices. It retrieves and interprets the instructions from the program stored in the main memory, and issues signals that cause other units of the system to execute them.

It does this through some special purpose registers and a decoder. The special purpose register called the *instruction register* holds the instruction to be currently executed, and the *program control register* holds the next instruction to be executed. The decoder interprets the meaning of each instruction supported by the CPU. Each instruction is also accompanied by a microcode, i.e., the basic directions to tell the CPU how to execute the instruction.

Arithmetic Logic Unit (ALU)

The ALU provides arithmetic and logic operations. This means that when the control unit encounters an instruction that involves an arithmetic operation (add, subtract, multiply, divide) or a logic operation (equal to, less than, greater than), it passes the control to the ALU. The ALU has the necessary circuitry to carryout these arithmetic and logic operations.

As an example, a comparison of two numbers (a logical operation) may require the control unit to load the two numbers in the requisite registers and then pass on the execution of the 'compare' function to the ALU. Figure 1.8 shows the basic structure of a CPU.

NOTES

NOTES

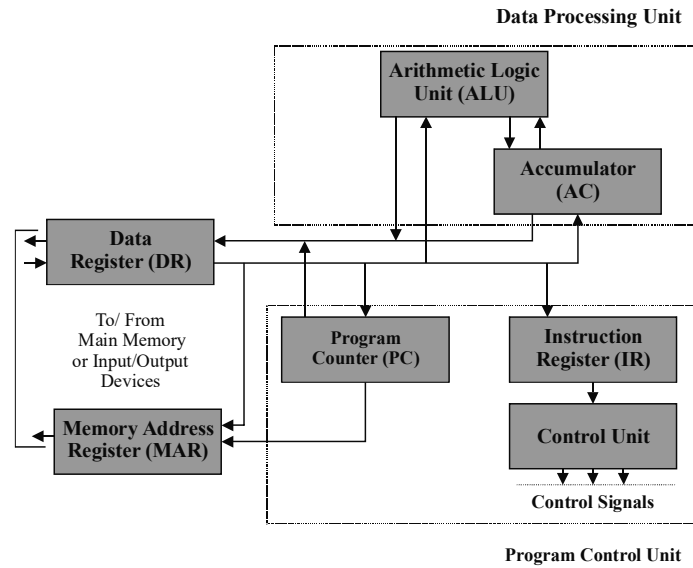


Fig. 1.8 Basic Structure of a CPU

Instruction Set

The primary function of the processing unit in the computer is to interpret the instructions given in a program and carry out the instructions. Processors are designed to interpret a specified number of instruction codes. Each instruction code is a string of binary digits. All processors have input/output instructions, arithmetic instructions, logic instructions, branch instructions and instructions to manipulate characters. The number and the type of instructions differ from processor to processor. The list of specific instructions supported by the CPU is termed as its *instruction set*. An instruction in the computer should specify the following:

- The task or operation to be carried out by the processor. This is termed as the opcode.
- The address(es) in memory of the operand(s) on which the data processing is to be performed.
- The address in the memory that may store the results of the data-processing operation performed by the instruction.
- The address in the memory for the next instruction to be fetched and executed. The next instruction which is executed is normally the next instruction following the current instruction in the memory. Therefore, no explicit reference to the next instruction is provided.

Instruction Representation

An instruction is divided into a number of fields and is represented as a sequence of bits. Each field constitutes an element of the instruction. A layout of an instruction is termed as the instruction format.

In most instruction sets, many instruction formats are used. An instruction is first read into an instruction register (IR), then the CPU extracts and processes the

required operands on the basis of references made on the instruction fields, and then decodes it. Since the binary representation of the instruction is difficult to comprehend, it is seldom used for representation. Instead, a symbolic representation is used.

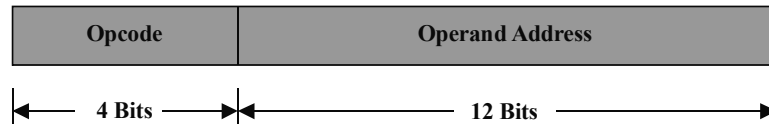


Fig. 1.9 A Sample Instruction Format

Examples of typical instructions:

Instruction	Interpretation	Number of Addresses
ADD A,B,C	Operation $A = B + C$ is executed	3
ADD A,B	$A = A + B$. In this case the original content of operand location is lost	2
ADD A	$AC = AC + A$. Here A is added to the accumulator	1

Typically, CPUs manufactured by different manufacturers have different instruction sets. This is why machine language programs developed for a particular CPU do not run on a computer with a different CPU (having a different instruction set).

Registers

The basic task performed by the CPU is instruction execution. Each instruction is executed using several small operations called micro-operations. Therefore, the basic issues relating to a CPU can be expressed as:

- It should be as fast as possible.
- The capacity of the main memory needed by the CPU is very large.

To understand further, let us define two relevant terms:

- **Cycle time of the CPU:** It is the time taken by the CPU to execute a well-defined shortest micro-operation.
- **Memory cycle time:** It is the speed at which the memory can be accessed by the CPU.

It has been found that the memory cycle time is approximately one-to-ten times higher than the CPU cycle time, that is why temporary storage is provided within the CPU in the form of CPU registers. The CPU registers are termed as fast memory and can be accessed almost instantaneously.

Further, the number of bits a register can store at a time is called the length of the register. Most CPUs sold today have 32-bit or 64-bit registers. The size of the register is also called the word size and indicates the amount of data that a CPU can process at a time. Thus, the bigger the word size, the faster the computer can process data.

NOTES

NOTES

The number of registers varies among computers, but typical registers found in most computers are:

- **Memory Address Register (MAR):** Specifies the address of memory location from which data is to be accessed (in case of read operation) or to which data is to be stored (in case of write operation).
- **Memory Buffer Register (MBR):** Receives data from the memory (in case of read operation) or contains the data to be written in the memory (in case of write operation).
- **Program Counter (PC):** Keeps track of the instruction that is to be executed next, after the execution of the current instruction.
- **Accumulator (AC):** Interacts with the ALU and stores the input or the output operand. This register, therefore, holds the initial data to be operated upon, the intermediate results and the final results of processing operations.
- **Instruction Register (IR):** Instructions are loaded in the IR before their execution, i.e., the instruction register holds the current instruction that is being executed.

The simplest form of instruction processing can be defined as a two-step process:

1. The CPU reads (fetches) instructions (codes) from the memory one at a time.
2. It executes or performs the operation specified by this instruction.

The instruction is fetched using the program counter (PC) that keeps track of the next instruction to be fetched. Normally, the next instruction in the sequence is fetched, as programs are executed in sequence. The fetched instruction is in the form of binary code and is loaded into an instruction register (IR) in the CPU. The CPU then interprets the instruction and performs the required action. In general, these actions can be divided into the following categories:

- Data transfer from CPU to memory or memory to CPU or from CPU to I/O or I/O to CPU.
- Data processing an arithmetic or a logic operation may be performed on the data by the CPU.
- **Sequence control:** This action is typically required for altering the sequence of execution. For example, if an instruction from location 50 specifies that the next instruction to be fetched should be from location 100, then the program counter will need to be modified to contain the location 100 (which otherwise would have contained 51). Execution of an instruction may involve any combination of these actions.

Processor Speed

The term processor has replaced the term central processing unit (CPU). The processors in personal computers or those embedded in small devices are often called **microprocessors**.

The speed at which the processor executes commands is called the processor speed or the clock speed. Every computer contains an internal clock (known as

the system clock) that regulates the rate at which the instructions are executed and synchronizes the various computer components. The processor requires a fixed number of clock cycles (electric pulses per second) to execute each instruction. Clock cycles are required to fetch, decode, and execute a single program instruction. Thus, the shorter the clock cycle, the faster is the processor.

In a computer, clock speed, therefore refers to the number of pulses per second generated by an oscillator that sets the tempo for the processor. It is usually measured in MHz (**Megahertz** - Million pulses per second) or GHz (**Gigahertz** - Billion pulses per second).

Computer clock speed has been roughly doubling every year. The Intel 8088, common in computers manufactured in the nineties, ran at 4.77 MHz. Today's personal computers run at clock speeds of a 100–1000 MHz and some even exceed one gigahertz. Although the processing speed in personal computers is measured in terms of megahertz, the processing speed of mini computers and mainframe systems is measured in terms of MIPS (millions of instructions per second) or BIPS (billions of instructions per second). This is because personal computers generally employ a single microprocessor chip as their CPU, while other classes of computers employ multiple processors to speed up their overall performance. Thus, a minicomputer with a speed of 500 MIPS is capable of executing 500 million instructions per second.

Clock speed is a measure of computer 'power,' but it is not always directly proportional to the performance level. If you double the speed of the clock, leaving all other hardware unchanged, you will not necessarily double the processing speed. The type of microprocessor, the bus architecture, and the nature of the instruction set all make a difference. In some applications, the amount of random access memory (RAM) is also important.

Instruction Cycle

When a CPU is given an instruction in machine language, this instruction is fetched from the memory by the CPU to execute. Instruction cycle (or fetch-and-execute cycle) refers to the time period, during which one instruction is fetched and executed by the CPU. An instruction cycle has four stages:

1. **Fetch:** In this step, an instruction is loaded from the memory into the CPU registers. All the instructions must be fetched before they can be executed.
2. **Decode:** In this step, the control unit decodes the instructions.
2. **Derive effective address of the instruction:** In this step, if the instruction has an indirect address then the effective address of the instruction from memory is read.
4. **Execute:** In this step, the action represented by the instruction is performed.

Steps 1 and 2, taken together, are called fetch cycle and these steps are the same for each instruction. Steps 3 and 4 are called execute cycle and these steps change with each instruction.

MIPS

Million Instructions Per Second (MIPS) is an old unit to measure a CPU's speed and power. It gives the number of machine instructions that a CPU can execute in

NOTES

one second. For example, the speed of IBM PC/XT computer is $\frac{1}{4}$ MIPS, while the speed of Pentium-based PCs is 100 MIPS. MIPS is used for referring to CPU speeds only.

NOTES

CPU Bus

The processor or the CPU bus is the set of wires that is used to carry information to and from the processor to other components of the computer. The flow of information through the processor bus is controlled by the motherboard of the computer. The processor bus connects the CPU to the main memory of the computer. The CPU bus is further divided into data bus and memory bus. The data bus transfers the actual data whereas the address bus transfers the information about where the data should go within the computer.

CISC vs RISC

The term 'CISC' (complex instruction set computer or computing) refers to computers designed with a full set of computer instructions that were intended to provide the required capabilities in the most efficient way. Later, it was discovered that by reducing the full set to only the most frequently used instructions, the computer could perform more work in a shorter amount of time for most applications. Since this was called reduced instruction set computing (RISC), there was now a need for a name for full-set instruction computers, hence the term CISC.

RISC (reduced instruction set computer or computing) is a microprocessor designed to perform a smaller number of types of computer instructions so that it can operate at a higher speed. Since each instruction type that a computer must perform requires additional transistors and circuitry, a larger list would mean the microprocessor becoming more complicated and slower in operation.

John Cocke of IBM Research in Yorktown, New York, originated the RISC concept in 1974 when he stated that about 20 per cent of the instructions in a computer performs 80 per cent of the work.

While today a number of microchips use the RISC concept — typically found in Sun Microsystems SPARC microprocessors and other medium-sized computers like IBM RS/6000 — Intel's Pentium microprocessors are CISC microprocessors.

1.4 BASIC COMPUTER ARCHITECTURE

There are three basic functions of a computer, which are as follows:

Data processing: A computer must be able to process data.

Data storage: A computer must store data. Even if data is supplied to a computer on the fly, for processing and producing the result immediately, the computer must be able to store that data temporarily. Apart from short-term data storage, it is equally important for a computer to perform a long-term storage function to store different files.

Data movement: A computer must be able to move data between itself and the

outside world. The computer-operating environment consists of devices that serve as data sources or destinations. When data is received from or delivered to a machine that is directly linked to a computer, the process is known as input/output and the devices are referred to as input/output devices. When data moves over longer distances, to or from a remote machine, the process is known as data communication.

Functional Units of a Computer

In its simplest form, a computer consists of five functionally independent components: input, output, memory, arithmetic and logic unit, and control unit as shown in Figure 1.10.

A computer accepts information in the form of a program and data through its input unit, which can be an electromechanical device such as a keyboard or from other computers over digital communication lines. The information received by the computer is either stored in the memory for later reference or used immediately by the arithmetic and logic unit (ALU) for performing the desired operations. Finally, the processed information in the form of results is displayed through an output unit. The control unit controls all activities taking place inside the computer. The ALU unit along with the control unit are collectively known as the central processing unit (CPU) or processor; and the input and output units are collectively known as the Input–Output (I/O) unit.

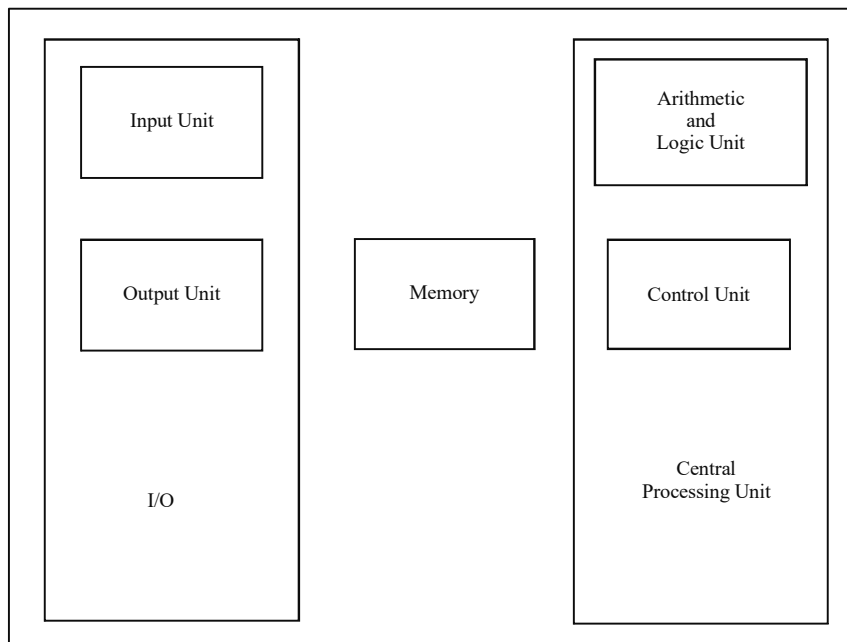


Fig. 1.10 Functional Units of a Computer

- **Input Unit:** A computer accepts input in coded form through an input unit. The keyboard is an input devices. Whenever a key is pressed, the binary code of the corresponding letter or digit is transferred to the memory unit or processor. Other types of input devices are mouse, punch card, joysticks, etc.
- **Memory Unit:** The task of the memory unit is to safely store programs as

NOTES

NOTES

well as input, output and intermediate data. The two different classes of memory are primary and secondary storage. Primary storage is a very fast memory and contains a large number of semiconductor cells capable of storing one bit of information. A group (of fixed size) of these cells is referred to as words and the number of bits in each word is referred to as word length, which typically ranges from 16 to 64 bits. When the memory is accessed, usually one word of data is read or written.

- **Processor Unit:** The processor unit performs arithmetic and other data processing tasks as specified by a program.
- **Control Unit:** It oversees the flow of data among the other units. The control unit retrieves the instructions from a program (one by one), which are safely kept in the memory. For each instruction, the control unit tells the processor to execute the operation marked by the instruction. The control unit supervises the program instructions and the processor manipulates the data as specified by the programs.
- **Output Unit:** The output unit receives the result of the computation, which is displayed on the screen or printed on paper using a printer.

Bus Structures

A shared communication path consisting of one or more connection lines is known as a bus and the transfer of data through this bus is known as bus transfer.

When data is read from or stored in memory, it is referred to as memory transfer.

The functional components of a computer must be connected in order to make a system operational. The CPU communicates with the other components via a bus. A bus is a set of wires that acts as a shared—but common—data path to connect multiple subsystems within the computer system. It consists of multiple lines, allowing the parallel movement of bits. Buses are low cost but very versatile and help connect devices with each other as well as the system. At any given point in time, only one device (be it a register, the ALU, memory or some other component) may use the bus. However, this sharing often results in a communications bottleneck. The speed of the bus is affected by its length as well as by the number of devices sharing it.

- **Data Bus:** It is used for the transmission of data. Data lines and the number of bits in a word are similar.
- **Address Bus:** It carries the address of the main memory location from where data can be accessed.
- **Control Bus:** It is used to indicate the direction of data transfer and to coordinate the timing of events during the transfer.

A digital computer consists of many processor registers and the transfer of information from one register to another is often required. Hence, paths must be provided so that such transfer operations can take place. Figure 1.11 shows the transfer among three registers R1, R2 and R3 through six data paths.

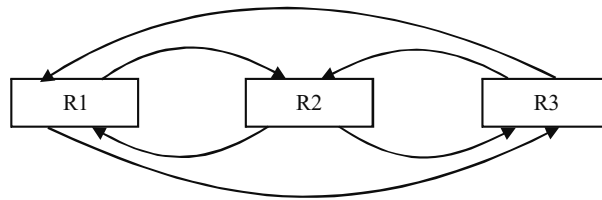


Fig. 1.11 Transfer among Three Registers

If different lines are used involving each register, the number of wires will increase considerably. Hence, a pair of common lines, one line for each bit of the register, is used for the transfer. This set of common lines through which binary data is transferred, one at a time, among registers is known as a *bus*. A common bus system can be constructed with the help of multiplexers and decoders. The multiplexer selects the source register whose binary information is then placed on the bus and the decoder selects one destination register to transfer the information to, from the bus. The construction of a bus system for four registers is shown in Figure 1.12. Two multiplexers have been used, one for the low-order significant bit and one for the high-order significant bit. If the register is of n bits, n multiplexers are required to produce n bus lines. These n lines in the bus are connected to n inputs of all the registers.

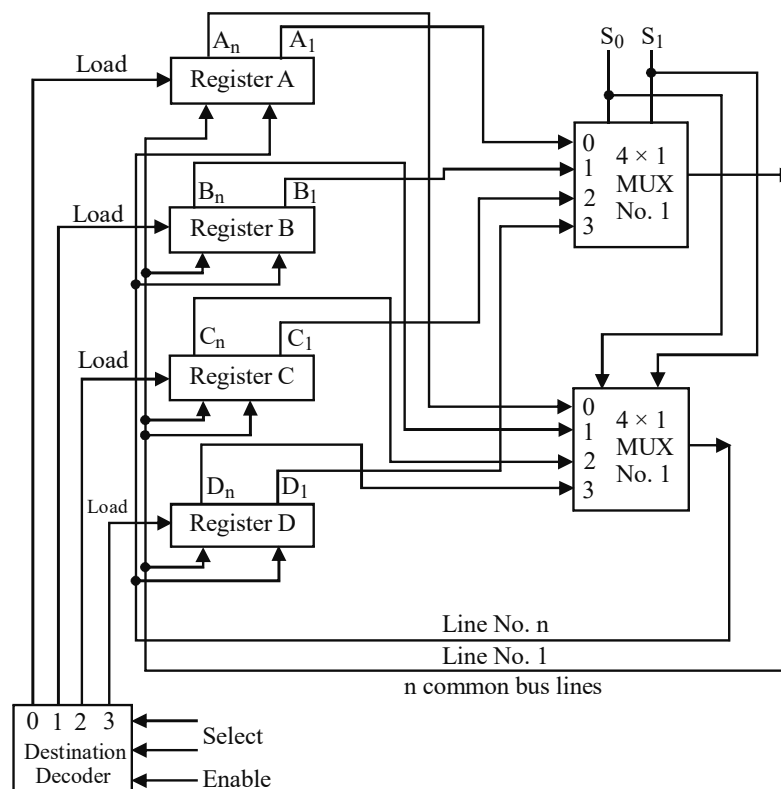


Fig. 1.12 Bus System for Four Registers

S_1 and S_0 are selection lines connected to selection inputs of all n multiplexers. The selection lines choose n bits of one register and transfers these to the common bus n lines. When S_1 and $S_0 = 00$, the 0 data inputs of all n multiplexers are selected and cause the n bits from register A to transfer to the n -line common bus, since the output of this register is connected to 0 data inputs of each multiplexer. Similarly,

NOTES

when S_1 and $S_0 = 01$, the content of register B is transferred into the n -line common bus and so on. The register that is selected for the four possible binary values is shown in Table 1.2.

NOTES

Table 1.2 Function Table for Bus in Figure 1.12

S_1	S_0	Register Selected
0	0	A
0	1	B
1	0	C
1	1	D

The shifting of data from a bus to one of the targeted registers is done with the help of the load control of that register. The load control of the particular register is activated by the outputs of the decoder when enabled. If the decoder is not enabled, no information from the bus will be transferred to the register although the multiplexers place the information of the source register onto the bus.

In general, for registers of n bits, n multiplexers are needed to construct a bus of n lines. The size of a multiplexer depends on the number of registers in the system. If there are K registers, the multiplexer's size will be $K \times 1$ since it multiplexes K data lines. To take an example, a general bus of 16 registers of 16 bits each needs 16 multiplexers of size 16×1 . Four selection lines are required. Also, the size of the destination decoder will be 4×16 .

Consider the statement

$$C \leftarrow B$$

The control function that enables this transfer must select register B as the source and register C as the destination registers. The content of register B is located on the bus and the content of the bus is then transferred to register C by starting its load control input.

Bus Organization

A bidirectional bus for carrying data between two units is called a data bus.

A unidirectional bus used to carry memory addresses is called memory bus.

The manner in which different buses are connected to form a common bus so that the CPU, memory and I/O devices can use the common bus, when required, is called bus organization.

A basic computer consists of a memory unit, a control unit and registers. There must be a path that can be used to transfer information between the memory and the registers or among registers. Using a common bus is the most efficient way of transferring information from source to destination in a system with multiple registers. Figure 1.13 shows the connection of eight registers and a 4096×16 memory unit of a common bus system. The eight registers are the Address Register (AR), Program Counter (PC), Data Register (DR), Accumulator (AC), Instruction Register (IR), Temporary Register (TR), Input Register (INPR) and Output Register (OUTR). Here, a 16-bit common bus has been used.

The outputs of seven registers and memory are linked to the common bus. The definite output chosen for the bus lines at any time is finalized by the binary value of the selected lines S_2 , S_1 and S_0 as shown in Table 1.3. The numbers along each output line shows the decimal equivalent of the required binary selection. When $S_2S_1S_0 = 011$, the 16-bit outputs of DR are placed on the bus lines.

The lines from the common bus are linked to the input of each register and the data inputs of the memory. The specific register whose LD (load) input is allowed gets the information from the bus. The memory gets the information from the bus when 'a write input' is allowed.

The memory puts its results on to the bus when the 'read input' is on and $S_2S_1S_0 = 111$.

The registers DR, AC, IR and TR are of 16- bits. Two registers, PC and AR, have 12 bits since they store addresses.

When the contents of AR and PC are placed on the bus, the four most significant bits are set to 0. When AR and PC receive data from the common bus, only the 12 least significant bits are transferred to the register.

The input registers INPR and OUTR have 8 bits; hence, they communicate only with the 8 least significant bits in the bus. INPR is connected to the bus for providing information. However, OUTR is connected to the bus only for receiving information from the bus. INPR receives a character from the I/O device, which is transferred to AC; and OUTR receives a character from AC and delivers it to an output device. No transfer takes place from OUTR to any of the other registers.

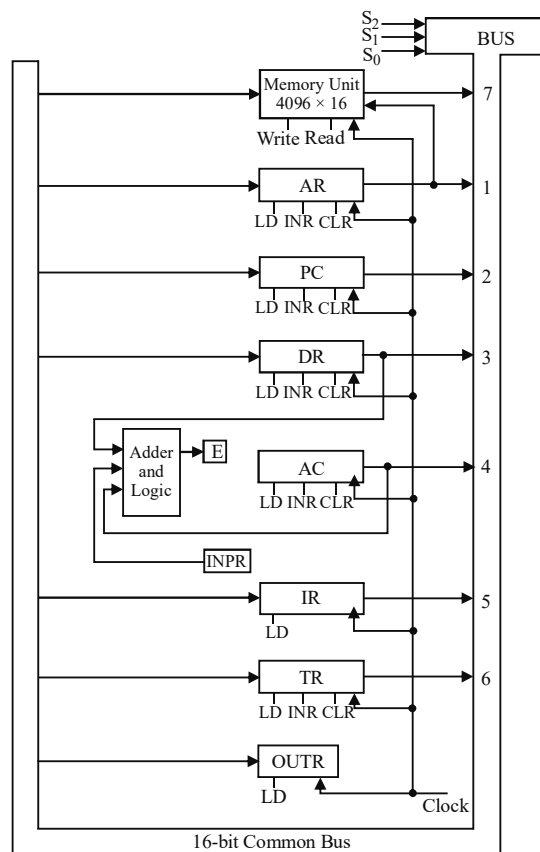


Fig. 1.13 Common Bus Organization

NOTES

NOTES

The 16-bit common bus receives information from six registers and the memory unit. In addition, the 16-bit common bus is linked to the inputs of six registers and the memory unit. Five registers have three control signals—LD (load), INR (increment) and CLR (clear). Two registers have only LD (load) control signal connected to the common bus. AR is also connected to the memory address. Thus, AR always specifies the memory address. During a memory write operation, the content of any register can be specified for the memory data and similarly during a memory read operation, any register except AC can receive data from the memory. The 16-bit AC receives inputs from the adder and logic circuit, which receives input from three registers. These three registers are 16-bit AC, 16-bit data register DR and 8-bit inputs, which come from input register INPR. The inputs from DR and AC are used for arithmetic and logic micro-operations. Table 1.3 shows the binary value of selection line $S_2S_1S_0$ that selects one of the registers.

For example, in order to transfer the contents of PC to AR (Address Register), the computer requires the following instructions:

- Set the selection variables $S_2S_1S_0 = 010$
- Transfer the contents of PC to the bus
- Enable LD input of AR
- Transfer contents of bus into AR

Table 1.3 Function Table

S_2	S_1	S_0	Register
0	0	1	AR
0	1	0	PC
0	1	1	DR
1	0	0	AC
1	0	1	IR
1	1	0	TR
1	1	1	Memory

Multiple Bus Organization

A two-bus structure used to connect the registers and the ALU of a processor is shown in Figure 1.14. All general-purpose registers are connected to both buses *A* and *B* to form a two-bus organization. The two operands required by the ALU are routed in one clock cycle; hence, the execution of instruction becomes faster since the ALU does not wait for the second operand, as is the case with single bus organization. Information passed on to the bus may be from general-purpose registers or special-purpose registers. In addition, the special-purpose registers are divided into two groups—one group at the left of the ALU connected to bus *A* and the other group is at the right of the ALU connected to bus *B*. The data from two special-purpose registers belonging to the same group cannot be transferred to the ALU at the same time.

The output of the ALU may be routed to either general-purpose registers or special-purpose registers. The ALU does not have any input buffer register and hence, both buses will be busy in carrying the operands during the binary operations. Therefore, the output of ALU is first stored in the output register. Transfer of the required operands and loading of the ALU output buffer register take place in one clock cycle. The content of the ALU output register is routed to the destination with the help of either bus *A* or bus *B* in the second clock cycle.

NOTES

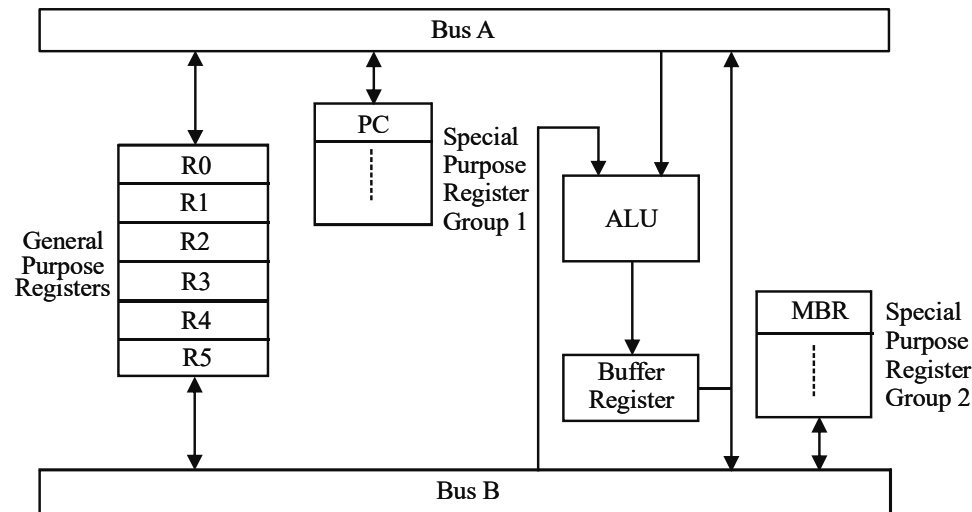


Fig. 1.14 Two-Bus Organization of the Data Path

The performance of a two-bus organization can be further improved by adding a third bus *C* at the output of ALU. The three-bus structure is shown in Figure 1.15. The addition of a third bus allows the system to perform operation, such as $R3 \leftarrow R1 + R2$ in one clock cycle as there are three separate buses in the system.

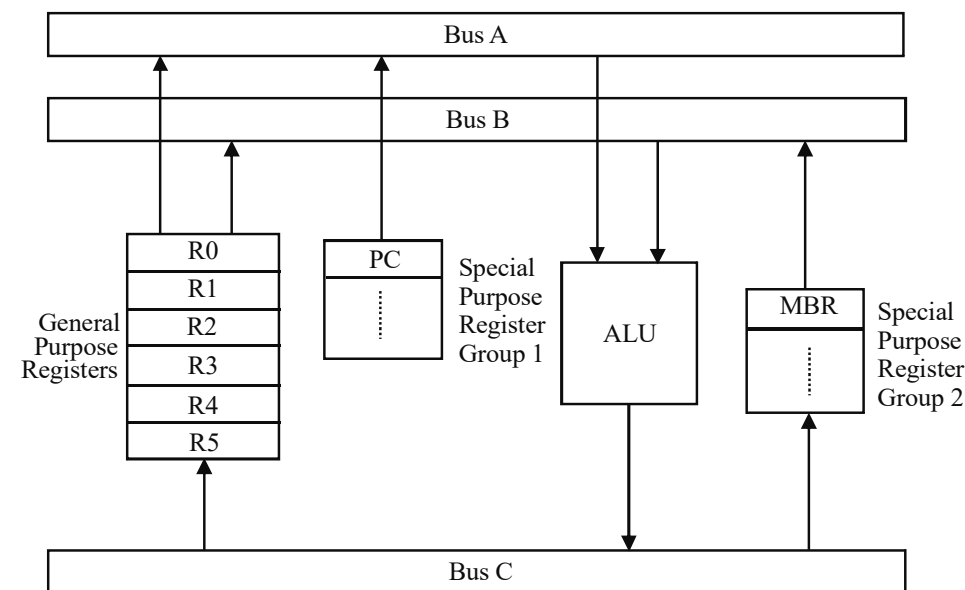


Fig. 1.15 Three-Bus Organization of the Data Path

Instructions and Instruction Sequencing

NOTES

An *instruction* is a command given to a computer to perform a specified operation on some given data. These instructions tell the CPU what to do. In other words, an instruction guides the CPU to perform work accordingly. The most common fields found in the instructions are the operation code and the operands. Each field specifies different information for the computer. The two important fields of an instruction are as follows:

- Op-code
- Operand

Thus,

$$\text{Instruction} = \text{Op-code} + \text{Operand}$$

Op-code (operation code) is an instruction field that specifies the particular operation to be performed by the instruction. Each operation has its unique op-code and may take several micro-operations to accomplish. MOV, ADD and SUB are examples of Intel 8086 op-codes.

Operand fields specify where to get the source and destination operands for the operation specified by the op-code. The source/destination of operands can be the memory or one of the general-purpose registers. The complete set of op-codes for a particular microprocessor defines the instruction set for that processor.

Instruction sequencing is the method by which instructions are selected for execution, i.e., the manner in which control of the processor is transferred from one instruction to another.

The simplest method of controlling the sequence of instruction execution is to have each instruction explicitly specify the address of the next instruction to be run. However, explicit inclusion of instruction addresses in all the instructions is disadvantageous as the instruction length increases. This results in increased cost of memory where the instructions are to be stored.

Instruction Execution

The sequence of operations performed by the CPU in processing an instruction is known as an instruction cycle. The time required to complete one instruction is called execution time.

To execute an instruction, the following three steps are required:

- *Fetch step*, during which a new instruction is read from the memory
- *Decode step*, during which the instruction is decoded
- *Execute step*, during which the operations specified by the instruction are executed

The instruction fetch operation is initiated by loading the contents of the program counter (PC) into the address register (AR) and it sends a read request to the memory. The contents of the PC is the address of the instruction to be run. The instruction read from the memory is then placed in the instruction register (IR) and the content of the PC is incremented so that it contains the address of the next

instruction in the program. After this, the instruction is decoded to determine the type of instruction that was just read. Finally, the instruction is executed to perform the operation specified by the instruction.

Let us consider an instruction, which adds the content of a memory location specified by register R_0 to the content of register R_2 and the result is to be stored in R_2 . The execution of this instruction is performed in the following steps:

- Step 1** Fetch and decode the instruction
- Step 2** Fetch the operand
- Step 3** Perform the operation (addition)
- Step 4** Store the result in R_2

NOTES

Check Your Progress

1. What are the three important parts of a computer system?
2. Give five advantages of using computers.
3. Who is considered as the father of modern digital computer?
4. What development was seen in the fourth generation microcomputers?
5. Give examples of third generation computers.
6. What is central processing unit?
7. Define the term bus transfer.

1.5 INPUT AND OUTPUT DEVICES

It is significant for computers to connect with the external world, i.e., receive and send data and information.

Computers have an input-output subsystem, referred to as I/O subsystem, which provides an efficient mode of communication between the central system and the outside world. Programs and data must be entered into the computer memory for processing, and results obtained from computations must be displayed or recorded for the user's benefit. This data must first be stored in the computer's memory after converting it into machine-readable form. The data will then be processed (average marks calculated) and sent from the memory to the output unit, which will present the data in a form that can be read by users.

The I/O devices that provide a means of communication between the computer and the outside world are known as peripheral devices. This is because they surround the CPU and the memory of a computer system. While input devices are used to enter data from the outside world into the primary storage, output devices are used to provide the processed results from primary storage to users.

Input devices are used to transfer user data and instructions to the computer. The most commonly used input devices can be classified into the following categories:

- Keyboard devices (general and special purpose, key-to-tape, key-to-disk, key-to-diskette)

NOTES

- Point-and-draw devices (mouse, trackball, joystick, light pen, touch screen)
- Scanning devices (optical mark recognition, magnetic ink character recognition, optical barcode reader, digitizer, electronic-card reader)
- Voice recognition devices
- Vision-input devices (webcam, video camera)

1.5.1 Keyboard Devices

Keyboard devices allow input into the computer system by pressing a set of keys mounted on a board, connected to the computer system. Keyboard devices are typically classified as general-purpose keyboards and special-purpose keyboards.

(i) General-purpose keyboard

The most popular keyboard used today is the 101-key with a traditional QWERTY layout, with an alphanumeric keypad, 12 function keys, a variety of special function keys, numeric keypad, and dedicated cursor control keys. It is so called because the arrangement of its alphanumeric keys in the upper-left row (as seen in the Figure 1.16).

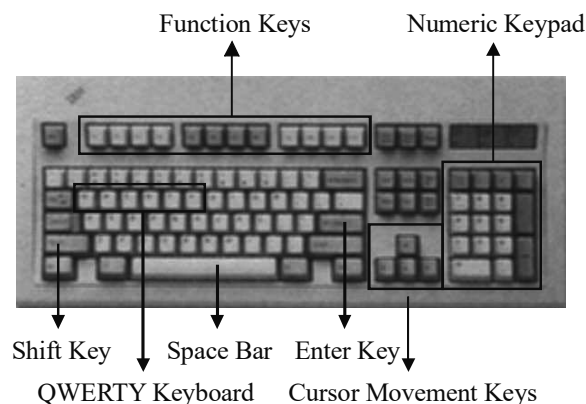


Fig. 1.16 QWERTY Keyboard Layout

- **Alphanumeric keypad:** This contains keys for the English alphabets, 0 to 9 numbers, special characters like * + - / [], etc.
- **12 Function keys:** These are keys labelled F1, F2 ... F12 and are a set of user-programmable function keys. The actual function assigned to a function key differs from one software package to another. These keys are also called soft keys since their functionality can be defined by the software.
- **Special function keys:** Special functions are assigned to each of these keys. The enter key, for example, is used to send the keyed-in data into the memory.
- **Numeric keypad:** This consists of keys with numbers (0 to 9) and mathematical operators (+ - * /) defined on them. It is usually located on the right side of the keyboard and supports quick entry of numerical data.
- **Cursor control keys:** They are defined by the arrow keys used to move the cursor in the direction indicated by the arrow (top, down, left, right).

Another popular key arrangement, called Dvorak system, was designed for easy learning and use. It was designed with the most common consonants in one part and all the vowels on the other part of the middle row of the keyboard. This key arrangement made the users use alternate keystrokes back and forth between both the hands. This keyboard was never commonly used.

(ii) Special-purpose keyboard

These have special-purpose keyboards to enable faster data entry and can be seen at the Automatic Teller Machines or the ATMs where the keyboard is required for limited functionality (support for some financial transactions) by the customers. These keyboards are specifically designed for special types of applications only.

(iii) Key-to-tape, key-to-disk, key-to-diskette

Used for data entry only; these processor-based workstations normally have a keyboard and a small monitor. The function of the processor is to check the accuracy of the data when it is being entered.

The screen displays data as it is being entered. These facilities are very useful and desirable during mass data entry and are therefore becoming very popular in data processing centers.

1.5.2 Point-and-draw Devices

The keyboard facilitates input of data only in text form. While working with display-based packages, we usually point to a display area and select an option from the screen (fundamentals of GUI applications). For such cases, the sheer user-friendliness of input devices that can rapidly point to a particular option displayed on screen and support its selection resulted in the advent of various point-and-draw devices.

Mouse

A mouse is a small input device used to move the cursor on a computer screen to give instructions to the computer and to run programs and applications. It can be used to select menu commands, move icons, size windows, start programs, close windows, etc. Nowadays, the mouse is the most important device in the functioning of a graphical user interface (GUI) of almost all computer systems. Figure 1.17 displays an optical mouse.



Fig. 1.17 Optical Mouse

You can click a mouse button, i.e., press and release the left mouse button, to select an item. You can *right click*, i.e., press and release the right mouse button to display a list of commands. You can *double click*, i.e., quickly press the left mouse button twice without any time gap between the press of the buttons, to open a program or a document. You can also *drag and drop*, i.e., place the

NOTES

cursor over an item on the screen and then press and hold down the left mouse button. Holding down the button, move the cursor to where you want to place the item and then release the button.

NOTES

1.5.3 Touch Screen

A touch screen is probably one of the simplest and most intuitive of all input devices. It uses optical sensors in or near the computer screen that can detect the touch of a finger on the screen. Once the user touches a particular screen position, sensors communicate the position to the computer. This is then interpreted by the computer to understand the user's choice for input. The most common usage of touch screens is in information kiosks where users can receive information at the touch of a screen. Figure 1.18 displays a touch screen.



Fig. 1.18 A Touch Screen

Touch pads

Input devices that are touch sensitive and take user input to manage onscreen pointer and other functions are known as touch pads. The touch pad enables the user to interact with the screen through the use of fingers that are dragged across the pad in various positions. The touch pad is used in place of the mouse which is an external peripheral device. Touch pads are usually present in notebooks and laptops as they provide convenience and space. Touch pads are touch as well as pressure sensitive and respond to finger drag and tapping combinations. In addition to the normal functions of a mouse, a user can also perform the scroll function by moving the finger at certain points, (the top and bottom corners of the right side of the pad). Touch pads are more convenient than a mouse as they put less pressure on the wrist and hand. The functionality of a touch pad is not limited by the manufacturer's user interface. They can be user programmed to recognize a combination of finger and tap moments to perform new actions as an input device. Touch pads have shown steady growth in market demand and user acceptability over all phases of their development cycles. Their growth is further expected to continue in the same way as the demand for touch-sensitive portable devices with more functionality and better appearance is on the rise. Figure 1.19 displays a touch pad.

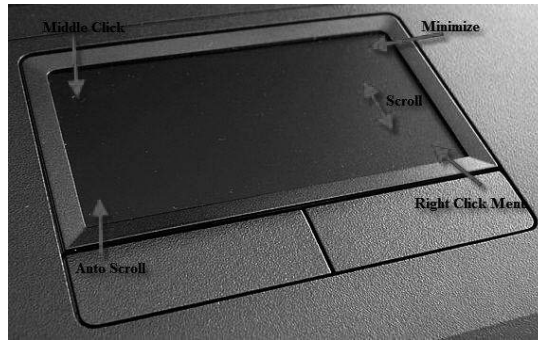


Fig. 1.19 A Touch Pad

Light pen

A light pen is a small input device used to select and display objects on a screen. It functions with a light sensor and has a lens on the tip of a pen-shaped device. The light receptor is activated by pointing the light pen towards the display screen and it then locates the position of the pen with the help of a scanning beam application to directly draw on screen. Figure 1.20 displays a light pen.

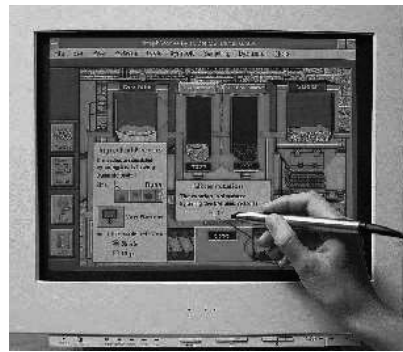


Fig. 1.20 A Light Pen being used to Display Objects on Screen

Trackball

A trackball is a pointing device that is much like an inverted mouse. It consists of a ball inset in a small external box or adjacent to, and in the same unit, as the keyboard of some portable computers. Figure 1.21 displays a trackball.



Fig. 1.21 A Trackball

It is more convenient and requires much less space than the mouse since here the whole device is not moved (as in the case of a mouse). Trackball comes in various

NOTES

shapes but supports the same functionality. Typical shapes used are a ball, a square and a button (typically seen in laptops).

NOTES

Joystick

A joystick is a vertical stick that moves the graphic cursor in the direction the stick is moved. It consists of a spherical ball, which moves within a socket and has a stick mounted on it. The user moves the ball with the help of the stick that can be moved left or right, forward or backward, to move and position the cursor in the desired location. Joysticks typically have a button on top that is used to select the option pointed by the cursor. Figure 1.22 displays a joystick.



Fig. 1.22 A Joystick

Video games, training simulators and control panels of robots are some common activities where joysticks are used.

3. Scanning Devices

Scanning devices are input devices used for direct data entry from the source document into the computer system. With the help of the scanner you can capture your images and documents and convert it into digital formats for easy storage on your computer. The scanner comprises of two major components, the first component lights up the page in order to capture the optical image and the second component converts that captured image into a digital format for viewing and storing it on your computer system. Figure 1.23 displays a hand-held scanner.



Fig. 1.23 A Hand-held Scanner

There are two types of scanners, *contact* and *laser*. Both illuminate the image first to calculate the reflected light and determine the value of the captured image. Hand-held contact scanners make contact as they are brushed over the printed matter to be read. Laser-based scanners are more versatile and can read data passed near the scanning area. Figure 1.24 displays a flat-bed scanner.



Fig. 1.24 A Flat-bed Scanner

Hand-held scanners are used where the information to be scanned or the volume of documents to be scanned is very low. They are much cheaper as compared to the flat-bed scanners. Capturing information using scanners reduces the possibility of human error typically seen during large data entry. The reduction in human intervention improves the accuracy of data and provides for timeliness of the information processed.

Source data automation is the recent development for data input technologies. Source data automation does not require any manual data entry; rather, it captures the data as a derivative of the routine business activity.

(i) Optical Mark Recognition (OMR)

The OMR devices can scan marks from a computer-readable paper. Such devices are used by universities and institutes to mark test sheets where the candidate selects and marks the correct answer from multiple choices given on a special sheet of paper. These marksheets are not required to be evaluated manually as they are fed in the OMR and the data is then transferred to the computer system for further evaluation. Figure 1.25 displays an OMR sheet.

Answer Sheet				
1.	a	b	c	d
	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	a	b	c	d
	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
3.	a	b	c	d
	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.	a	b	c	d
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Page 1 of 4				

Fig. 1.25 Sheet Read by an OMR Device

NOTES

NOTES

The actual technique used by an OMR device once again involves focussing a light on the page being scanned, thereby detecting the reflected light pattern for the marks. Pencil marks made by the user reflect the light determining which responses are marked.

(ii) Magnetic Ink Character Recognition (MICR)

Magnetic ink character recognition is like an optical mark recognition device and is used only in the banking industry. MICR devices scan cheque numbers directly from the cheque leaflets and then automatically feed them in the computer systems for further use, doing the job quickly, accurately and efficiently. Figure 1.26 displays magnetic ink characters on a bank cheque.

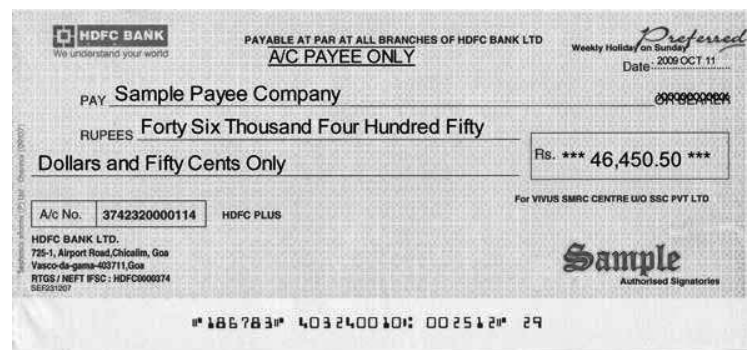


Fig. 1.26 A Bank Cheque using MICR Technology

Banks using MICR technology print cheque books on special types of paper. The necessary details of the bank (like the bank's identification code, relevant account number and cheque number) are pre-printed on the cheques using an ink that contains iron oxide particles that can be magnetized.

MICR readers are used to read and sort cheques and deposits. An MICR reader-sorter reads the data on the cheques and sorts the cheques for distribution to other banks and customers or for further processing.

(iii) Optical Bar Code Reader (OBR)

Data coded in the form of small vertical lines forms the basis of bar coding. Alphanumeric data is represented using adjacent vertical lines called *bar codes*. These are of varying widths and the spacing between them are used to uniquely identify books, merchandise in stores, postal packages, etc. Figure 1.27 shows a bar code used on a book for its unique identification.



Fig. 1.27(a) An Example of a Bar Code



Fig. 1.27(b) A Bar Code Reader

A bar code reader uses laser beam technology. The laser beam is moved across the pattern of bars in a bar code. These bars reflect the beam in different ways. The reflected beam is then sensed by a light-sensitive detector, which then converts the light patterns into electrical pulses, thereby transmitting them to logic circuits for further conversion to alphanumeric value. Bar code devices are available as hand-held devices.

NOTES

(iv) Digitizer

Digitizers are used to convert drawings or pictures and maps into a digital format for storage into the computer. A digitizer consists of a digitizing or graphics tablet, which is a pressure-sensitive tablet, and a pen with the same X and Y coordinates as on the screen. Some digitizing tablets also use a crosshair device instead of a pen. The movement of the pen or crosshair is reproduced simultaneously on the display screen. When the pen is moved on the tablet, the cursor on the computer screen moves simultaneously to the corresponding position on the screen (X and Y coordinates). This allows the user to draw sketches directly or input existing sketched drawings easily. Digitizers are commonly used by architects and engineers as a tool for computer-aided designing (CAD). Figure 1.28 displays a digitizing tablet.



Fig. 1.28 A Digitizing Tablet

(v) Electronic-card reader

Card readers are devices that also allow direct data input into a computer system. The electronic-card reader is connected to a computer system and reads the data encoded on an electronic card and transfers it to the computer system for further processing. Electronic cards are plastic cards with data encoded on them and meant for a specific application. A popular example of electronic cards are the plastic cards issued by banks to their customers for use in ATMs. Electronic cards are also used by many organizations for controlling access of various types of employees to physically secured areas. Figure 1.29 displays an access card security system.

NOTES



Fig. 1.29 An Access Card Security System

Based how the data is encoded, electronic cards may be either magnetic strip cards or smart cards. Magnetic strip cards have a magnetic strip on the back of the card. Data stored on magnetic strips cannot be read with the naked eye, a useful way to maintain confidential data. Smart cards, going a stage further, have a built-in microprocessor chip where data can be permanently stored. They also possess some processing capability making them suitable for a variety of applications. To gain access, for example, an employee inserts a card or badge in the reader. This device reads and checks the authorization code before permitting the individual to enter a secured area. Since smart cards can hold more information as compared to magnetic strip cards, they are gaining popularity.

1.5.5 Voice Recognition Devices

Voice recognition devices consist of a microphone attached to the computer system. A user speaks into the microphone to input data. The spoken words are then converted into electrical signals (this is in the analog form). A digital-to-analog converter then converts the analog form to digital form (0s and 1s) that can be interpreted by the computer. The digitized version is then matched with the existing pre-created dictionary to perform the necessary action. Voice recognition devices have limited usage today because they have several problems. Not only do they require the ability to recognize who is speaking, but also what is being said (the message). This difficulty arises primarily because people speak with different accents and different tones and pitches. The computer requires a large vocabulary to be able to interpret what is being said. Thus, voice recognition systems are successful in a limited domain.

Vision-input Devices

Vision input devices allow data input in the form of images. It mostly consists of a digital camera, which focuses on the object whose picture is to be taken. The camera creates the image of the object in digital format, which can then be stored within the computer.

The speech recognition system digitizes the voice input and this system similarly compares the digitized images to be interpreted to the prerecorded digitized images of the database of your computer system. After finding the right match, it sends it for further processing or pre-defined action.

Web camera

A web camera is a video capturing device attached to the computer system, mostly using a USB port used for video conferencing, video security, as a control input device. Figure 1.30 displays a web camera.



Fig. 1.30 A Web Camera

NOTES

1.5.6 Output Devices

An electromechanical device that accepts data from the computer and translates it into a form that can be understood by the external world is known as an output device. The processed data, stored in the memory of the computer, is sent to an output unit, which then transforms the internal representation of data into a form that can be read by the users.

Normally, the output is produced on a display unit like a computer monitor or can be printed through a printer on paper. At times, speech outputs and mechanical outputs are also used for some specific applications.

Output produced on display units or speech output that cannot be touched is referred to as softcopy output while output produced on paper or material that can be touched is known as hardcopy output. A wide range of output devices are available today and can be broadly classified as follows:

- Display devices (monitors, multimedia projectors)
- Speakers
- Printers (dot matrix, inkjet, laser)
- Plotters (flatbed, drum)

1. Display Devices

A display device is the most essential peripheral of a computer system. Initially, alphanumeric display terminals were used that formed a 7×5 or 9×7 array of dots to display text characters only. As a result of the increasing demand and use of graphics and GUIs, graphic display units were introduced. These graphic display

NOTES

units are based on series of dots known as pixels used to display images. Every single dot displayed on the screen can be addressed uniquely and directly.

Since each dot can be addressed as a separate unit, it provides greater flexibility for drawing pictures. Display screen technology may be of the following categories:

- (i) **Cathode ray tube (CRT):** The CRT consist of an electron gun with an electron beam controlled with electromagnetic fields and a phosphate-coated glass display screen structured into a grid of small dots known as pixels. The image is created with the electron beam produced by the electron gun, which is thrown on the phosphor coat displayed by the electromagnetic field.
- (ii) **Liquid crystal display (LCD):** Liquid Crystal Display (LCD) was first introduced in the 1970s in digital clocks and watches, and is now widely being used in computer display units. The Cathode Ray Tube (CRT) was replaced with the Liquid Crystal Display (LCD) making it slimmer and more compact. However, the image quality and the image color capability got comparatively poorer.

The main advantage of LCD is its low energy consumption. It finds its most common usage in portable devices where size and energy consumption are of main importance.

- (i) **Projection display:** Projection display technology is characterized by replacing the personal size screen with large screens upon which the images are projected. It is attached to the computer system and the magnified display of the computer system is projected on a large screen.
- (ii) **Monitors:** Monitors use a cathode ray tube (CRT) to display information. It resembles a television screen and is similar to it in other respects as displayed in Figure 1.31.

The monitor is connected to a keyboard for manual input of characters. The screen displays information as it is keyed in, enabling a visual check of the input before it is transferred to the computer. It is also used to display the output from the computer and hence serves as both an input and an output device.



Fig. 1.31 A CRT Monitor

The monitor is the most commonly used input/output device today and is also known as a soft copy terminal. A printing device is usually required to provide a hard copy of the output.

Multimedia Projector

A multimedia projector is an output device which is used to project information from the computer onto a large screen so that it can be viewed by a large group of people. Prior to this, the standard mode of making presentations was to make transparencies and project them using an overhead projector. This was a tedious and time-consuming activity, since for every change in the subject matter a new transparency had to be prepared. And of course, since electronic cut, copy and paste was not possible, this meant additional work.

A multimedia projector can directly be plugged into a computer system and the information projected on a large screen thereby making it possible to present information to a large audience. The presenter can also use a pointer to emphasize specific areas of interest in the presentation.

LCD flat screens connected to the computer systems for projecting the LCD image using an overhead multimedia projector are widely used today.

Speakers

Speakers are used to produce music or speech from programs, a speaker port (a port is a connector in your computer wherein you can connect an external device) allows connection of a speaker to a computer. Speakers can be built into the computer or can be attached separately. Figure 1.32 displays speakers.



Fig. 1.32 Speakers

1.5.7 Hard Copy Devices

Printers are used for generating output printed on paper. There is a huge range of commercially available printers today (estimated to be 1500 different types). These printers can be classified into categories based on the following:

- Printing technology
 - Printing speed
 - Printing quality
- (i) **Printing technology:** Printers can be classified as impact or non-impact printers, based on the technology they use for producing output. Impact printers work on the mechanism similar to a manual typewriter where the printer head strikes on the paper and leaves the impression through

NOTES

NOTES

an inked ribbon. Dot matrix printers and character printers fall under this category. Non-impact printers use chemicals, inks, toners, heat or electric signals to print on the paper and they do not physically touch the paper while printing.

(ii) **Printing speed:** The number of characters printed in a unit of time reflects the printing speed. Based on speed, these may be classified as character printer (prints one character at a time), line printers (prints one line at a time), and page printers (print the entire page at a time). Printer speed is measured in terms of characters-per-second or *cps* for a character printer, lines-per-minute or *lpm* for a line printer, and pages-per-minute or *ppm* for a page printer.

(iii) **Printing quality:** The quality determined by the resolution of printing and is characterized by the number of dots that can be printed per linear inch, horizontally or vertically. It is measured in terms of dots-per-inch or *dpi*. Printers can be classified as near-letter-quality or NLQ, letter-quality or LQ, near-typeset-quality or NTQ and typeset-quality or TQ, based on their printing quality. NLQ printers have resolutions of about 300 dpi, LQ of about 600 dpi, NTQ of about 1200 dpi and TQ of about 2000 dpi. NLQ and LQ printers are used for ordinary printing in day-to-day activities, while NTQ and TQ printers are used to produce top-quality printing, typically required in the publishing industry.

The different types of printers are as follows:

(i) Dot matrix

Dot matrix printers are the most widely used impact printers in personal computing. These printers use a print head consisting of a series of small metal pins that strike on a paper through an inked ribbon, leaving an impression on the paper through the ink transferred. Characters thus produced are in a matrix format. The shape of each character, i.e., the dot pattern, is obtained from information held electronically. Figure 1.33 displays characters formed using dots and dot matrix printer.

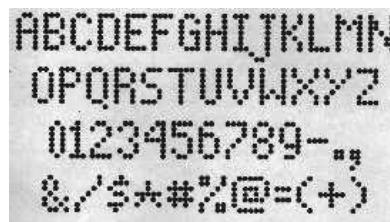


Fig. 1.33(a) Characters Formed using Dots



Fig. 1.33(b) Dot Matrix Printer

The speed, versatility and ruggedness, combined with low cost, tend to make such printers particularly attractive in the personal computer market. Typical printing speeds in case of dot matrix printers range between 40–1000 characters per second (cps). In spite of all these features in dot matrix printer technology, the low print quality gives it a major disadvantage.

(ii) Inkjet

Inkjet printers are based on the use of a series of nozzles for propelling droplets of printing ink directly on almost any size of paper. They, therefore, fall under the category of non-impact printers. The print head of an inkjet printer consists of a number of tiny nozzles that can be selectively heated up in a few microseconds by an IC register. When this happens, the ink near it vaporizes and is ejected through the nozzle to make a dot on the paper placed in front of the print head. The character is printed by selectively heating the appropriate set of nozzles as the print head moves horizontally. Figure 1.34 displays an inkjet printer.



Fig. 1.34 An Inkjet Printer

(iii) Laser

Laser printers work on the same printing technology as photocopiers, using static electricity and heat to print with a high quality powder substance known as toner. Figure 1.35 displays a laser printer.



Fig. 1.35 A Laser Printer

Laser printers are capable of converting computer output into print, page by page. Since characters are formed by very tiny ink particles, they can produce very high quality images (text and graphics). They generally offer a wide variety of character fonts, and are silent and fast in use. Laser printers are faster in printing speed than other printers discussed above. Laser printers can print from 10 pages to 100 pages per minute, depending upon the make/model. Laser is high quality, high speed, high volume and non-impact technology that works on almost any kind of

NOTES

paper. Even though this technology is more expensive than inkjet printers, it is preferred because of its unmatched features, such as high quality, high-speed printing and noiseless and easy-to-use operations.

NOTES

Plotters

Plotters are used to make line illustrations on paper. They are capable of producing charts, drawings, graphics, maps and so on. A plotter is much like a printer but is designed to print graphs instead of alphanumeric characters. Based on the technology used, there are mainly two type of plotters: *pen plotters* and *electrostatic plotters*. Pen plotters, also known as flatbed plotters, draw images with multicolored pens attached to a mechanical arm. Electrostatic plotters, also known as drum plotters, work on the same technology as laser printers.

Flatbed plotters and drums are the most commonly used plotters.

(i) Flatbed plotters

Flatbed plotters have a flat base like a drawing board on which the paper is laid [as shown in Figure 1.36(a)]. One or more arms, each of them carrying an ink pen, moves across the paper to draw. The arm movement is controlled by a microprocessor (chip). The arm can move in two directions, one parallel to the plotter and the other perpendicular to it (called the x and y directions). With this kind of movement, it can move very precisely to any point on the paper placed below.

The computer sends the commands to the plotter which are translated into x and y movements. The arm moves in very small steps to produce continuous and smooth graphics. The size of the plot in a flatbed plotter is limited only by the size of the plotter's bed. The advantage of flatbed plotters is that the user can easily control the graphics. He can manually pick up the arm anytime during the production of graphics and place it on any position on the paper to alter the position of graphics to his choice.

However, flatbed plotters occupy a large amount of space.

(ii) Drum plotters

Drum plotters move the paper with the help of a drum revolver during printing [as shown in Figure 1.36(b)]. The arm carrying a pen moves only in one direction, perpendicular to the direction of the motion of the paper. It means that while printing, the plotter pens print on one axis of the paper and the cylindrical drum moves the paper on the other axis. With this printing technology, the plotter has an advantage to print on unlimited length of paper but on a limited width. Drum plotters are compact and lightweight compared to flatbed plotters. This is one of the advantages of such plotters.

The disadvantage of drum plotters, however, is that the user cannot freely control the graphics when they are being created. Plotters are more expensive when compared to printers. Typical application areas for plotters include Computer-Aided Engineering (CAE) applications, such as Computer-Aided Design (CAD) and Computer-Aided Manufacturing (CAM), architectural drawing and map drawing.

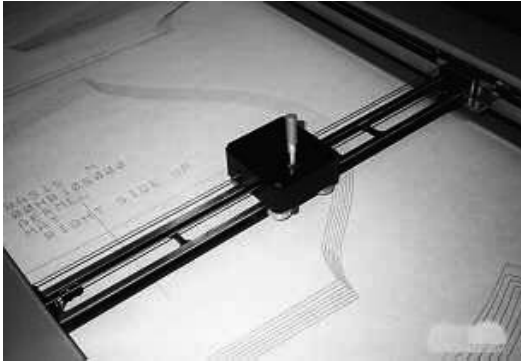


Fig. 1.36(a) Top View of a Flatbed Plotter



Fig. 1.36(b) A Drum Plotter

NOTES

1.6 STORAGE MEDIA

The most common properties used for characterizing and evaluating the storage unit of the computer system are expressed below:

1. **Storage capacity:** Represents the size of the memory. It is the amount of data that can be stored in the storage unit. Primary storage units have less storage capacity as compared to secondary storage units. The capacity of internal memory and main memory can be expressed in terms of the number of words or bytes. The capacity of external or secondary storage, on the other hand, is measured in terms of bytes.
2. **Storage cost:** Cost is another key factor that is of prime concern in a memory system. It is usually expressed per bit. It is obvious that lower costs are desirable. It is worth noting that with the increase in the access time for memories, the cost decreases.
3. **Access time:** The time required to locate and retrieve the data from the storage unit. It is dependant on the physical characteristics and access mode used for that device. Primary storage units have faster access time as compared to secondary storage units.
4. **Access mode:** Memory comprises various locations. Access mode is the mode in which information is accessed from the memory. The user can access memory devices in any of the following ways:
 - (a) **Random access memory (RAM):** This refers to the mode in which any memory location can be accessed in any order in the same amount of time. Ferrite and semiconductor memories, which usually constitute the primary storage or main memory, are of this nature.
 - (b) **Sequential access:** Memories that can be accessed only in a pre-defined sequence are sequential access memories. Here, since sequencing through other locations precedes the arrival at a desired location, the access time varies according to the location. Information on a sequential device can be retrieved in the same sequence in which it was stored. Songs stored on a cassette, that can be accessed only one by one, is an example of sequential access. Typically, magnetic tapes are sequential access memory.

NOTES

(c) **Direct access:** Sometimes, the information is neither accessed randomly nor in sequence but something in between. In this type of access, a separate read/write head exists for each track, and on a track, you can access the information serially. This semi-random mode of access exists in magnetic disks.

5. **Permanence of storage:** If the storage unit can retain the data even after the power is turned off or interrupted, it is termed as non-volatile storage. And, if the data is lost once the power is turned off or interrupted, it is called volatile storage. It is obvious from these properties that the primary storage units of the computer systems are volatile, while the secondary storage units are non-volatile. A non-volatile storage is definitely more desirable and feasible for storage of large volumes of data.

Primary Storage Devices

1. Static and Dynamic RAM

The main memory is the central storage unit in a computer system. It is a relatively large and fast memory. It is used to store programs and data during computer operations. The principal technology used for the main memory is based on semiconductor-integrated circuits. There are two possible modes in which the integrated circuit RAM chips are available. These modes are: *static* and *dynamic*.

The static RAM (SRAM) stores binary information using clocked sequential circuits. The stored information remains valid only as long as power is applied to the unit. On the other hand, dynamic RAM (DRAM) stores binary information in the form of electric charges that are applied to capacitors inside the chip. The stored charge on the capacitors tends to discharge with time and so must be periodically recharged by refreshing the dynamic memory. The dynamic RAM offers larger storage capacity and reduced power consumption. Therefore, large memories use dynamic RAM, while static RAM is mainly used for specialized applications.

The different types of memory discussed above are both of the read/write type. What about a memory where only one of the operations is possible, e.g., if we allow only reading from the memory (cannot change the information in the memory)? The memory might have some major importance; like an important bit of the computer's operating system which normally does not change can be stored in this kind of memory. Such a memory is called ROM (Read Only Memory).

2. Read-only Memory (ROM)

Most of the memory in a general-purpose computer is made of RAM integrated circuit chips, but a portion of the memory may be constructed using ROM chips. Originally, RAM was used to refer to random-access memory, but now we use the term read/write memory to distinguish it from read-only memory (since ROM is also random access). RAM is used for storing the bulk of the programs and data that are subject to change, while ROM is used to store programs that are permanently resident in the computer and do not change once the production of the computer is completed.

Among other things, the ROM portion of the main memory is used for storing an initial program called the bootstrap loader, whose function is to get the computer software operating when power is turned on. Since RAM is volatile, its contents are destroyed when power is turned off. The contents of ROM remain unchanged even after the power is turned off and on again.

Read-only memories can be manufacturer-programmed or user-programmed. When the data is burnt into the circuitry of the computer by the manufacturer, it is called manufacturer-programmed ROM; for example, a personal computer manufacturer may store the boot program permanently in the ROM chip of the computers manufactured by it. Such chips are supplied by the manufacturer and are not modifiable by users. This is an inflexible process and requires mass production. Thus, a new kind of ROM, known as Programmable Read-only Memory (PROM), was designed. This is also non-volatile in nature. It can be written only once using some special equipment. The supplier or the customer can electrically perform the writing process in PROM.

In both ROM and PROM, you can perform write operations only once and you cannot change whatever you have written. But what about the cases where you mostly read but also write a few times? Another type of memory chip called EPROM (Erasable Programmable Read-only Memory) was developed to take care of such situations. EPROMs are typically used by R&D personnel who experiment by changing micro-programs on the computer system to test their efficiency.

Further, EPROM chips are of two types: EEPROMs (Electrically EPROM) in which high voltage electric pulses are used to erase stored information, and UVEPROM (Ultra Violet EPROM) in which stored information is erased by exposing the chip for a while to ultraviolet light.

Figure 1.37 summarizes the various types of Random Access Memories.

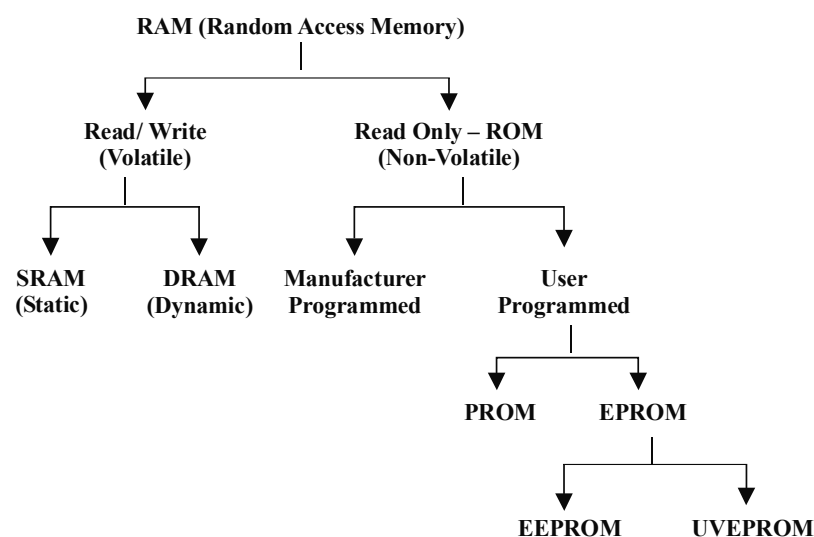


Fig. 1.37 Types of Random Access Memories

NOTES

Secondary Storage Devices

NOTES

As discussed earlier, RAM is a volatile memory with limited storage capacity. The cost of RAM is also relatively higher as compared to secondary memory. Logic dictated that a relatively cheaper media, showing some sort of permanence of storage, be used. As a result, additional memory called *external* or *auxiliary memory* or *secondary storage* is used in most computers.

The magnetic medium was found to be long lasting and fairly inexpensive, and therefore, became an ideal choice for large storage requirements. The use of magnetic tapes and disks as storage media is very common. As optical technology is advancing, optical disks are turning out to be one of the major secondary storage devices.

1. Cache Memory

Cache memories are small, fast memories placed between the CPU and the main memory. They are faster than the main memory with access times closer to the speed of the CPU. Caches are fast, but very expensive. So, they are used only in small quantities. As an example, caches of size 64K, 128K are normally used in PC-386 and PC-486, which can have 1 to 8 MB of RAM or even more. Cache memories provide fast speed memory retrieval without compromising the size of the memory.

If the memory is so small, would it be advantageous to increase the overall speed of memory? This can be answered with the help of a phenomenon known as locality of reference. Let us examine what this means.

If we analyse a large number of typical programs, we would find that memory references at any given interval of time tend to be confined to a few localized areas in the memory. This phenomenon is called the property of locality of reference. This is true because most of the programs typically contain iterative loops (like 'for' or 'while' loops). During the execution of such programs, the same set of instructions (within the loop) are executed many times. The CPU repeatedly refers to the set of instructions in the memory that constitute the loop. Every time a specific subroutine is called, its set of instructions is fetched from the memory. Thus loops and subroutines tend to localize the references to memory for fetching instructions.

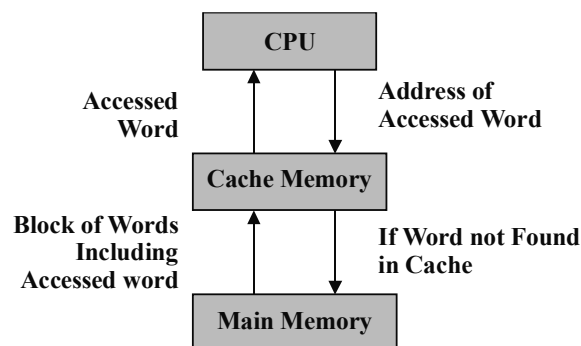


Fig. 1.38 Functioning of the Cache Memory

Based on the locality of reference, we understand that the cache has a copy of certain portions of main memory. First, the memory read or write operation is checked with the cache, and in case of availability of desired data in the cache, it is used directly by the CPU. Otherwise, a block of words is read from main memory to cache and the word is then used by the CPU from cache. Figure 1.38 explains the functioning of the cache memory.

NOTES

1.6.1 Traditional Storage-Magnetic Tapes

Magnetic tapes are used for storing files of data that are sequentially accessed or not used very often and are stored off-line. They are typically used as backup storage for archiving of data.

In case of magnetic tapes, a tape (plastic ribbon usually 1/2 inch or 1/4 inch wide and 50 to 2400 feet long) is wound on a spool and its other end is threaded manually on a take-up spool. The beginning of the tape (BOT) is indicated by a metal foil called a *marker*. When a write command is given, a block of data (records are usually grouped in blocks of two or more) is written on the tape. The next block is then written after a gap (called Inter Block Gap or IBG). A series of blocks are written in this manner. The end of tape (EOT) is indicated by an end-of-tape marker which is a metal foil stuck in the tape. After the data is written, the tape is rewound and kept ready for reading. Figure 1.39(a) shows the data organization on a magnetic tape. Figure 1.39(b) and 1.39(c) show magnetic tape reels and magnetic tape cartridges.

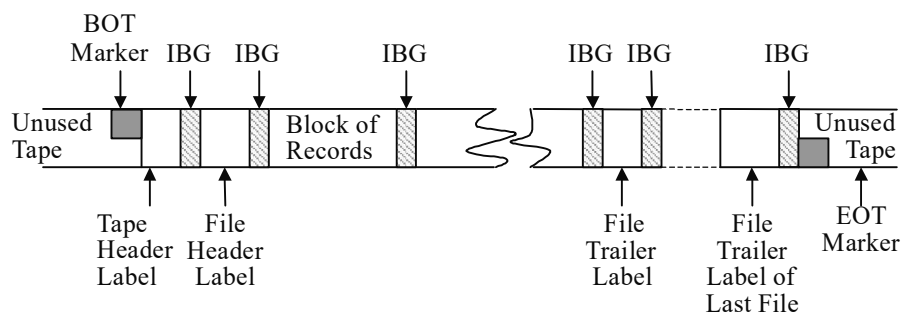


Fig. 1.39(a) Data Organization on a Magnetic Tape



Fig. 1.39(b) Magnetic Tape Reel

NOTES



Fig. 1.39(c) Magnetic Tape Cartridge

The tape is read sequentially, i.e., data can be read in the order in which the data has been written. This implies that if the desired record is at the end of the tape, all the earlier records have to be read before it is reached. A typical example of a tape can be seen in a music cassette, where, to listen to the fifth song one must listen to, or traverse, the earlier four songs. The access time of information stored on tape is therefore very high as compared to that stored on a disk.

The storage capacity of the tape depends on its data recording density and the length of the tape. The data recording density is the amount of data that can be stored or the number of bytes that can be stored per linear inch of tape. The data recording density is measured in BPI (Bytes per inch).

Thus,

Storage capacity of a tape = Data recording density × Length of tape

It is worth noting that the actual storage capacity for storing user data is much less owing to the file header labels, file trailer labels, BOT and EOT markers, and the use of IBGs.

Some commonly used magnetic tapes are:

- 1/2 inch tape reel
- 1/2 inch tape cartridge
- 1/4 inch streamer tape
- 4 mm DAT (Digital Audio Tape) – typical capacity of 4GB to 14 GB

1.6.2 Magnetic Disks

Magnetic disks are direct-access medium, and so are the most popular online secondary storage devices. Direct-access devices are also called random-access devices because information is literally available at random or in any order. There is direct access to any location on the device. Thus, approximately equal access time is required for each location. An example of this is a music CD where if you wish to listen to the fifth song, you can directly select the fifth track without having to fast forward the previous four.

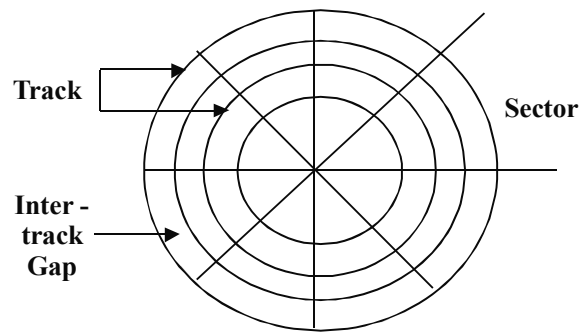


Fig. 1.40 Logical Layout of a Magnetic Disk

A magnetic disk refers to a circular plate made of metal or plastic and coated with magnetized material (as shown in Figure 1.40). Often both sides of the disk are used. Data is recorded on the disk in the form of magnetized and non-magnetized spots (not visible to the naked eye) representing 1s and 0s.

(i) Disk devices

A disk drive is a peripheral device used to store and collect information. It can be removable or fixed, high capacity or low capacity, fast or slow speed, and magnetic or optical.

Structurally, a drive is the object inside which a disk(s) is either permanently or temporarily stored. While a disk contains the media on which the data is stored, a drive contains the machinery and circuitry required for implementing the read/write operations on the disk.

The disk looks literally like a flat circular disk. The computer writes information onto the disk, where it is stored in the same form as it is stored on a cassette tape. Disks, as such, are just magnetically coated rolls or circular disks which are divided into sectors and tracks. The data is accordingly stored and numbered with respect to the track and sector number on the disk. Only the structure of the media is different from one to another. Examples of removable disk drives are DVD, CD-ROM, floppy disk drive, etc. The examples of non-removable disk drives include hard disk.

The method of accessing data could be sequential access (Magnetic Tape Drives) or random access (HDD, DVD), where the read/write head can directly go to any location on the disk and perform action.

(ii) Diskettes

Based on the size and packaging of the disks, they can be classified into two types — floppy disks and hard disks. Further, disks that are permanently attached to the unit assembly and cannot be removed by the occasional user are called hard disks. A drive using removable disks is called a floppy disk drive.

(iii) Floppy disks

The disks used with a floppy disk drive are small, removable disks made of plastic, and coated with magnetic recording material. There are two sizes commonly used, with diameters of 5.25 and 3.5 inches.

NOTES

NOTES

- The 5.25 inch disk is a 5.25 inch diameter floppy disk. Earlier, such disks recorded data only on one side and were called single-sided (SS) disks. Today both the surfaces are used for recording and are called double-sided (DS) disks. These are available in two capacities — double density (DD), and high density (HD), where density refers to the number of bits that can be stored per square inch area.

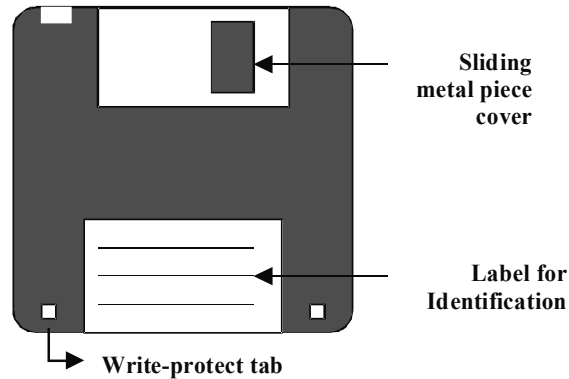


Fig. 1.41 A 3.5 Inch Floppy Disk

- The 3.5 inch disk is a disk of 3½ inch diameter. These record data on both sides and are therefore double sided disks. Figure 1.41 displays a 3.5 inch floppy disk. These disks come in three different capacities — double density, high density and very high density. These are smaller and can store more data than can the 5.25 inch disks.

The storage capacity for any disk can be calculated as:

$$\text{Storage capacity} = \text{Number of recording surfaces} \times \text{Number of tracks per surface} \times \text{Number of sectors per track} \times \text{Number of bytes per sector}$$

Thus, for a 3.5 inch high density disk which has 80 tracks, 18 sectors/track, and 512 bytes/sector, the disk storage capacity can be calculated as follows:

$$2 \times 80 \times 18 \times 512 = 14,74,560 \text{ bytes or } 1.4 \text{ MB (approximately)}$$

Table 1.4 provides the necessary details and associated storage capacities of various kinds of floppy disks.

Floppy disks were extensively used in personal computers as a medium for distributing software to computer users. Nowadays, CDs or DVDs are used for that purpose.

Table 1.4 Storage Capacities of Floppy Disks

Size (diameter in inches)	No. of Recording Surfaces	No. of Tracks	No. of Sectors/Tracks	No. of Bytes/Sector	Storage Capacity (approx)
5.25	2	40	9	512	3,68,640 bytes or 360kB
5.25	2	80	15	512	12,28,800 bytes or 1.2 MB
3.5	2	40	18	512	7,37,280 bytes or 720 kB
3.5	2	80	18	512	14,74,560 bytes or 1.4 MB
3.5	2	80	36	512	29,49,120 or 2.8 MB

NOTES

(ii) Hard disks

Unlike floppy disks, hard disks are made up of rigid metal. The sizes for the disk platters range between 1 to 14 inches in diameter. Depending on the way they are packaged, hard disks can be categorized as disk packs or Winchester disks.

- **Disk packs** consist of two or more hard disks mounted on a single central shaft. Because of this, all disks in a disk pack rotate at the same speed. It consists of separate read/write heads for each surface (excluding the upper surface of the topmost disk platter and the lower surface of the bottommost disk platter, as mentioned earlier). Disk packs are removable in the sense that they can be removed and kept offline when not in use (typically stored away in plastic cases) as displayed in figure 1.42. They have to be mounted on the disk drive before they can be used. Thus different disk packs can be mounted on the same disk drive at different instances, thereby providing virtually unlimited (modular) storage capacity.



Fig. 1.42 A Disk Pack

- **Winchester disks** also consist of two or more hard disk platters mounted on a single central shaft but are of the fixed type. The disk platters are sealed in a contamination-free container. Due to this fact, all the disk platters, including the upper surface of the topmost disk platter and the lower surface of the bottommost platter, are used for storing data. So, even though

Winchester disks have limited storage capacity as opposed to disk packs, they can store larger amounts of data as compared to the same number of disk platters. Figure 1.43 displays a winchester disk.

NOTES

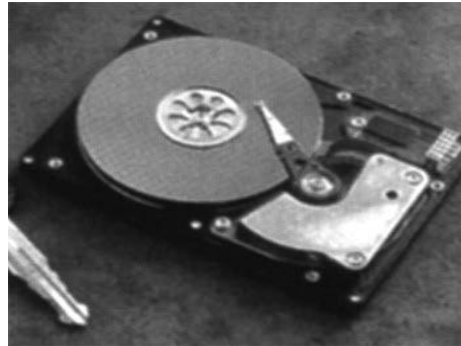


Fig. 1.43 A Winchester Disk

Another kind of disk called the zip disk is very common today. This consists of a single hard disk platter encased in a plastic cartridge. Such a disk typically has a capacity of about 100 MB. The zip drive can further be fixed or portable. The fixed zip drive is permanently connected to the computer system while the portable ones can be carried around and connected to any computer system for the duration of its use. In both cases, however, the zip cartridge (the actual storage medium) is portable just like a floppy, albeit with a nearly 100 times larger storage capacity. Figure 1.44 displays zip disks and zip drive.



Fig. 1.44 Zip Disks and Zip Drive

Optical Storage Devices

(i) Optical disks

Optical disks are storage devices with huge storage capacities. They are a relatively new storage medium and use laser beam technology for writing and reading of data.

Optical disks consist of one large track that starts from the outer edge and spirals inward towards the center (this is unlike the magnetic disk in which tracks are concentric circles on the disk platter). An optical disk is also split into sectors but these are of the same length regardless of their location on the track. Data is therefore packed at maximum density over the disk.

The storage capacity of an optical disk is determined as follows:

$$\text{Storage Capacity} = \text{Number of sectors} \times \text{Number of bytes per sector}$$

(Note that we do not consider the number of tracks since there is only one track in this case)

Thus, a 5.25 inch optical disk having 3,30,000 sectors and storing 2,352 bytes per sector, will have a storage capacity

$$3,30,000 \times 2352 = 77,61,60,000 \text{ bytes or } 740 \text{ MB (approx.)}$$

The technology used in optical disks uses laser beams to write and read data as opposed to the read/write head used in magnetic disks. Data is recorded by etching microscopic pits (burnt surface) on the disk surface. A high-intensity laser beam is used to etch the pits, while for retrieving data, a low-intensity laser beam is used. Figure 1.45 displays an optical disk formats.



Fig. 1.45 An Optical Disk and Disk Drive

There are three optical memory devices that are becoming increasingly popular in various computer applications: CD-ROM, WORM, and Erasable Optical disks. We shall discuss these in the succeeding section.

CD-ROM: The compact disk read-only memory (CD-ROM) is a direct extension of the audio CD. It is generally made from a resin called polycarbonate that is coated with aluminium to form a highly reflective surface. The information on a CD-ROM is stored as a series of microscopic pits on the reflective surface (using a high-intensity laser beam). The process of recording information on these disks is known as 'mastering'. This is so-called because this master disk is then used to make a 'die' that is used to make copies.

The information is retrieved from a CD-ROM using a low-powered laser that is generated in an optical disk drive unit. The disk is rotated and the laser beam is aimed at the disk. The intensity of the laser beam changes when it encounters a pit. A photo-sensor detects the change in intensity, hence recognizing the digital signals that are recorded on the surface of the CD-ROM and converts them into electronic signals of 1s and 0s.

As the name suggests, information stored in a CD-ROM can be read only. It cannot be modified in any way. It is therefore useful for applications in which there is a database of information that is useful as it is and does not need changing in any way, for example, a directory such as Yellow Pages. CD-ROMs are very

NOTES

NOTES

handy in the distribution of large amounts of information to a large number of users. The strengths of CD-ROMs lie in the fact that they provide for:

- Large storage capacity for information/data.
- Fast and inexpensive mass replication.
- Suitable for archival storage since they are removable disks.

The weaknesses of CD-ROMs are:

- They are read-only and cannot be updated.
- Access time in them is greater than that of a magnetic disk.

WORM: The drawbacks of CD-ROM were partially resolved by the introduction of WORM ('write-once, read many').

In some applications, only a limited copies of compact disks are required to be made. This makes the CD-ROM production economically not viable. This is because manufacturers duplicate CD-ROMs by using expensive duplication equipment. To overcome such a problem, write-once read-many CDs have been developed.

WORM disks allow users to create their own CDs by using a CD-Recordable (CD-R) drive. This can be attached as a peripheral device to the computer system. WORM disks recorded like this can be read by any CD-ROM drive.

Erasable optical disk: The erasable optical disk is the latest development in optical disks. Like magnetic disks, the data in the erasable optical disk can be changed repeatedly. Erasable optical disks are also known as rewritable optical disks.

These disks integrate the magnetic and optical disk technologies to enable rewritable storage with the laser-beam technology and so are also called magneto-optical disks. In such systems, a laser beam is used along with a magnetic field to read or write information on a disk that is coated with a magnetic material.

To write, the laser beam is used to heat a specific spot on the magnetic coated material. At this increased temperature, a magnetic field is applied so as to change the polarization of that spot, thereby recording the data that is required. It may be noted here that this process does not cause any physical changes in the disk. Hence, it can be repeated many times. The degree of rotation of the polarized laser beam reflected from the surface is detected to perform reading function. This implies that as the disk spins, the polarized spots pass under the laser beam, and depending on their orientation or alignment, some of them reflect the light while others scatter it. This produces patterns of 'on' and 'off' that are converted into electronic signals of binary 1s and 0s.

The capacity of an erasable disk is very high in comparison to a magnetic disk; for example, a 5.25 inch optical disk can store around 650 MB of data while Winchester disks normally can store a maximum capacity of 320 MB. This is why magneto-optical disks are ideal for multimedia applications that require large storage capacities.

DVD: Known as the Digital Versatile Disk or the Digital Video Disk, it has the same physical dimensions as that of a CD-ROM, but it can hold up to 4.4GB data on a Single layer disk and up to 8.47GB on a dual layer disk with the maximum data transfer of 27MB/s at 20x speeds. The laser used to read/write data on a DVD is much more precise and has a wavelength of 650nm, which is one reason why a DVD can hold more data.

HD-DVD: This is a high density, mostly single-sided, double-layered optical disc that can hold up to 15GB on a single layer and 30GB on a dual layer disc. The read/write speed on an HD-DVD varies between 36 MBPS to 72 MBPS. These were primarily designed for the storage of high definition videos and large volumes of data. The basic look and feel of an HD-DVD drive and disk is the same as that of a CD-ROM and DVD except that it uses a laser of a different wavelength and the microscopic structure of storage on a disk is different.

Blu-Ray Disc: Another high density optical storage media format is gaining increasing popularity these days. Its main uses are high-definition video and data storage. A dual layer Blu-ray Disc can store 50 GB, almost six times the capacity of a double-layer DVD (or more than 10 times if single-layer). The data can be read from 36 MBPS to 432 MBPS which is higher than HD-DVD. When used for HD video playback, the video is encoded on it in MPEG-2, AVC, VC-1 (H.264) format, which is the same as HD-DVD.

With the wide variety of storage choices available and constant increase in their number due to changes in technology, it is impossible to say that a single medium of storage can suit the needs of all. Therefore, it is necessary to make the choice of right media based upon factors like capacity, speed, durability, life span and cost of media.

Computer Output Microfilm (COM)

COM refers to a process characterized by copying/printing data from media located on personal computers, mini, or mainframe computers onto a microfilm. It comprises a high-speed recorder, which transfers digital data onto a microfilm applying laser technology, and a processor which develops the microfilm once exposed to the light source.

A computer output microfilm device translates information normally held on magnetic tape into miniature images on a microfilm (also called microfiche—'fiche' pronounced as 'fish'). The device displays information as characters on a CRT screen and then using photographic methods, records the display onto the film. Drawings and images can also be displayed along with narrative text.

A special reader/printer can be subsequently used to view the processed film. The reader operates on a 'back projection' principle displaying one frame at a time on a translucent screen, typically about A4 size. The printer can then be used selectively to produce a hard copy of what is presented on screen.

Figure 1.46 identifies the various steps in COM production.

NOTES

NOTES

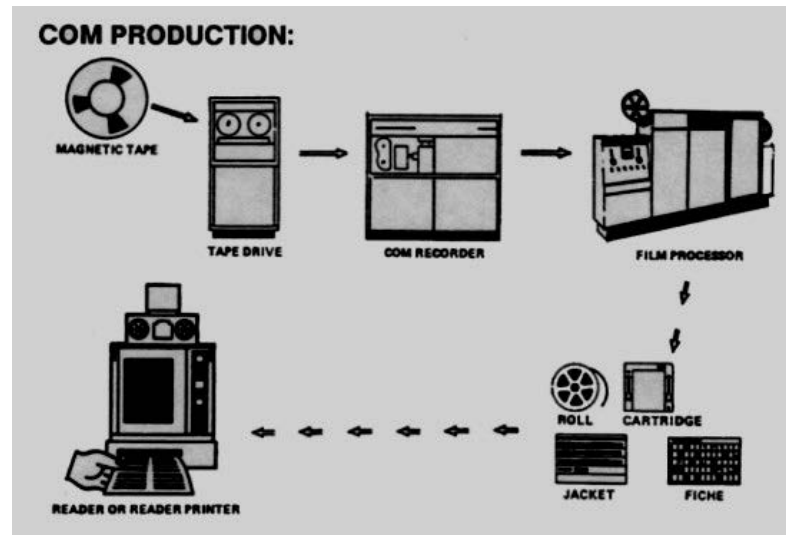


Fig. 1.46 Steps in a COM Production

A COM system provides an easy and compact way of recording and storing information, and subsequently retrieving desired pieces of information. It offers various advantages like reduction of paper, reduction in cost (since it is cheaper than most electronic media), improved quality (COM technology provides superior image quality), and electronic record retention/archiving.

COM is best suited for data requiring long-term storage. This is because microfilm is less volatile than magnetic media like disks and tapes. COM stores data in a very compact format. It is to be noted that up to 270 pages can be contained in a single 4×6 inch fiche.

Conversion of magnetic tapes to microfilms is cost-effective for closed files. However, in case of highly active data or data requiring regular updating, using microfilms may not be as efficient as retaining the information online. It is useful for data that must be archived for long periods of time and referenced only occasionally, e.g., information that must be archived to comply with legal regulations, information maintained by insurance companies, banks, government agencies, and various other organizations of this type.

Measuring drive performance

Disk performance can be categorized by the speed at which the data can be read or written. Over the years there have been changes in disk drive interface, rotation speeds, number of heads and cylinders and storage format, all of which have led to a decrease in data access time.

The various types of disks currently available in the market are:

- IDE – Integrated Drive Electronics
- EIDE – Enhanced Integrated Drive Electronics
- SCSI – Small Computer System Interface
- SATA I & II – Serial Advanced Technology Attachment

There are two standard methods for accessing and writing data on a disk — sequential access and random access.

Sequential Access is when you read or write on the disk blocks in sequential or continuous order, that is, one block after another. Examples of where Sequential Access is used in computing or data storage would be the backing up of data onto tape drives or the process of writing data onto CDs and DVDs. Any storage medium based on magnetic tape, VHS, audio cassette etc, are read and written by Sequential Access.

Random Access, as the name suggests, is performed when the hard drive head needs to read/write data from/at various locations on the disk. In this case, the disk heads move rapidly from one place to another and the seek time to access data is increased because it involves mechanical operations. Most of the disk operations performed during routine computer work are random access. This is also the reason why random access time is more important while measuring disk performance than sequential access time. While data is written onto optical media sequentially, data on CDs and DVDs can be read randomly.

For Random Access, the *average seek time* and *average latency time* are added to come up with the total time it takes for the disk to read and write data on it. The average seek time is the time it takes to move the head arm from one position to another, and average latency time is the time it takes for the required data block to come under the head for the read/write operations. The average latency time depends on the rotations per minute (RPM) of the disk, which is the speed at which the magnetic or optical disk rotates.

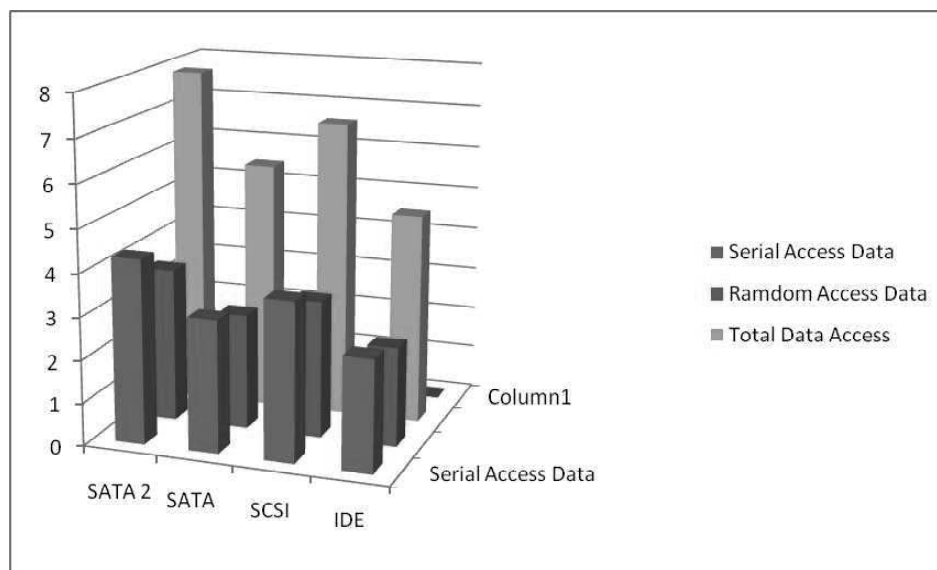


Fig. 1.47 Data Accessed by Different Disk Drives

The above bar chart shows the relative amount of data accessed by each type of disk drive.

NOTES

NOTES

Check Your Progress

8. What are scanning devices?
9. How is an output produced in computer?
10. How does a laser printer work?
11. What is cache memory?
12. State about the disk drives.

1.7 ANSWERS TO ‘CHECK YOUR PROGRESS’

1. A computer primarily constitutes three integral components, viz. input device, central processing unit (CPU) and output device.
2. The five advantages of using computers are speed, storage, and accuracy, diligence, versatility.
3. Charles Babbage, a 19th century professor at Cambridge University, is considered the father of the modern digital computer.
4. In the fourth generation computers, Large Scale Integration (LSI) were developed which could fit hundreds of components onto a single chip. By 1980s, Very Large Scale Integration (VLSI) squeezed thousands of components onto a single chip and later. Ultra Large Scale Integration (ULSI) increased that number to millions.
5. Mainframes as well as minicomputers are some of the examples of third generation computers.
6. The central processing unit (CPU) is the most important component of a digital computer that interprets the instructions, and processes the data contained in computer programs. CPU works as the brain of the computer and performs most of the calculations.
7. A shared communication path consisting of one or more connection lines is known as a bus and the transfer of data through this bus is known as bus transfer.
8. Scanning devices are input devices used for direct data entry from the source document into the computer system. With the help of the scanner, you can capture your images and documents and convert them into digital formats for easy storage on your computer.
9. Normally, the output is produced on a display unit like a computer monitor or can be printed through a printer on paper. At times, speech outputs and mechanical outputs are also used for some specific applications.
10. Laser printers work on the same printing technology as photocopiers, using static electricity and heat to print with a high quality powder substance known as toner.

11. Cache memories are small, fast memories placed between the CPU and the main memory. They are faster than the main memory with access times closer to the speed of the CPU.
12. A disk drive is a peripheral device used to store and collect information. It can be removable or fixed, high capacity or low capacity, fast or slow speed, and magnetic or optical.

NOTES

1.8 SUMMARY

- A computer primarily constitutes three integral components, viz. input device, central processing unit (CPU) and output device.
- Any computer system essentially consists of three important parts, namely, input device, central processing unit (CPU) and output device. You have also learnt about the history of computers as well as the various generations of computer.
- The history of computer development can be divided into distinct phases which are often referred to as generations of computing devices. 'Generation' in computer terminology is a 'step' in technology.
- Each generation of computers is characterized by a major technological development that fundamentally changes the way computers operate, resulting in increasingly smaller, cheaper, more powerful, efficient and reliable devices with decreasing energy consumption and lesser generation of heat.
- Hybrid computers are a mixture of digital and analog computers. A hybrid computer uses the best characteristics of digital and analog computers. It helps the user to process both continuous and discrete data. Hybrid computers are generally used for weather forecasting and industrial process control.
- A PC is a small single-user microprocessor-based computer that sits on your desktop and is generally used at homes, offices, and schools. As the name implies, PCs were mainly designed to meet the personal computing needs of individuals.
- The central processing unit (CPU) is the most important component of a digital computer that interprets the instructions, and processes the data contained in computer programs.
- The ALU provides arithmetic and logic operations. This means that when the control unit encounters an instruction that involves an arithmetic operation (add, subtract, multiply, divide) or a logic operation (equal to, less than, greater than), it passes the control to the ALU.
- The I/O devices that provide a means of communication between the computer and the outside world are known as peripheral devices. This is because they surround the CPU and the memory of a computer system.

NOTES

- Keyboard devices allow input into the computer system by pressing a set of keys mounted on a board, connected to the computer system. Keyboard devices are typically classified as general-purpose keyboards and special-purpose keyboards.
- A mouse is a small input device used to move the cursor on a computer screen to give instructions to the computer and to run programs and applications. It can be used to select menu commands, move icons, size windows, start programs, close windows, etc.
- Scanning devices are input devices used for direct data entry from the source document into the computer system. With the help of the scanner you can capture your images and documents and convert it into digital formats for easy storage on your computer.
- An electromechanical device that accepts data from the computer and translates it into a form that can be understood by the external world is known as an output device.
- Storage devices may be classified according to their: storage capacity, storage cost, access time and access mode. Storage devices may be divided into primary storage devices like the RAM and the ROM and secondary storage device like the cache memory, magnetic tapes, magnetic disks, hard disks, optical storage devices like the optical disks, CD-ROMs, WORMs, erasable optical disk, DVDs, etc.

1.9 KEY TERMS

- **Inputting:** It is the process wherein the user feeds in the set of commands or instructions to process data into the computer system.
- **Storing:** The process of recording data and information so that it can be retrieved for use whenever required.
- **Processing:** Performing arithmetic or logical operations on data to convert them into useful information.
- **Outputting:** The process of providing results to the user which in the form of visual display and/or printed reports.
- **Controlling:** Directing the sequence and the manner in which all these previous tasks are carried out.
- **Hybrid Computers:** It is the combination of digital and analog computers and uses the best features of both.
- **Tablet PC:** It is a mobile computer that looks like a notebook or a small writing slate but uses a stylus pen or your finger tip to write on the touch screen.
- **Personal Digital Assistant (PDA):** It is a small, palm sized, hand-held computer which has a small color touch screen with audio and video features.

- **CPU:** The control unit and arithmetic logic unit are together known as the central processing unit. It is the brain of any computer system.
- **Mouse:** A small input device used to move the cursor on a computer screen to give instructions to the computer and to run programs and applications.
- **Touch Pads:** The input devices that are touch sensitive and take user input to manage onscreen pointer and other functions.
- **Light Pen:** A small input device used to select and display objects on a screen.
- **Trackball:** A pointing device that is like an inverted mouse.
- **Scanning Devices:** Input devices used for direct data entry from the source document into the computer system.

NOTES

1.10 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is a computer?
2. Name the stages of evolution of computers.
3. What is a digital computer?
4. What are supercomputers?
5. Define the term arithmetic logic unit.
6. State about the instructions sequencing.
7. Define the term touch pads.
8. What are the basic criteria for classifying printers?
9. State about the magnetic tape.

Long-Answer Questions

1. Describe the basic tasks that can be performed by a computer system.
2. Briefly describe the various generations of the computer.
3. Discuss about the various types of computer with the help of appropriate examples.
4. Explain how the control unit coordinates with the various components of a computer system.
5. Briefly explain the computer architecture with the help of diagram.
6. Discuss briefly input/output devices with the help of appropriate examples.
7. Briefly describe the functioning of a laser printer giving appropriate examples.
8. What are magnetic disks? Discuss its functions.

1.11 FURTHER READING

NOTES

Shepherd, Robert D. *Introduction to Computers and Technology*. EMC/Paradigm Publishing, 1998.

Kumar, A. Anand. *Fundamentals of Digital Circuits*. New Delhi: Prentice-Hall of India, 2003.

Yarborough, John M. *Digital Logic Applications and Design*. Thomson Learning, 1996.

Roth, Charles. *Fundamentals of Logic Design*. Thomson Learning, 2001.

Mano, Morris. *Digital Logic and Computer Design*. New Delhi: Prentice-Hall of India, 1979.

UNIT 2 CONCEPTS OF DATA REPRESENTATION AND SOFTWARE

NOTES

Structure

- 2.0 Introduction
- 2.1 Objectives
- 2.2 Data Representation
 - 2.2.1 Decimal Number System
 - 2.2.2 Binary Number System
 - 2.2.3 Inter-Conversion between two Different Number System
- 2.3 Software Concepts
 - 2.3.1 Software Qualities
 - 2.3.2 Types of Software
 - 2.3.3 Programming Language
 - 2.3.4 Object Oriented Language
- 2.4 History and Evolution of Operating Systems
 - 2.4.1 Functions of OS
 - 2.4.2 Operating System as Resource Manager
 - 2.4.3 Types of Operating System
 - 2.4.4 BIOS
- 2.5 Answers to ‘Check Your Progress’
- 2.6 Summary
- 2.7 Key Terms
- 2.8 Self-Assessment Questions and Exercises
- 2.9 Further Reading

2.0 INTRODUCTION

On hearing the word number, we immediately think of the familiar decimal number system with its ten digits: 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9. These are called Arabic numerals. Our present number system provides modern mathematicians and scientists with great advantages over those of previous civilizations, and is an important factor in our advancement. Since fingers are the most convenient tools nature has provided, human beings use them in counting. The decimal number system followed naturally from this usage.

A computer cannot operate without any instructions and is based on a logical sequence of instructions in order to perform a function. These instructions are known as a ‘computer program’, and constitute the computer software. System software constitutes all the programs, languages and documentation provided by the manufacturer in the computer. These programs provide the user with an access to the system so that he can communicate with the computer and write or develop his own programs. On the other hand, users install specific software programs based on their requirements.

An operating system (OS) is the software on a computer that manages the way different programs use the hardware and regulates the ways that are used to

NOTES

control the computer. In other words, an operating system is a primary control program that manages all the functions and applications of a computer. The three main components of an OS are control programs, system service programs and utility programs.

In this unit, you will study about the data representation, decimal and binary system, conversion of binary, octal, hexadecimal, software, types of software, software qualities, programming languages, classification of high level language, object oriented languages, history and evolution of operating system, function of an operating system, types of operating system, operating system as resource manager, BIOS.

2.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the basics of decimal and binary system
- Describe the conversions of binary octal hexadecimal systems
- Explain the concepts of software
- Discuss the various types of software
- Explain about the software quality and programming language
- Define object oriented languages
- Understand the history and evolution of software
- Describe the function of an operating system
- Explain the various types of operating system
- Define the significance of BIOS

2.2 DATA REPRESENTATION

2.2.1 Decimal Number System

The number system which utilizes ten distinct digits, i.e., 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9 is known as the decimal number system. It represents numbers in terms of groups of ten, as shown in Figure 2.1.

We would be forced to stop at 9 or to invent more symbols if it were not for the use of positional notation. It is necessary to learn only ten basic numbers and the positional notational system in order to count any desired figure.

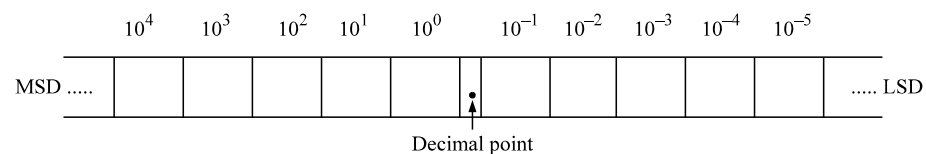


Fig. 2.1 Decimal Position Values as Powers of 10

NOTES

Decimal number system has a base or radix of 10. Each of the ten decimal digits 0 through 9 has a place value or weight depending on its position. The weights are units, tens, hundreds, and so on. The same can be written as the power of its base as $[10^0, 10^1, 10^2, 10^3 \dots]$, etc. Thus, the number 1993 represents quantity equal to $1000 + 900 + 90 + 3$. Actually, this should be written as $\{1 \times 10^3 + 9 \times 10^2 + 9 \times 10^1 + 3 \times 10^0\}$. Hence, 1993 is the sum of all digits multiplied by their weights. Each position has a value ten times greater than the position to its right.

Example 2.1 Show the representation of the number 379.

Solution:

100	10	1
10^2	10^1	10^0
3	7	9
3×100	$+ 7 \times 10$	$+ 9 \times 1$

$\therefore 379_{10} = 3 \times 100 + 7 \times 10 + 9 \times 1$
 $= 3 \times 10^2 + 7 \times 10^1 + 9 \times 10^0$

In this example, 9 is the least significant digit (LSD) and 3 is the most significant digit (MSD).

Example 2.2 Show the representation of the number 1936.469.

Solution: $1936.469_{10} = 1 \times 10^3 + 9 \times 10^2 + 3 \times 10^1 + 6 \times 10^0 + 4 \times 10^{-1}$
 $+ 6 \times 10^{-2} + 9 \times 10^{-3}$
 $= 1000 + 900 + 30 + 6 + 0.4 + 0.06 + 0.009 = \mathbf{1936.469}$

It is seen that powers are numbered to the left of the decimal point starting with 0 and to the right of the decimal point starting with -1 .

The general rule for representing numbers in the decimal system by using positional notation is as follows:

$$a_n a_{n-1} \dots a_2 a_1 a_0 = a_n 10^n + a_{n-1} 10^{n-1} + \dots + a_2 10^2 + a_1 10^1 + a_0 10^0$$

Here, n is the number of digits to the left of the decimal point.

2.2.2 Binary Number System

A number system that uses only two digits, 0 and 1, is called the **binary number system**. The binary number system is also called a **base two system**. The two symbols, 0 and 1, are known as **bits** (binary digits).

The binary system groups numbers by twos and by powers of two, shown in Figure 2.2. The word binary comes from a Latin word meaning two at a time.

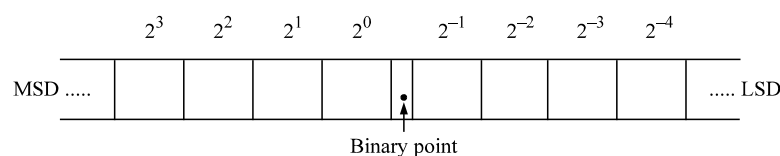


Fig. 2.2 Binary Position Values as a Power of 2

The weight or place value of each position can be expressed in terms of 2, and as $2^0, 2^1, 2^2$, etc. The least significant digit has a weight of $2^0 (= 1)$. The

NOTES

second position to the left of the least significant digit is multiplied by $2^1 (= 2)$. The third position has weight equal to $2^2 (= 4)$. Thus, the weights are in the ascending powers of 2 or 1, 2, 4, 8, 16, 32, 64, 128, etc.

The numeral 10_{two} (one, zero, base two) stands for two, the base of the system.

In binary counting, single digits are used for none and one. Two-digit numbers are used for 10_{two} and 11_{two} [2 and 3 in decimal numerals]. For the next counting number, 100_{two} (4 in decimal numerals) three digits are necessary. After 111_{two} (7 in decimal numerals) four digit numerals are used until 1111_{two} (15 in decimal numerals) is reached, and so on. In a binary numeral, every position has a value two times the value of the position to its right.

A binary number with 4 bits is called a nibble and a binary number with eight bits is known as a byte.

Example 2.3 Show the representation of the number 1011_2 .

Solution:
$$1011_2 = 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0$$

$$= 1 \times 8 + 0 \times 4 + 1 \times 2 + 1 \times 1$$

$$\therefore 1011_2 = 8 + 0 + 2 + 1 = 11_{10}$$

In general,

$$[b_n b_{n-1} \dots b_2 b_1 b_0]_2 = b_n 2^n + b_{n-1} 2^{n-1} + \dots + b_2 2^2 + b_1 2^1 + b_0 2^0$$

Example 2.4 Show the representation of the binary number 10101.011_2 .

Solution:

1	0	1	0	1	.	0	1	1
2^4	2^3	2^2	2^1	2^0	.	2^{-1}	2^{-2}	2^{-3}
(MSD)						(LSD)		

$$\therefore 10101.011_2 = 1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$$

$$+ 0 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3}$$

$$= 16 + 0 + 4 + 0 + 1 + 0 + 0.25 + 0.125 = 21.375_{10}$$

In each binary digit, the value increases in powers of two starting with 0 to the left of the binary point and decreases to the right of the binary point starting with power -1 .

Why is a Binary Number System is Used?

Binary number system is used in digital computers because all electrical and electronic circuits can be made to respond to the two states concept. For instance, a switch can be either opened or closed, only two possible states exist. A transistor can be made to operate either in cut-off or saturation; a magnetic tape can be either magnetized or non-magnetized; a signal can be either HIGH or LOW; a punched tape can have a hole or no hole. In all these cases, each device is operated in any one of the two possible states and the intermediate condition does not exist. Thus, 0 can represent one of the states and 1 can represent the

other. Hence, binary numbers are convenient to use in analysing or designing digital circuits.

Binary fractions

A binary fraction can be represented by a series of 1s and 0s to the right of a binary point. The weights of digit positions to the right of the binary point are given by 2^{-1} , 2^{-2} , 2^{-3} , and so on.

Example 2.5 Show the representation of the binary fraction 0.1011.

$$\begin{aligned} \text{Solution: } 0.1011 &= 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} + 1 \times 2^{-4} \\ &= 1 \times 0.5 + 0 \times 0.25 + 1 \times 0.125 + 1 \times 0.0625 \\ 0.1011_2 &= \mathbf{0.6875}_{10} \end{aligned}$$

Mixed numbers

Mixed numbers contain both integer and fractional parts. The weights of mixed numbers are

$$\begin{array}{ccccccc} 2^3 & 2^2 & 2^1 & \cdot & 2^{-1} & 2^{-2} & 2^{-3}, \text{ etc.} \\ & & & \uparrow & & & \\ & & & \text{Binary point} & & & \end{array}$$

Example 2.6 Show the representation of the mixed binary number 1011.101.

$$\begin{aligned} \text{Solution: } 1011.101_2 &= 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} + 0 \times 2^{-2} \\ &\quad + 1 \times 2^{-3} \\ &= 1 \times 8 + 0 \times 4 + 1 \times 2 + 1 \times 1 + 1 \times 0.5 + 0 \times 0.25 + 1 \\ &\quad \times 0.125 \\ \therefore [1011.101]_2 &= \mathbf{[11.625]}_{10} \end{aligned}$$

When different number systems are used, it is customary to enclose the number within big brackets and the subscripts indicate the type of the number system.

2.2.3 Inter-Conversion between two Different Number System

1. Binary to Decimal Conversion

A binary number can be converted into decimal number by multiplying the binary 1 or 0 by the weight corresponding to its position and adding all the values.

Example 2.7 Convert the binary number 110111 to decimal number.

$$\begin{aligned} \text{Solution: } 110111_2 &= 1 \times 2^5 + 1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 \\ &= 1 \times 32 + 1 \times 16 + 0 \times 8 + 1 \times 4 + 1 \times 2 + 1 \times 1 \\ &= 32 + 16 + 0 + 4 + 2 + 1 \\ &= \mathbf{55}_{10} \end{aligned}$$

We can streamline binary to decimal conversion by the following procedure:

Step 1. Write the binary, i.e., all its bits in a row.

NOTES

NOTES

Step 2. Write 1, 2, 4, 8, 16, 32, ..., directly under the binary number working from right to left.

Step 3. Omit the decimal weight which lies under zero bits.

Step 4. Add the remaining weights to obtain the decimal equivalent.

The same method is used for binary fractional number.

Table 2.1 lists the binary numbers from 0000 to 10000. Table 2.2 lists powers of 2 and their decimal equivalents and the number of K. The abbreviation K stands for $2^{10} = 1024$. Therefore, 1K = 1024, 2K = 2048, 3K = 3072, 4K = 4096, and so on. Many personal computers have 64K memory, this means that computers can store up to 65,536 bytes in the memory section.

Table 2.1 Binary Numbers

Decimal	Binary
0	0
1	01
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001
10	1010
11	1011
12	1100
13	1101
14	1110
15	1111
16	10000

Table 2.2 Powers of 2

Powers of 2	Equivalent	Abbreviation
2^0	1	
2^1	2	
2^2	4	
2^3	8	
2^4	16	
2^5	32	
2^6	64	
2^7	128	
2^8	256	
2^9	512	
2^{10}	1024	1K
2^{11}	2048	2K
2^{12}	4096	4K
2^{13}	8192	8K
2^{14}	16384	16K
2^{15}	32768	32K
2^{16}	65536	64K

2. Decimal to Binary Conversion

There are several methods for converting a decimal number to a binary number. The first method is simply to subtract values of powers of 2 which can be subtracted from the decimal number until nothing remains. The value of the highest power of 2 is subtracted first, then the second highest and so on.

Example 2.8 Convert the decimal integer 29 to the binary number system.

Solution: First the value of the highest power of 2 which can be subtracted from 29 is found. This is $2^4 = 16$.

Then, $29 - 16 = 13$

The value of the highest power of 2 which can be subtracted from 13, is 2^3 , then $13 - 2^3 = 13 - 8 = 5$. The value of the highest power of 2 which can be subtracted from 5, is 2^2 . Then $5 - 2^2 = 5 - 4 = 1$. The remainder after subtraction is 1 or 2^0 . Therefore, the binary representation for 29 is given by,

$$29_{10} = 2^4 + 2^3 + 2^2 + 2^0 = 16 + 8 + 4 + 0 \times 2 + 1$$

$$= 1 \quad 1 \quad 1 \quad 0 \quad 1$$

$$[29]_{10} = [11101]_2$$

Similarly, $[25.375]_{10} = 16 + 8 + 1 + 0.25 + 0.125$

$$= 2^4 + 2^3 + 0 + 0 + 2^0 + 0 + 2^{-2} + 2^{-3}$$

$$[25.375]_{10} = [11011.011]_2$$

This is a laborious method for converting numbers. It is convenient for small numbers and can be performed mentally, but is less used for larger numbers.

Double-Dabble Method

A popular method known as **double-dabble method**, also known as divide-by-two method, is used to convert a large decimal number into its binary equivalent. In this method, the decimal number is repeatedly divided by 2, and the remainder after each division is used to indicate the coefficient of the binary number to be formed. Notice that the binary number derived is written from the bottom up.

Example 2.9 Convert 199_{10} into its binary equivalent.

Solution:

$199 \div 2 = 99 + \text{remainder}$	1	(LSB)	
$99 \div 2 = 49 + \text{remainder}$	1	↑	
$49 \div 2 = 24 + \text{remainder}$	1		
$24 \div 2 = 12 + \text{remainder}$	0		
$12 \div 2 = 6 + \text{remainder}$	0		
$6 \div 2 = 3 + \text{remainder}$	0		
$3 \div 2 = 1 + \text{remainder}$	1		
$1 \div 2 = 0 + \text{remainder}$	1		(MSB)

The binary representation of 199 is, therefore, 11000111. Checking the result we have,

$$[11000111]_2 = 1 \times 2^7 + 1 \times 2^6 + 0 \times 2^5 + 0 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0$$

$$= 128 + 64 + 0 + 0 + 0 + 4 + 2 + 1$$

$$\therefore [11000111]_2 = [199]_{10}$$

Notice that the first remainder is the LSB and last remainder is the MSB. This method will not work for mixed numbers.

Decimal Fraction to Binary

The conversion of decimal fraction to binary fractions may be accomplished by using several techniques. Again, the most obvious method is to subtract the highest value of the negative power of 2, which may be subtracted from the decimal fraction. Then, the next highest value of the negative power of 2 is subtracted from the remainder of the first subtraction, and this process is continued until there is no remainder or to the desired precision.

NOTES

NOTES

Example 2.10 Convert decimal 0.875 to a binary number.

Solution: $0.875 - 1 \times 2^{-1} = 0.875 - 0.5 = 0.375$
 $0.375 - 1 \times 2^{-2} = 0.375 - 0.25 = 0.125$
 $0.125 - 1 \times 2^{-3} = 0.125 - 0.125 = 0$

$\therefore [0.875]_{10} = [0.111]_2$

A much simpler method of converting longer decimal fractions to binary consists of repeatedly multiplying by 2 and recording any carriers in the integer position.

3. Octal to Decimal Conversion

An octal number can be easily converted to its decimal equivalent by multiplying each octal digit by its positional weight.

Example 2.11 Convert $(376)_8$ to decimal number.

Solution: The process is similar to binary to decimal conversion except that the base here is 8.

$$[376]_8 = 3 \times 8^2 + 7 \times 8^1 + 6 \times 8^0$$

$$= 3 \times 64 + 7 \times 8 + 6 \times 1 = 192 + 56 + 6 = [254]_{10}$$

The fractional part can be converted into decimal by multiplying it by the negative powers of 8.

4. Decimal to Octal Conversion

The methods used for converting a decimal number to its octal equivalent are the same as those used to convert from decimal to binary. To convert a decimal number to octal, we progressively divide the decimal number by 8, writing down the remainders after each division. This process is continued until zero is obtained as the quotient, the first remainder being the LSD.

The fractional part is multiplied by 8 to get a carry and a fraction. The new fraction obtained is again multiplied by 8 to get a new carry and a new fraction. This process is continued until the number of digits have sufficient accuracy.

Example 2.12 Convert $[416.12]_{10}$ to octal number.

Solution: Integer part $416 \div 8 = 52 + \text{remainder } 0 \text{ (LSD)}$

$52 \div 8 = 6 + \text{remainder } 4$

$6 \div 8 = 0 + \text{remainder } 6 \text{ (MSD)}$

Fractional part $0.12 \times 8 = 0.96 = 0.96 \text{ with a carry of } 0$

$0.96 \times 8 = 7.68 = 0.68 \text{ with a carry of } 7$

$0.68 \times 8 = 5.44 = 0.44 \text{ with a carry of } 5$

$0.44 \times 8 = 3.52 = 0.52 \text{ with a carry of } 3$

$0.52 \times 8 = 4.16 = 0.16 \text{ with a carry of } 4$

$$\begin{array}{r}
 0.16 \times 8 = 1.28 = 0.28 \text{ with a carry of } 1 \\
 0.28 \times 8 = 2.24 = 0.24 \text{ with a carry of } 2 \\
 0.24 \times 8 = 1.92 = 0.92 \text{ with a carry of } 1
 \end{array}$$

$$\therefore [416.12]_{10} = [640.07534121]_8$$

This gives approximate answer. If more accuracy is needed, then continue multiplying by 8 until you have as many digits as necessary for your application.

5. Binary to Octal Conversion

The simplest procedure is to use the **binary-triplet method**. The binary digits are grouped into groups of three on each side of the binary point with zeros added on either side if needed to complete a group of three. Then, each group of 3 bits is converted to its octal equivalent. Note that the highest digit in the octal system is 7.

Example 2.13 Convert $[11001.101011]_2$ to octal number.

Solution: Binary $[11001.101011]_2$

Divide into groups of 3 bits	011	001	.	101	011
	↓	↓		↓	↓
	3	1		5	3

Note that a zero is added to the left-most group of the integer part. Thus, the desired octal conversion is $[31.53]_8$.

6. Hexadecimal to Binary Conversion

Hexadecimal numbers can be converted into binary numbers by converting each hexadecimal digit to 4 bit binary equivalent using the specific code. If the hexadecimal digit is 3, it should not be represented by 2 bits $[11]_2$, but it should be represented by 4 bits as $[0011]_2$.

Example 2.14 Convert $[EC2]_{16}$ to binary number.

Solution:	Hexadecimal number	E	C	2
		↓	↓	↓
	Binary Equivalent	1110	1100	0010
\therefore	$[EC2]_{16} =$	[1110	1100	0010]

7. Binary to Hexadecimal Conversion

Conversion from binary to hexadecimal is easily accomplished by partitioning the binary number into groups of four binary digits, starting from the binary point to the left and to the right. It may be necessary to add zero to the last group, if it does not end in exactly 4 bits. Each group of 4 bits binary must be represented by its hexadecimal equivalent.

Example 2.15 Convert $[10011100110]_2$ to hexadecimal number.

Solution: Binary number $[10011100110]_2$

Grouping the above binary number into 4-bits, we have

NOTES

	0100	1110	0110	
Hexadecimal equivalent	↓	↓	↓	
	4	E	6	

NOTES

$$\therefore [10011100110]_2 = [4E6]_{16}$$

8. Hexadecimal to Decimal Conversion

As in octal, each hexadecimal number is multiplied by the powers of 16, which represents the weight according to its position and finally, desired result is get by adding all the values.

Another way of converting a hexadecimal number into its decimal equivalent is to first convert the hexadecimal number to binary and then convert from binary to decimal.

Example 2.16 Convert $[B6A]_{16}$ to decimal number.

Solution: Hexadecimal number $[B6A]_{16}$

$$\begin{aligned} [B6A]_{16} &= B \times 16^2 + 6 \times 16^1 + A \times 16^0 \\ &= 11 \times 256 + 6 \times 16 + 10 \times 1 = 2816 + 96 + 10 = [2922]_{10} \end{aligned}$$

9. Hexadecimal to Octal Conversion

This can be accomplished by first writing down the 4-bit binary equivalent of hexadecimal digit and then partitioning it into groups of 3 bits each. Finally, the 3-bit octal equivalent is written down.

Example 2.17 Convert $[2AB.9]_{16}$ to octal number.

Solution: Hexadecimal number	2	A	B	.	9	
	↓	↓	↓		↓	
4 bit numbers	0010	1010	1011	.	1001	
3 bit pattern	001	010	101	011	.	100 100
	↓	↓	↓	↓		↓ ↓
Octal number	1	2	5	3	.	4 4
\therefore	$[2AB.9]_{16} = [1253.44]_8$					

10. Octal to Hexadecimal Conversion

It is the reverse of the above procedure. First the 3-bit equivalent of the octal digit is written down and then partitioned into groups of 4 bits, then the hexadecimal equivalent of that group is written down.

Example 2.18 Convert $[16.2]_8$ to hexadecimal number.

Solution: Octal number	1	6	.	2
	↓	↓		↓
3 bit binary	001	110	.	010
4 bit pattern		1110	.	0100
		↓		↓
Hexadecimal		E	.	4
\therefore	$[16.2]_8 = [E.4]_{16}$			

11. Octal to Decimal Conversion

An octal number can be easily converted into its decimal equivalent by multiplying each octal digit by its positional weight.

Example 2.19 Convert $(376)_8$ into its decimal form.

Solution: The process is similar to binary to decimal conversion except the base here is 8.

$$\begin{aligned} [376]_8 &= 3 \times 8^2 + 7 \times 8^1 + 6 \times 8^0 \\ &= 3 \times 64 + 7 \times 8 + 6 \times 1 = 192 + 56 + 6 = [254]_{10} \end{aligned}$$

The fractional part can be converted into decimal by multiplying it by the negative powers of 8 as shown in the following examples.

Example 2.20 Convert $(0.4051)_8$ into its decimal form.

Solution: $[0.4051]_8 = 4 \times 8^{-1} + 0 \times 8^{-2} + 5 \times 8^{-3} + 1 \times 8^{-4}$

$$= 4 \times \frac{1}{8} + 0 \times \frac{1}{64} + 5 \times \frac{1}{512} + 1 \times \frac{1}{4096}$$

$$\therefore [0.4051]_8 = [0.5100098]_{10}$$

Example 2.21 Convert 6327.45_8 into its decimal form.

Solution: $[6327.45]_8 = 6 \times 8^3 + 3 \times 8^2 + 2 \times 8^1 + 7 \times 8^0 + 4 \times 8^{-1} + 5 \times 8^{-2}$

$$= 3072 + 192 + 16 + 7 + 0.5 + 0.078125$$

$$[6327.45]_8 = [3287.578125]_{10}$$

12. Decimal to Octal Conversion

The methods used for converting a decimal number to its octal equivalent are the same as those used to convert decimal to binary. To convert a decimal number to octal, we progressively divide the decimal number by 8, writing down the remainders after each division. This process is continued until zero is obtained as the quotient, the first remainder being the LSD.

The fractional part is multiplied by 8 to get a carry and a fraction. The new fraction obtained is again multiplied by 8 to get a new carry and a new fraction. This process is continued until the number of digits have sufficient accuracy.

Example 2.22 Convert $[416.12]_{10}$ into its octal form.

Solution: Integer part

$416 \div 8 = 52 + \text{remainder of}$	0	(LSD)	↑
$52 \div 8 = 6 + \text{remainder of}$	4		↑
$6 \div 8 = 0 + \text{remainder of}$	6	(MSD)	↑

Fractional part

$0.12 \times 8 = 0.96 = 0.96$ with a carry of	0	
$0.96 \times 8 = 7.68 = 0.68$ with a carry of	7	
$0.68 \times 8 = 5.44 = 0.44$ with a carry of	5	

NOTES

NOTES

$$\begin{aligned}
 0.44 \times 8 &= 3.52 = 0.52 \text{ with a carry of } 3 \\
 0.52 \times 8 &= 4.16 = 0.16 \text{ with a carry of } 4 \\
 0.16 \times 8 &= 1.28 = 0.28 \text{ with a carry of } 1 \\
 0.28 \times 8 &= 2.24 = 0.24 \text{ with a carry of } 2 \\
 0.24 \times 8 &= 1.92 = 0.92 \text{ with a carry of } 1
 \end{aligned}$$

$$\therefore [416.12]_{10} = [640.07534121]_8$$

Example 2.23 Convert $[3964.63]_{10}$ into its octal form.

Solution: Integer part

$$\begin{aligned}
 3964 \div 8 &= 495 \text{ with a remainder of } 4 \text{ (LSD)} \\
 495 \div 8 &= 61 \text{ with a remainder of } 7 \\
 61 \div 8 &= 7 \text{ with a remainder of } 5 \\
 7 \div 8 &= 0 \text{ with a remainder of } 7 \text{ (MSD)}
 \end{aligned}$$

$$\therefore [3964]_{10} = [7574]_8$$

Fractional part

$$\begin{aligned}
 0.63 \times 8 &= 5.04 = 0.04 \text{ with a carry of } 5 \\
 0.04 \times 8 &= 0.32 = 0.32 \text{ with a carry of } 0 \\
 0.32 \times 8 &= 2.56 = 0.56 \text{ with a carry of } 2 \\
 0.56 \times 8 &= 4.48 = 0.48 \text{ with a carry of } 4 \\
 0.48 \times 8 &= 3.84 = 0.84 \text{ with a carry of } 3 \text{ [LSD]}
 \end{aligned}$$

$$\therefore [3964.63]_{10} = [7574.50243]_8$$

Note that the first carry is the MSD of the fraction. More accuracy can be obtained by continuing the process to obtain octal digits.

13. Octal to Binary Conversion

Since 8 is the third power of 2, we can convert each octal digit into its three-bit binary form and from binary to octal form. All three-bit binary numbers are required to represent the eight octal digits of the octal form. Octal number system is often used in digital systems, especially for input/output applications. Each octal digit represented by three-bits is shown in Table 2.3.

Example 2.24 Convert $[675]_8$ into its binary form.

Solution:

Octal digit	6	7	5	
	↓	↓	↓	
Binary	110	111	101	
\therefore	$[675]_8 =$	$[110$	111	$101]_2$

Example 2.25 Convert $[246.71]_8$ into its binary form.

Solution:

Octal digit	2	4	6	.	7	1
	↓	↓	↓		↓	↓
Binary	010	100	110		111	001
\therefore	$[246.71]_8 =$	$[010$	100	110	$.$	$111 001]_2$

Table 2.3 Octal to Binary Conversion

Octal digit	Binary equivalent
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111
10	001 000
11	001 001
12	001 010
13	001 011
14	001 100
15	001 101
16	001 110
17	001 111

NOTES

14. Binary to Octal Conversion

The simplest procedure is to use **binary-triplet method**. The binary digits are grouped into groups of three on each side of the binary point with zeroes added on either side if needed to complete a group of three. Then, each group of three bits is converted into its octal equivalent. Note that the highest digit in the octal system is 7.

Example 2.26 Convert $[11001.101011]_2$ into its octal form.

Solution: Binary 11001.101011

Divide into groups of three bits 011 001 . 101 011
 ↓ ↓ ↓ ↓
 3 1 5 3

Note that a zero is added to the left-most group of the integer part. Thus, the desired octal conversion is $[31.53]_8$.

Example 2.27 Convert $[11101.101101]_2$ to octal.

Solution: Binary 11101.101101

Divide into groups of three bits 011 101 . 101 101
 ↓ ↓ ↓ ↓
 3 5 5 5

∴ $[11101.101101]_2 = [35.55]_8$

NOTES

2.3 SOFTWARE CONCEPTS

A computer cannot operate without any instructions and is based on a logical sequence of instructions in order to perform a function. These instructions are known as a 'computer program', and constitute the computer software. The sequences of instructions are based on algorithms that provide the computer with instructions on how to perform a function. Thus, it is impossible for a computer to process without software, a term attributed to John W. Tukey in 1958.

Different kinds of software designs have been developed for particular functions. Popular computer software include interpreter, assembler, compiler, operating systems, networking, word processing, accounting, presentation, graphics, computer games, etc. The computer software is responsible for converting the instructions in a program into a machine language facilitating their execution.

Software engineers develop computer software depending on basic mathematical analysis and logical reasoning. Before implementation, the software undergoes a number of tests. Thus, the programming software allows you to develop the desired instruction sequences, whereas in the application software the instruction sequences are predefined. Computer software can function from only a few instructions to millions of instructions; for example, a word processor or a web browser. Figure 2.3 shows how software interacts between user and computer system.

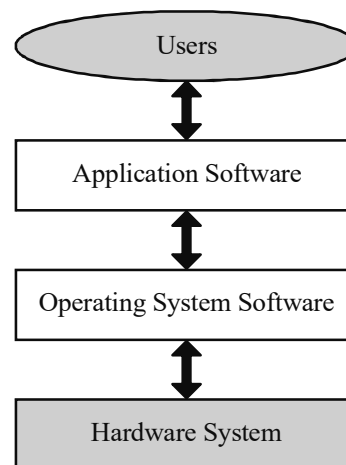


Fig. 2.3 Interaction of Software between User and a Computer System

On the functional basis, software is categorized as follows:

- **System software:** It helps in the proper functioning of computer hardware. It includes device drivers, operating systems, servers and utilities.
- **Programming software:** It provides tools to help a programmer in writing computer programs and software using various programming languages. It includes compilers, debuggers, interpreters, linkers, text editors and an integrated development environment (IDE).

- **Application software:** It helps the end users to complete one or more specific tasks. The specific applications include industrial automation, business software, computer games, telecommunications, databases, educational software, medical software and military software.

NOTES

2.3.1 Software Qualities

Following points describe how quality of softwares suffer if proper guidelines are not followed :

1. Software requirements are the foundation for measuring the quality. If requirements are not properly specified there will be a lack of quality.
2. If the development criteria defined by specified standards to guide the manner in which software is engineered are not followed, then this surely results in lack of quality.
3. Software quality is suspect, if software conforms to its explicit requirements but fails to meet a set of implicit requirements that often goes unmentioned.

Software quality is a combination of various factors that will vary across different applications and the customers who request them.

McCall's Quality Factors

Software quality can be affected by two categories of factors:

- (1) Factors that can be directly measured (e.g., defects per function-point).
- (2) Factors that can be measured only indirectly (e.g., usability or maintainability).

Measurement must occur in each case. One must compare the software (documents, programs, and data) to some data and arrive at an indication of quality.

McCall, Richards, and Walters propose a useful categorization of factors that affect software quality. These software quality factors, shown in Figure 2.4, focus on three important aspects of a software product-its operational characteristics, its ability to undergo change and its adaptability to new environments.

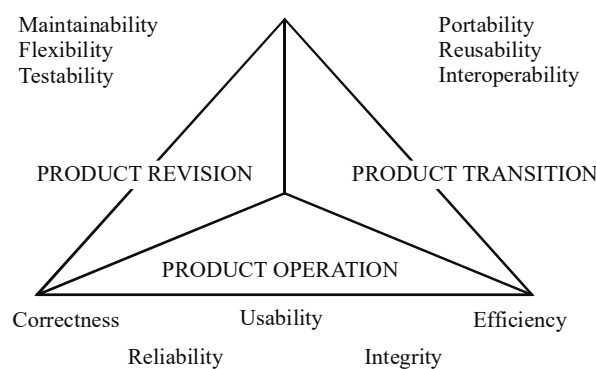


Fig. 2.4 McCall's Software Quality Factors

Referring to Figure 2.4, following descriptions are provided by McCall:

- **Efficiency:** The amount of computing resources and code required by a program to perform its function.

NOTES

- *Correctness*: The level up to which a program satisfies its specification and fulfills the customer's objectives.
- *Usability*: Efforts needed to learn, operate, prepare input and interpret output of a program.
- *Reliability*: The level to which a program can be expected to perform its intended function without failure.
- *Flexibility*: Efforts required in modifying an operational program.
- *Portability*: Efforts required in transferring the program from one hardware and/or software system environment to another.
- *Integrity*: Efforts required for Controlling the unauthorized access to software or data by persons.
- *Maintainability*: Effort required in locating and fixing an error in a program.
- *Testability*: Efforts required for testing a program to ensure that it performs its intended function.
- *Reusability*: Extent to which a program [or parts of a program] can be reused in other applications.
- *Interoperability*. Efforts required to couple one system to another.

To develop direct measures of these quality factors in some case is very difficult or even impossible. Therefore, for each of the factors a set of metrics are defined and used to develop expressions according to the following relationship:

$$Fq = c1 \times m1 + c2 \times m2 + \dots + cn \times mn$$

where:

cn are regression coefficients.

Fq is a software quality factor.

mn are the metrics that affect the quality factor.

Unfortunately, many of the metrics can be measured only subjectively. The metrics that is used to 'grade' specific attributes of the software may be in the form of a checklist. The grading scheme proposed by McCall and others is a 0 for low to 10 for high scale. The following metrics are used in the grading scheme:

- *Completeness*: It is the extent up to which full implementation of required function has been achieved.
- *Data commonality*: Standard data structures and types are used throughout the program.
- *Communication commonality*: The level to which standard interfaces, bandwidth and protocols are used.
- *Auditability*: The ease with which conformance to standards can be checked.

- *Accuracy*: The precision of computations and control.
- *Conciseness*: The compactness of the program in terms of lines of code.
- *Consistency*: The use of uniform design and documentation techniques throughout the software development project.
- *Error tolerance*: The damage that occurs when the program encounters an error.
- *Execution efficiency*: The run-time performance of a program.
- *Expandability*: The degree to which architectural, data or procedural design can be extended.
- *Generality*: The breadth of potential application of program components.
- *Hardware independence*: The degree to which the software is decoupled from the hardware on which it operates.
- *Instrumentation*: The degree to which the program monitors its own operation and identifies errors that do occur.
- *Modularity*: The functional independence of program components.
- *Operability*: The ease of operation of a program.
- *Security*: The availability of mechanisms that control or protect programs and data.
- *Self-documentation*: The degree to which the source code provides meaningful documentation.
- *Simplicity*: The degree to which a program can be understood without difficulty.
- *Software system independence*: The degree to which the program is independent of nonstandard programming language features, operating system characteristics and other environmental constraints.
- *Traceability*: The ability to trace a design representation or actual program component back to requirements.
- *Training*: The degree to which the software assists in enabling new users to apply the system.

The relationship between software quality factors and these metrics is shown in Table 2.4.

NOTES

Table 2.4 Quality Factors and Metrics

NOTES

Software quality metric / Quality factor	Correctness	Reliability	Efficiency	Integrity	Maintainability	Flexibility	Testability	Portability	Reusability	Interoperability	Usability
Audit ability				x			x				
Accuracy		x									
Communication commonality										x	
Completeness	x										
Complexity		x				x	x				
Concision			x		x	x					
Consistency	x	x			x	x					
Data commonality										x	
Error tolerance		x									
Execution efficiency			x								
Expandability						x					
Generality						x		x	x	x	
Hardware independence								x	x		
Instrumentation				x	x		x				
Modularity		x			x	x	x	x	x	x	
Operability			x								x
Security				x							
Self-documentation					x	x	x	x	x		
Simplicity		x			x	x	x				
System independence											
Traceability	x										
Training											x

FURPS

FURPS is set of software quality factors that have been developed by Hewlett-Packard. FURPS stands for functionality, usability, reliability, performance and supportability. The FURPS quality factors are defined as follows:

- *Functionality*: By evaluating the feature set and capabilities of the program, the generality of the delivered functions and the security of the overall system, functionality can be assessed.

- *Usability*: By considering human factors, consistency, overall aesthetics and documentation, usability can be assessed.
- *Reliability*: It is evaluated by measuring the mean-time-to-failure (MTTF), the frequency and severity of failure the ability to recover from failure, the accuracy of output results, and the predictability of the program.
- *Performance*: It is measured by processing speed, consumption of resources, response time, efficiency and throughput.
- *Supportability*: It combines the extensibility, adaptability, serviceability (these three attributes are better represented by a more common term, maintainability), testability, compatibility, configurability.

The FURPS quality factors can be used to establish quality metrics for each step in the software engineering process.

ISO 9126 Quality Factors

To identify the key quality attributes for computer software the ISO 9126 standard was developed. Six key quality attributes are identified by this standard. These attributes are described as follows:

- *Functionality*: The extent to which the software satisfies stated needs as indicated by the following subattributes: suitability, accuracy, interoperability, compliance and security.
- *Reliability*: The amount of time that the software is available for use as indicated by the following subattributes: maturity, fault tolerance, recoverability.
- *Usability*: The degree to which the software is easy to use as indicated by the following subattributes: understandability, learnability, operability.
- *Efficiency*: The degree to which the software makes optimal use of system resources as indicated by the following subattributes: time behaviour, resource behaviour.
- *Maintainability*: The ease with which repair may be made to the software as indicated by the following subattributes: analysability, changeability, stability, testability.
- *Portability*: The ease with which the software can be transposed from one environment to another as indicated by the following subattributes: adaptability, installability, conformance, replaceability.

ISO 9126 factors provide a basis for indirect measures and an excellent checklist for assessing the quality of a system.

2.3.2 Types of Software

Today, software is a significant aspect of almost all fields including business, education, medicine, etc. The basic requirement for software is a distinct set of procedures. Thus, software can be used in any domain that can be described in logical and related steps and every software is developed with the aim of catering

NOTES

NOTES

to a particular objective, such as data processing, information sharing, communication, etc. Software is based on the type of applications that are as follows:

- **System software:** This type of software is involved managing and controlling the operations of a computer system. System software is a group of programs rather than one program and is responsible for using computer resources efficiently and effectively. Operating system, for example, is system software, which controls the hardware, manages memory and multitasking functions and acts as an interface between applications programs and the computer.
- **Real-time software:** This is based on observing, analysing and controlling real life events as they occur. Manually, a real-time system guarantees a response to an external event within a specified period of time. The real-time software, for instance, is used for navigation in which the computer must react to a steady flow of new information without interruption. Most defence organizations all over the world use real time software to control their military hardware.
- **Business software:** This kind of software is functional in the domain of management and finance. The basic aspect of a business system comprises payroll, inventory, accounting and software that permits users to access relevant data from the database. These activities are usually performed with the help of specialized business software that facilitates efficient framework in the business operation and in management decisions.
- **Engineering and scientific software:** This software has developed as a significant tool used in the research and development of next generation technology. Applications, such as study of celestial bodies, study of undersurface activities and programming of orbital path for space shuttle, are heavily dependent on engineering and scientific software. This software is designed to perform precise calculations on complex numerical data that are obtained during real-time environment.
- **Artificial Intelligence (AI) software:** Certain problem-solving techniques are non-algorithmic in nature and primarily require this type of software. The solutions to such problems normally cannot be arrived at using computation or straightforward analysis. Such problems need particular problem-solving techniques including expert system, pattern recognition and game playing. Also, it constitutes various kinds of searching techniques, such as the application of heuristics. The function of (AI) is to add certain degree of intelligence into the mechanical hardware to have the desired work done in an agile manner.
- **Web-based software:** This category of software performs the function of an interface between the user and the Internet. There are various forms in which data is available online, such as text, audio or video format, linked with hyperlinks. For the retrieval of web pages from the Internet a web

browser is used, which is a web-based software. The software incorporates executable instructions written in special scripting languages, such as common gateway interface (CGI) or active server page (ASP). Apart from providing navigation on the web, this software also supports additional features that are useful while surfing the Internet.

- **Personal Computer (PC) software:** This software is primarily designed for personal use on a daily basis. The past few years have seen a marked increase in the personal computer software market from normal text editor to word processor and from simple paintbrush to advance image-editing software. This software is used mostly in almost every field, whether it is database management system, financial accounting package, or a multimedia based software. It has emerged as a versatile tool for daily life applications.

Software can also be classified in terms of the relationship between software users or software purchasers and software development.

- **Commercial off-the-shelf (COTS):** This comprises the software without any committed user before it is put up for sale. The software users have less or no contact with the vendor during development. It is sold through retail stores or distributed electronically. This software includes commonly used programs, such as word processors, spreadsheets, games, income tax programs, as well as software development tools, such as, software testing tools and object modelling tools.
- **Customized or bespoke:** This software is designed for a specific user, who is bound by some kind of formal contract. Software developed for an aircraft, for example, is usually done for a particular aircraft making company. They are not purchased ‘off-the-shelf’ like any word processing software.
- **Customized COTS:** In this classification, a user can enter into a contract with the software vendor to develop a COTS product for a special purpose, that is, software can be customized according to the needs of the user. Another growing trend is the development of COTS software components—the components that are purchased and used to develop new applications. The COTS software component vendors are essentially parts stores which are classified according to their application types. These types are listed as follows:
 - **Stand-alone software:** A software that resides on a single computer and does not interact with any other software installed in a different computer.
 - **Embedded software:** A software that pertains to the part of unique application involving hardware like automobile controller.
 - **Real-time software:** In this type of software the Operations are executed within very short time limits, often microseconds, e.g., radar software in air traffic control system.
 - **Network software:** In this type of software, software and its components interact across a network.

NOTES

NOTES

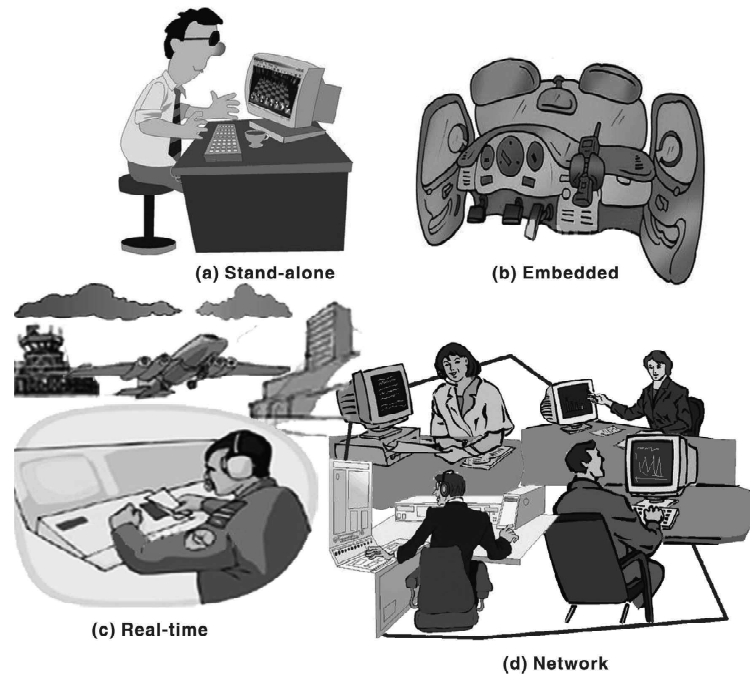


Fig. 2.5 Types of Customized COTS

2.3.3 Programming Language

A computer language essentially implies a language that is understandable to the computer. It is the computer's native language. Computer languages serve the same purpose as human languages. They are a means of communication. Let us understand the similarities and differences between computer languages and human languages.

Languages that we speak daily, such as English, Hindi, French or German are known as natural or human languages. It constitutes words and rules known as lexicon and syntax, respectively. These words are joined to make meaningful phrases according to the rules of the syntax. A computer language also consists of lexicon and syntax, i.e., characters, symbols and rules of usage that allow the user to communicate with the computer.

The primary difference between a natural language and computer language is that natural languages have a large set of words (vocabulary) to choose from while computer languages have a limited or restricted set of words. Thus, fewer words but more rules characterize a computer language.

All problems to be solved by the computer need to be broken down into discrete logical steps before the computer can execute them. The process of writing such instructions in a computer or programming language is called programming or coding.

Since as computer hardware has improved over the years, programming languages have also moved from machine-oriented languages (that used strings of binary 0s and 1s) to problem-oriented languages (that use common English terms). All computer languages can, however, be classified under the following categories:

- Machine language (first-generation language)

- Assembly language (second-generation language)
- High-Level language (third-generation language)

Classification of Computer Languages

Computer languages are classified as follows:

The computer can understand only a binary-based language. This is a combination of 0s and 1s. Instructions written using sequences of 0s and 1s constitute a are known as machine language. First-generation computers used programs written in machine language.

A major drawback of machine language is that it is highly complex and difficult to use. Also, it consumes a lot of time and requires a substantial effort on the part of the programmer. Thousands of machine language instructions are needed to carry out simple tasks, such as listing a few addresses for mails. Any instruction in machine language is divided into two components:

- (i) Command: Also called the 'operation code' or opcode including addition, multiplication, etc.
- (ii) Operand: Refers to the address of the data on which the function has to be performed.

A general machine language instruction is presented as follows:

OP Code	Operand
001	010001110

The number of operands varies with each computer and is therefore computer dependent.

It can be concluded that in order to develop computer programs in machine language, the programmer will be required to remember a lot of operation codes and addresses of the data items based on the storage location and also information regarding the internal structure of the computer. Thus, using machine language can be highly complicated and liable to errors. Identifying these errors and introducing changes had become increasingly difficult leading programmers to seek better options.

Assembly Language

The development of assembly language marked the beginning of the evolution of programming languages. In assembly language mnemonics (symbolic codes) were used to present operation codes as well as strings of characters to represent addresses. Instructions in assembly language may appear as follows:

Operation	Operation address
READ	M
ADD	L

Certain important facts about assembly language are as follows:

- Assembly language was designed to replace each machine code by an understandable mnemonic and each address with a simple alphanumeric

NOTES

NOTES

string. It was matched to the processor structure of a particular computer and was therefore (once again) machine dependent. This meant that programs written for a particular computer model could not be executed on another one. In other words, an assembly language program lacked portability.

- A program written in assembly language needs to be translated into machine language before the computer can execute it. This is done by a special program called 'Assembler' which takes every assembly language program and translates it into its equivalent machine code.
- The assembly language program is known as the source program, while the equivalent machine language program is known as the object program. It may be useful to know that the assembler is a system program supplied by the computer manufacturer. Second-generation computers used assembly language.
- The lack of portability of programs (written using machine or assembly languages) between various computer systems led to the development of high-level languages. Since they allowed a programmer to overlook a lot of low-level particulars of the hardware of the computer system, they were called high-level language programs.
- It was obvious that if the syntax, mnemonics and rule and regulations of the programming language were closer to the natural language, it would be easier for the programmer to program and the lesser the possibility of introducing errors (or bugs) into the program. Hence, third-generation languages, which were algorithmic and procedural, came into being in the mid-1950s. They were designed to solve a particular problem. They contained commands that are particularly suited to one type of application. For example, a number of languages were designed to process scientific or mathematical problems. Others emphasized on commercial applications. These languages varied very little between different computer systems, unlike machine or symbolic languages. But a compiler or an interpreter program was required to translate these machine codes. Once again, the high-level program is called the source code while its equivalent machine language program is referred to as the object code.
- Easy-to-learn feature, machine independence, easier maintenance and portability contributed to the popularity of high-level languages. Slow program execution was the main disadvantage since programs needed to be converted into machine language (by an interpreter or a compiler) before they could be executed.

A program written in assembly language needs to be translated into machine language before the computer can execute it. This is done by a special program called 'Assembler' which takes every assembly language program and translates it into its equivalent machine code. The assembly language program is known as the source program, while the equivalent machine language program is known as the object program. It may be useful to know that the assembler is a system program supplied by the computer manufacturer. Second-generation computers used assembly language.

NOTES

The lack of portability of programs (written using machine or assembly languages) between various computer systems led to the development of high-level languages. Since they allowed a programmer to overlook a lot of low-level particulars of the hardware of the computer system, they were called high-level language programs. And it was obvious that if the syntax, mnemonics and rule and regulations of the programming language were closer to the natural language, it would be easier for the programmer to program and the lesser the possibility of introducing errors (or bugs) into the program. Hence, third-generation languages, which were algorithmic and procedural, came into being in the mid-1950s. They were designed to solve a particular problem. They contained commands that are particularly suited to one type of application, for example, a number of languages were designed to process scientific or mathematical problems. Others emphasized on commercial applications. These languages varied very little between different computer systems, unlike machine or symbolic languages. But a compiler or an interpreter program was required to translate these machine codes. Once again, the high-level program is called the source code while its equivalent machine language program is referred to as the object code.

Easy-to-learn features, machine independence, easier maintenance and portability contributed to the popularity of high-level languages. Slow program execution was the main disadvantage since programs needed to be converted into machine language (by an interpreter or a compiler) before they could be executed.

High-Level Languages

Some of the high-level languages have been discussed as follows:

1. Fortran

FORTAN (FORmula TRANslation) was the first high-level language developed by John Backus at IBM in 1956.

FORTRAN has a number of versions with FORTRAN IV being one of the earlier popular versions. In 1977, the American National Standards Institute (ANSI) published standards for FORTRAN with a view to standardizing the form of the language used by manufacturers. This standardized version is called FORTRAN 77.

2. Cobol

COBOL (Common Business Oriented Language), the first language used for commercial applications, was developed under the leadership of Grace Hopper, a US Navy programmer, with a group of computer manufacturers and users in 1959. The maintenance and further growth of the language was handed over to a group called CODASYL (CONference on Data Systems Languages).

It is written using statements that resemble simple English and can be understood easily; for example, to add two numbers (stored in variables A and B), a simple statement in COBOL would be: ADD A TO B GIVING C.

COBOL was standardized by ANSI in 1968 and in 1974. COBOL became the most widely used programming language for business and data processing applications.

NOTES

3. Basic

BASIC (Beginner's All-purpose Symbolic Instruction Code) was developed as a teaching tool for undergraduate students in 1966 by John Kemeny and Thomas Kurtz, two professors at Dartmouth College. Eventually BASIC was used as the main language amongst the personal computer users.

A minimum version of BASIC was standardized by ANSI and is so simple that it has been incorporated in every subsequent version of BASIC. Some versions of BASIC include MBASIC (Microsoft BASIC), and CBASIC (Compiler based BASIC).

One of the newer versions of BASIC, commonly known as Visual Basic, has also evolved from the original BASIC language. It contains various statements and functions that can be used to create applications for a Windows or GUI environment.

4. Pascal

PASCAL was designed by Nicholas Wirth, a Swiss professor, in 1971. It was developed as a more structured language used for teaching which Wirth named after the French mathematician Blaise Pascal, who also designed the first successful mechanical calculator. His primary aim was to provide a language that supported beginners learning good problem solving and programming techniques.

In addition to manipulation of numbers, PASCAL supports manipulation of vectors, matrices, strings of characters, records, files and lists, thereby supporting non-numeric programming. Hence, it has proved to be an attractive language for professional computer scientists.

PASCAL has been standardized by ISO (International Standards Organization) and ANSI.

5. PL/1

(Programming Language 1) was developed by IBM in the 1960s and was the first language that was attempted to be used for a variety of applications rather than one particular area like business or science or artificial Intelligence.

6. LISP

(LISP Processing) was developed in the early 1950s but was implemented in the 1959 by John McCarthy at the Massachusetts Institute of Technology. It became a standard language with the artificial intelligence community and was a program that could easily handle recursive.

7. C

This language was developed by Dennis Ritchie of Bell Laboratories in order to implement the operating system UNIX.

8. C ++

This language was developed by the Bjarne Stroustrup of Bell Laboratories by enhancing C. C++ is also used to write procedural programs like C but the reason for its increased popularity is perhaps because of its capability to handle the rigours

of object-oriented programming. C and C++ are the most extensively used general-purpose languages amongst programming experts.

9. JAVA

Java is again an object-oriented language like the C++ but is a simplified version with extra features. It is less prone to programming errors. It was developed for writing programs that could be safely and easily executed through the Internet. It is free from any kind of common virus threats. It is basically a network-oriented language that can develop website pages with enhanced multimedia features using small java programs known as java applets, Java is a secure to use over the Internet and is a platform independent language.

Generation of Computer Languages

Computer languages have changed with every changing generation of computers. Some of these new languages are:

1. Fourth-Generation Languages (4 GL)

These are non-procedural languages, which suggests that they present the objective but not the procedure to achieve it. The main features of fourth-generation languages are as follows:

- They constitute simple instructions.
- They are user-friendly, enabling the user to write programs and achieve the desired goal.
- They eliminate the need of a professional programmer for writing programs.

FORTH, was the first fourth-generation language and was developed by Charles Morre, an American astronomer, in 1970.

Its main application is in the industrial and scientific control applications. Another example of fourth-generation language is FOCUS.

2. Fifth-Generation Languages (5 GL)

Fifth-generation languages have developed as a result of research in the area of artificial intelligence. They are, however, early stage.

- (i) **PROLOG (PROgramming LOGic):** A logic programming language. PROLOG a general purpose language often associated with artificial intelligence and computational linguistics. It has a purely logical subset, called pure Prolog, as well as a number of extralogical features.

Prolog was developed in the early 1970s by two French computer scientists, Alain Colmerauer and Philippe Roussel.

Some of other popular languages include:

- (ii) **GPSS (General Purpose System Simulator):** Used for modelling physical and environmental events
- (iii) **SNOBOL (String-Oriented Symbolic Language):** Designed for pattern matching and list processing.
- (iv) **LOGO (a version of LISP):** Developed in the 1960s to help children learn about computers.

NOTES

- (v) **PILOT** (Programmed Instruction Learning, Or Testing): Used in writing instructional software.

2.3.4 Object Oriented Language

NOTES

The new discipline, object-orientation emerged during the 1960s. By the end of 1960s the first language of this type-originated. Object-orientation is a form of organization and this can be performed to some degree with any programming language. Simula was the first language of this category but it never became widely used. It was the archetype of the concepts of object-oriented programming and has been highly influential on its successors. The designers of Simula were Ole-Johan Dahl and Kristen Nygaard, who worked at the Norwegian Computing Centre in Oslo. Their effort and work became successful when they provided a Simula I prototype compiler running on a UNIVAC 1107 mainframe in 1964.

Smalltalk was one of the languages that embraced the concepts of class and message of Simula. This project commenced in the early 70s at Xerox PARC. The main aim was to create the quintessential dynamic object-oriented language. This allows new classes, objects and behaviour to be appended by representing the class hierarchy, objects and methods of a program as meta-data at run-time.

One of the goals of object-oriented languages is the creation of flexible programs. A minor shift in requirement of a program should not demand an entire rewrite. Also, if other programs share any of its features, it should be able to adapt the relevant sections to save time and money. The success of any object-orientation language depends on the feature that when evolving a new language from an existing one, the old source code need not be rewritten and so, the new techniques could be progressively mastered. On this ground, the most successful language is C++. C was invented by Dennis Ritchie of Bell Laboratories in 1972. It became the most prevalent language of its time. C++, which was credited to Bjarne Stroustrup of AT&T Bell Laboratories Computing Science Research Centre in 1983, had few incompatibilities with C, permitting gradual familiarization with the procedural improvements and features to support data abstraction and the object model.

Microsoft also made an Object-oriented language and gave their own version of BASIC called Visual BASIC. Similarly Borland, had their own successful Turbo Pascal environment and later introduced this methodology to their product and released their own visual tool, Delphi. The architects of Smalltalk worked hard to overcome its shortcomings and their research centered on supplementing Lisp with the Common Lisp Object System. Similarly, commercial object-oriented version of COBOL came into picture as Micro Focus Object COBOL.

The discussion of the object-oriented programming language is not complete without the reference of Java which was delivered by Sun Microsystems in 1995. The main objective of Java was to let devices, peripherals and appliances possess a common programming interface (platform independence). A Java compiler generates a pseudo code which must be executed on any computer having an appropriate interpreter known as a Java virtual machine (JVM). With a similar syntax and many improvements to C++, like garbage collection, its deployment has become extremely fast. The survey says that over 40 per cent of the global

businesses asked are already using Java and over 30 per cent intend to use it soon.

All these languages possess certain distinguishing characteristics: classes, inheritance, polymorphism and ease of modification.

Table 2.5 lists of some of the present day object-oriented programming languages:

Table 2.5 Important Object-Oriented Programming Languages

Sl No	Language	Sl No	Language
1	A#	15	BLAZE
2	ABCL	16	Common Objects
3	Abel	17	Blue
4	Actor	18	Bob
5	Acttalk	19	Concurrent
6	Ada	20	ConstraintLisp
7	ADES	21	cooC
8	ADVSYS	22	COOL
9	Agora	23	CSSA
10	ALLOY	24	Denali
11	ALTRAN	25	DinnerBell
12	AppleScript	26	DOWL
13	ASDL	27	DRAGOON
14	Beta	28	DROOL
29	CommonLoops	42	Object Lisp
30	HERAKLIT	43	Pascal Plus
31	ISLisp	44	POP++
32	Java	45	SmallWorld
33	MEROON	46	Turbo Pascal
34	UC++	47	VDM++
35	Venari	48	POOL2
36	XDL	49	PopTalk
37	XLISP	50	Omega
38	XScheme	51	O'small
39	yacc	52	Oz
40	Z++	53	PHOCUS
41	Neon	54	Object-Oriented

NOTES

Check Your Progress

1. Define the term binary number system.
2. State about the decimal to octal conversion.
3. What is computer software?
4. What are the two categories of factors that affect software quality?
5. Define the term computer language.
6. What language can a computer understand?

2.4 HISTORY AND EVOLUTION OF OPERATING SYSTEMS

NOTES

Operating systems have been evolving from the early 1950s. The chronological order of their evolution is given as follow.

1950s: Serial processing operating systems were used during the 1950s. The utilization of CPU was very low.

1960s: Batch operating systems evolved that could execute a batch of jobs. The utilization of CPU improved. For example, mainframe systems.

1970s: Multiprogrammed operating systems were developed, where CPU usage was multiplexed among various jobs present in the memory. When a job was waiting for I/O, the CPU could execute other jobs. CPU was not allowed to sit idle. CPU resource was utilized to a maximum.

Time sharing systems also evolved during this decade. It was an extension of the multiprogrammed systems. In this the processor time was shared among the jobs available in the memory. This facilitated interactive computing. In time sharing system the switching of jobs was done very fast and each job was given a CPU time for a fraction of a second. The response time in these systems was very short. Time sharing systems were also called multitasking systems.

1980s: Desktop systems or personal computers were introduced in the 1980s. These systems concentrated more on users' convenience and responsiveness rather than on improving the utilization of the CPU.

Multiprocessor systems were introduced to facilitate parallel computing. These systems increased the throughput (throughput is the number of jobs executed per unit time) and reliability.

Distributed systems were also developed parallelly. These operating systems facilitated network communication like LAN, MAN and WAN.

1990s: In the 1990s, real time systems were introduced. This operating system used time constraint technology, i.e., operational deadlines from procedure to system response.

The ultimate goal of an easy to use and human-friendly operating system is to transform the computer into a useful, user-friendly, acceptable and affordable machine for all the uses. To achieve this ultimate goal, the following three major objectives must be realized:

1. Convenience or ease of use

Operating systems hide the idiosyncrasies of hardware by providing abstractions for ease of use. Abstraction hides the low-level details of the hardware and provides high-level user-friendly functions to use a hardware. For example, consider using a hard disk which records information in terms of magnetic polarities. A hard disk consists of many cylinders and tracks and sectors within a track, and the read-

write heads for writing and reading bits to a sector. A user of a computer system cannot convert the data to be recorded in a format needed for the disk and issue low-level commands to address the appropriate track/ read–write head and write the data into the right sector. There are software programs called device drivers for every device types to handle this task. The programmer just issues read/ write commands through the OS and the OS passes it to the driver which translates this command to low-level commands that a hard disk can understand. The OS may issue the same command for reading/ writing to a tape drive or to a printer. However, the device drivers translate it into the proper low-level commands understood by the addressed device. The same is true for any other electromechanical and electronic devices connected to a computer.

An application program uses the built in functionalities of an operating system. The OS provides functionalities to other programs through the use of software interrupts or system calls. Library functions are provided for all standard input/ output operations. A programmer can reuse them and concentrate on the coding of major logical aspects of the problem at hand.

2. Efficient allocation and utilization of resources

A computer system has various resources like the central processing unit (CPU), memory and I/O devices. Every part of the resource is controlled by the OS. Executing programs or processes may request for the use of resources and the OS sanctions the request and allocates the requested resources, if available. The allocated resources will be taken back from the processes normally after the use of the same is completed. That is, the OS is considered as the manager of the resources of a computer system. The allocation of resources must be done after analysing the overall efficiency and utilization. That is, the allocation must be done to improve the system through-put (number of jobs executed per unit time) and the use of resources in a fair manner.

3. Ability to evolve

As the time passes, the users need new and improved services for accomplishing some task with the computer. The OS provides many services to the user community. In some situations, as you use the system for doing new task, some of the existing services may introduce undesirable side effects causing inefficient utilization of resources or may even compromise with the security of the system itself. So, it is desirable to remove the existing implementations of such services and introduce new implementations. Also, there can be bugs in the code that may emerge over time as you use the services for various purposes. The design of the OS must include provisions for easy introduction of new services and removing, improving or replacing existing services. Also, the design must provide easy interfacing facility to connect and communicate with new types of hardware devices and upgraded versions of the existing hardware devices.

Hardware Components of Computer System

You understand from the preceding discussion that operating systems control the hardware components of computers. So, let us study the major hardware

NOTES

NOTES

components of computers. Figure 2.6 depicts the four major hardware components of a computer system. They are:

- Processor or central processing unit (CPU)
- Input/output (I/O) devices
- Memory
- Buses.

Let us see each these components in detail.

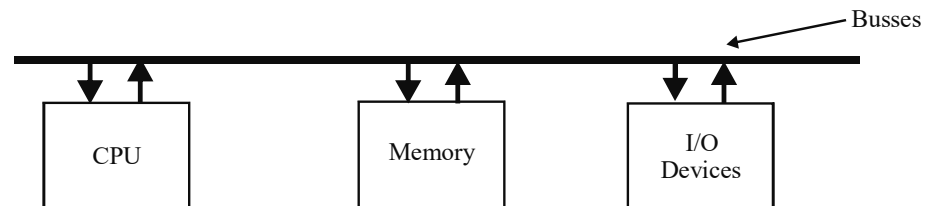


Fig. 2.6 Hardware Components of a Computer

- 1. Central Processing Unit (CPU):** The CPU consists of various registers for holding instructions and the operands of instructions and other functions, arithmetical/ logical unit for carrying out arithmetic and logic operations dictated by the instructions and the control/ execution unit for generating the control signals needed for fetching instructions, operands and executing instructions. A modern CPU also includes the primary cache memory for speedy access of instruction and data.

As registers will be referred to at many points in the discussion of operating systems, some more details of same are necessary. There are various other functions or purposes for which registers are used. Registers are used for holding operands or results of computation are called accumulators. There is a special register to act as instruction pointer or program counter (PC). The status or state of CPU at any point during computation, that is recorded and normally kept in a register is called a status register. The top of a stack used for holding temporary and local variables is pointed to by a register called the stack pointer.

- 2. Input/ Output Devices:** The input/output devices of a computer include keyboard, monitor, mouse, pen, printers, scanners, secondary storage devices like hard disk, floppy disk, CD ROMs, etc. These are used for the interaction of the computer with real-world objects. A keyboard, mouse and monitor are needed for the human interaction with the computers. The program and data are stored in the permanent secondary storage devices before execution. A huge amount of storage space at low cost is feasible with secondary or permanent storage devices. However, the secondary storage devices are very slow as compared to the random access memory (RAM) the primary or main memory storage devices.
- 3. Memory:** The main memory or RAM is the place where normally the program and data are stored before the execution begins. Later, when you study virtual memory technique, we will see that only some part of the program or data needs to be loaded in the main RAM before the start of

execution. Other parts will be brought as and when required during the execution time. The speed of access (reading and writing) of RAM is normally about one tenth of CPU Register/ cache memory access speed. So, a high-speed cache memory is normally placed between the CPU and the main memory (RAM).

- 4. Buses:** Bus is the electrical interconnection wires that are used for connecting the CPU, memory and I/O devices together so that the data can be sent and received by each of the components. Each of the devices/ components are connected to common buses. The CPU communicates with memory and I/O devices through the buses. There are various buses, between a pair of components such as address bus, data bus and control bus.

Operating System Services

In a computer system, users have to run their programs, store programs or data permanently on secondary storage devices. They also have to determine the malfunctions occurring programs and locate the information needed to identify the reasons for errors. Processes (running programs) may have to communicate with each other for sharing of data and to cooperate in execution. Processes may also need to read data from input devices like keyboard and write data to output devices like printer and monitor. It will not be practicable for each user program to build these facilities for accomplishing their task. Moreover, such facilities if built into each user program, may lead to a system that is unsecured and undependable. So, the operating system must provide the services needed to accomplish all these tasks. The services provided by an operating system for these and other purposes may be grouped under the following four headings:

1. Program execution, control and communication
2. Service for I/O operations
3. File manipulation
4. Error detection

1. Program execution, control and communication services

A computer system must provide facilities for the execution and control of programs (process). A program needs memory for loading the program before the start of execution and other resources including the CPU for execution. In a system, there will be many users and each user will be executing many tasks all of which need memory, the CPU and other resources. These resources will be normally limited and shared among processes in a space and/ or time in a multiplexed fashion. In some situations, an application may be executed as a multiple cooperating process possibly on different computers, connected through network. It may be necessary to exchange information between processes. Also, processes will be competing for using the available resources. So, these functions should be done in a supervisory mode by the operating system. Thus, an OS must provide the following services:

- (i) Creation and deletion of processes
- (ii) Suspension and resumption of processes

NOTES

NOTES

- (iii) Synchronizing the execution of processes
- (iv) Communication between processes on single or different computers
- (v) Prevention and handling of deadlock situations
- (vi) Keeping track of the amount of memory spaces allocated to different processes and the addresses of such memory spaces
- (vii) Allocation of main memory when requested
- (viii) De-allocating the main memory used by the process when it terminates and adds it to the free space list for future allocation to other processes.

2. I/O operations

User programs need to perform input and output (I/O) operations on different types of devices during their execution. Each device has its own controller and associated commands for I/O. It will be highly inefficient in case a user needs to learn all these low-level commands and incorporate these in the program. The operating system must hide the hardware details of the devices from the users and provide high-level commands like Open, Read and Write to handle I/O on all devices irrespective of their types. Also, from the protection of files and efficiency point of view, it is not advisable to do the I/O operations directly by a user program. Manufacturer of each device type will provide a program that is appropriate to the device driver. The device drivers can take high-level commands like open, read and write as inputs and translate them into low-level commands that the controller of the device can understand. So, an OS must provide the services like Open, Read, and Write to handle all devices in a uniform manner.

3. File manipulation

Files are necessary for storing data and programs permanently. There are several different file types like text, image, executable, music, video and program source files. A program takes input from some files and writes output to some other files. Generally speaking, organized information is stored in files. Files are created on secondary storage devices. Filesystem is that part of the OS program that is used to create, manipulate, organize and maintain files. Files are organized using directories in a Filesystem. Users need not bother about how the file is stored and retrieved from secondary storage. It is the duty of the operating system to hide the hardware specific details of secondary storages. The OS provides the following services to manipulate and maintain the files in the system:

- (i) Create files and directories
- (ii) Open files
- (iii) Close files
- (iv) Read files
- (v) Write files
- (vi) Rename files
- (vii) Copy files
- (viii) Delete files and directories
- (ix) List files

- (x) Search files and directories.
- (xi) Allocate, de-allocate and manage secondary storage space

4. Error detection

A running program may have errors or bugs, which need to be detected and corrected. Errors may be due to many reasons including trying to execute illegal instruction or accessing of memory (intentionally or unknowingly) that is not allocated for a user or infinite loop in a process or trying to execute an illegal operation on I/O devices. The operating system monitors the functioning of different programs to detect errors and localize them. Users can confidently execute their programs without the worry of malfunctioning.

NOTES

Components of Operating Systems

An operating system is a complex and normally huge software used to control and coordinate the hardware resources like a CPU, memory and I/O devices to enable easy interaction of the computer with human and other applications. The objects or entities that an operating system manages or deals with include processes, memory space, files, I/O devices and networks. Describe each of these entities have been discussed in the following subsections.

Process

A process is simply the program in execution. For every program to execute, the operating system creates a process. A process needs resources like a CPU, memory and I/O devices to execute a program. These resources are under the control of the operating system. In a computer, there are many programs in the state of execution. Hence, a large number of processes demanding various resources also need to be maintained and managed in an operating system. When the execution is finished, the resources held by that process will be returned back to the operating system.

Memory Space

As mentioned earlier, the execution of a program needs memory. The available memory is divided among various programs that need to be executed simultaneously (concurrently) in a time-multiplexed way. In normal type of memory, only one memory location can be accessed at a time. So, in a uni-processor environment, because the secondary storage devices are much slower than the main memory for fast overall execution speed the programs to be executed concurrently are loaded into memory and kept ready, awaiting the CPU. The memory is space multiplexed to load and execute more than one program in a time interleaved way. Even if more than one CPU is available, only one program memory area can be accessed at a time. However, instructions that do not need main memory access (access can be from local memory or from CPU cache) can be executed simultaneously in a multiprocessor scenario.

Files

Files are used to store sequences of bits, bytes, words or records. Files are an abstract concept that are used to represent storage of logically similar entities or

NOTES

objects. A file may be a data file or a program file. A file with no format for interpreting its content is called a plain unformatted file. Formatted file content can be interpreted and is more portable as one knows in what way the data inside is structured and what it represents. Example of file formats are JPEG, MPEG, GIF, EXE, MP3, etc.

Input/ Output(I/O) Devices

The input/output devices of a computer include keyboard, monitor, mouse, pen, joystick, scanner, printer, modem and secondary storage devices like hard disk, floppy disk, CD ROMs, etc. Primary memory like RAM is volatile and the data or program stored there will be lost when you switch off the power to the computer system or when you shutdown the computer. Secondary storage devices are needed to permanently preserve program and data. Also, as the amount of primary storage that can be provided will be limited due to reasons of cost, a secondary storage is needed to store the huge amount of data that is needed for various applications.

Network

Network is the interconnection system between one computer and the other computers located in the same desk/room, same or adjacent building, in any geographical location in the world. It is possible to have wired or wireless network connections to other computers located anywhere in the world.

2.4.1 Functions of OS

From the preceding discussion, it is clear that the operating system has various functions to carry out. So, the operating system consists of many subsystems to accomplish its various tasks. The major components or subsystems that are needed for an operating system are:

- Process management system
- Process scheduling system
- Memory management system
- File and secondary storage management system
- I/O management system
- Networking system
- Protection system
- User interface system

The following subsections describe each of these subsystems.

• Process management system

An operating system must provide many services for the execution of user programs and system programs. Programs are executed as processes. A process is a program in execution. A program can be executed by many processes on behalf of different users. For example, a C++ compiler will be executed by many users as different processes. A running program has various data structures to represent its context. Code, data, program counter (PC), registers, open files, and other resources in use all together represent the context. A process encapsulates

the whole context of a running program. Once the execution of a process is completed, it should be removed, killed or deleted from the system. A process requesting for the resource that is not currently available must be made to wait or suspended. It must be resumed when the resources are available. More than one process sharing resources must be synchronized for the right and consistent use of such shared resources. A process should not be allowed to wait for a resource that is already held by another waiting process in a circular chain. In some cases, a job has to be executed as a multiple cooperating process. So, for the execution, control and deadlock-free cooperation of processes, the process management subsystem must provide services for the following:

- o Creation and deletion of processes
- o Suspension and resumption of processes
- o Synchronizing execution of processes
- o Communication between processes
- o Prevention and handling of deadlock situations

Modern operating systems also support the concept of threads which are lightweight processes. A process may have many threads to improve the response time of an application. In such a case, services similar for process management can be used for thread management.

• **Process scheduling system**

In a computer system, there will be many processes in a ready state, waiting for the CPU for execution. When and in what order to allocate the CPU for different processes, in the ready state, is called process scheduling or CPU scheduling. This scheduling is done based on the various scheduling algorithms that are used for this purpose. In a multiuser computer system, the next process to be executed is selected based on the scheduling algorithm. Also, the process will be allocated the CPU for a duration that is dictated by the scheduling algorithms. After the expiry of permitted time, a hardware timer interrupts the current execution to enable the operating system to select the next process to execute, based on the scheduling algorithm. The context of the outgoing process must be saved and the context of the new incoming process must be restored before handing over the CPU to the newly selected process. This is called a context switch. There are various scheduling algorithms that are in use for optimizing the performance of computer systems. Common among them are, round robin, first-in-first-out (FIFO), shortest time first, shortest remaining time next, etc. So, the process scheduling subsystem must provide the following functions for the execution of programs:

- o Services for selecting next process to execute based on some scheduling algorithm
- o Context switching
- o Restarting the timer to interrupt at the end of the time duration allotted for the new process
- o Transfer control to the new process

NOTES

NOTES

• Memory management system

The main memory is arranged as an array of bytes or words as dictated by the low-level architecture of a CPU. Every memory location has a different address. These addresses can be from zero to the size of memory that can be accommodated in the CPU. This is called the address space of a CPU. These are the memory addresses that the CPU can access directly. As mentioned earlier, in a multiuser system, there will be many programs loaded and existing simultaneously in the main memory. The available main memory will be allocated to such programs/processes. Processes may also request memory during the execution time. So, the operating system must allocate memory to suffice the needs of different processes out of the free memory available at any time. When the processes complete their execution, its memory must be taken back to the free memory list. When sufficient main memory is not available, the memory occupied by some of the waiting processes or a part of the executing processes may be written to secondary storage. The memory thus freed, may be allocated to the demanding processes. So, the memory management subsystem has to provide the following facilities for the execution of programs in order to:

- o Keep track of the amount of memory spaces allocated to different processes and the addresses of such memory spaces
- o Allocate main memory when requested
- o De-allocate memory used by the process when it terminates and add it to the free space list for future allocation to other processes

• File and secondary storage management system

File is an abstract concept for recoding the memory locations where the information is recorded on storage media. Files can also be considered as information folders. Information is recorded in secondary storage as sequences of bits, bytes, words or records. Files are stored on secondary storage media. Hard disk (magnetic disks), magnetic tapes and optical disks such as compact disk read-only memory. CD ROM, digital versatile disk—DVD, etc. are examples of secondary storage devices. Different storage devices are characterized by speed, capacity, data transfer rate and access methods. Files are organized into directories based on the information content or the purpose for which it is used or based on the owner name. The various functions that a file management subsystem must provide are:

- o Creating and deleting files
- o Creating and deleting directories
- o Opening, reading, writing and appending files and directories
- o Allocating storage block to files and de-allocating storage blocks to free list
- o Keeping track of the usage of storage by files
- o Backing up files

• I/O management system

Users of a computer system interact mainly through I/O devices. Through the keyboard and mouse, you can apply inputs and commands to the computer. Through

NOTES

the monitor you can get back the responses from the computer. I/O devices normally communicate with the computer through hardware-level interrupts. The operating system has to respond to such events by executing interrupt service routines stored as a part of its code. The I/O subsystem provides utilities for all types of communication needed with the I/O devices. Moreover, every device type has its own special program called the device driver for interaction and communication with the operating system.

• **Networking system**

Modern operating systems that support the execution of programs in a remote computer connected through a communication network are called network operating systems. Network operating systems are loosely coupled software on loosely coupled hardware (independent computers connected to the network). Such a system must provide remote login facility for executing programs and remote copy command facility.

It is also possible to have a system with tightly coupled software on loosely coupled hardware. Such a system is called distributed operating system. A distributed operating system creates the impression of a large uni-processor computer out of many networked computers. A single operating system runs in all the computers and are tightly coupled through software. Files in such a system appear to have the same name (with different access privileges) to different users working in any of the networked computers. A distributed system must provide a single global inter-process communication mechanism so that any process can communicate with any other process. Distributed operating system must also provide schemes for global protection of files.

• **Protection system**

In a multiuser system, there will be many users who will be executing programs simultaneously/ concurrently. Also, different users will have their own resources like programs and data files. Some users may permit their resources to be shared by other users in a controlled way. An operating system must provide facilities for protecting the shared use of resources like programs and files and also provide facilities for controlled access rights to such resources by the users. The protection system must ensure that each resource of the system is accessed correctly and only by those processes that are allowed to do so.

• **User interface system**

Users interact with a computer using a command interpreter or a shell program. Human users can also communicate through a human computer interface. They can enter commands through a terminal such as keyboard, mouse and monitor, by typing on the keyboard or clicking on an icon or button on the desktop. The shell interprets commands and executes the operations by invoking operating system services. A user program can directly invoke the operating system services through software, interrupts also known as system calls.

2.4.2 Operating System as Resource Manager

The operating system can be looked upon from two angles, namely, Extended Machines and the Resource Manager. When looking from the extended machines viewpoint,

NOTES

users are provided with a virtual machine which is quite easy for usage by the operating system. From the resource manager's viewpoint, various parts of the system are efficiently managed by the operating system. The designing of the structure of the OS can be done in the form of a monolithic system, a chain of various layers, a virtual machine system, an exokernel or through the use of the client-server model. Some of the fundamental concepts of OS can be stated as processes, memory management, I/O management, the file systems as well as the security.

A virtual machine (VM) can be defined as the software implementation of a system such as a computer by which the programs are implemented similar to a physical machine. Present usage involves virtual machines having indirect relation with any real hardware.

There are two major categories of virtual machines.

1. System virtual machine
2. Process virtual machine

This categorization is based on the extent to which they resemble any real machine and their use.

Complete system platform is provided by a system virtual machine that supports execution of a complete operating system (OS). On the contrary, a single process is supported by a process virtual machine. For both types of virtual machines, an essential characteristic is the limitation of resources and abstraction to which the software running within is subjected. In a nutshell, it cannot go beyond its virtual world.

For example, a program created in Java gets services from the JRE (Java Runtime Environment) which is an application to which commands are given and the expected results are received from Java software. Thus, such Java application provides services to program and acts as a 'virtual machine' that takes place of an operating system or hardware for running this program.

1. System virtual machines

The system virtual machines, also known as the hardware virtual machines permits the sharing of basic physical machine resources amid various virtual machines, where every machine runs its separate operating system. The virtual machine monitor or hypervisor is the software layer that provides virtualization. The execution of the hypervisor can be done on bare hardware, i.e., Type 1 or native VM or on top of an operating system, i.e., Type 2 or hosted VM.

Important merits of the virtual machine include:

- The existence of more than one OS environments on a single system, which is in complete seclusion from one another
- An instruction set architecture (ISA) is provided by the virtual machine that is quite dissimilar from that of the real machine
- A request provisioning, preservation, high accessibility and disaster revival system.

One of the main disadvantages of a virtual machine is that it is less efficient as compared to a real machine when the hardware is indirectly accessed by it.

NOTES

The guest operating systems wherein all multiple VMs run their separate operating system are commonly used for the purpose of server consolidation. In this, various services which are used for running on distinct machines for the purpose of avoiding any type of interruption are executed in separate VMs over the same physical machine. A use like this can be normally termed as Quality-of-Service isolation (QoS isolation).

The similarity among the guest OSES is not required, making it feasible to execute various OSES on the sole computer like Microsoft Windows and Linux, or earlier versions of an OS for assisting software which has not been transferred to the new version yet. In embedded systems, using virtual machines by which various guest OSES are assisted is turning to be trendy; a general use is in assisting a real-time operating system simultaneously as a high-level OS like Linux or Windows.

Sandboxing an unreliable OS is another use, as it is a system that is yet to be developed. There are other benefits to the virtual machines for the development of the OS, such as efficient debugging access and reboots that are swift in speed.

A level of isolation inside a single operating system is provided by a substitutive technique like the Solaris Zones. As compared to a VM, this does not have that much of isolation.

2. Process Virtual Machines

A process VM also known as an application virtual machine executes in the form of a usual application within an OS assisting only one process. It is generated at the initialization of the process and removed at the time of exiting from the process. The basic purpose provided by it is to present a platform-independent programming environment through which details of the underlying hardware or operating system can be abstracted and programs for executing in a similar way on any platform can be permitted.

In Java programming language, such a type of VM that is executed by the use of the Java Virtual Machine is very familiar. Various other instances include the Parrot Virtual Machine that plays the role of an abstraction layer for various interpreted languages, as well as the .NET Framework that executes on a VM termed as the Common Language Runtime (CLR).

Techniques for Virtualization

The techniques for virtualization have been discussed here.

Emulation of the Underlying Raw Hardware (Native Execution)

Such an approach may be defined as a complete virtualization of the hardware that may be executed through the use of a Type 1 or Type 2 hypervisor. A Type 1 hypervisor executes on the hardware in a direct manner whereas a Type 2 hypervisor executes on another OS like Linux. Any of the OS that is assisted by the basic hardware, may be executed by any virtual machine. Therefore, two or more 'guest' operating systems can be executed at the same time by the user, in different 'private' virtual computers.

NOTES

The pioneer system using this concept was IBM's CP-40, the first (1967) version of IBM's CP/CMS (1967–1972) and the precursor to IBM's VM family (1972–present). With the VM architecture, most users run a relatively simple interactive computing single-user operating system, CMS, as a 'guest' on top of the VM control program (VM-CP). This approach kept the CMS design simple, as if it were running alone; the control program provides multitasking and resource management services in a 'behind the scenes' manner. In addition to CMS, VM users can run any of the other IBM operating systems, such as MVS or z/OS. z/VM is the current version of VM and is used to support hundreds or thousands of virtual machines on a given mainframe. Some installations use Linux for zSeries to run Web servers, where Linux runs as the operating system within several virtual machines.

Note: Intel and AMD have introduced features in their x86 processors to enable virtualization in hardware.

Emulation of a Non-native System

Virtual machines can also perform the role of an emulator, allowing software applications and operating systems written for another computer processor architecture to be run.

Some virtual machines emulate hardware that only exists as a detailed specification. For example, consider the following:

- One of the first was the p-code machine specification, which allowed programmers to write Pascal programs that would run on any computer running virtual machine software that correctly implemented the specification.
- The specification of the Java virtual machine.
- The Common Language Infrastructure virtual machine, which is at the heart of the Microsoft .NET initiative.

This technique allows diverse computers to run any software written to that specification; only the virtual machine software must be written separately for each type of computer on which it runs.

Operating System-level Virtualization

Operating System-level Virtualization is a server virtualization technology which virtualizes servers on an operating system (kernel) layer. It can be thought of as partitioning—a single physical server is sliced into multiple small partitions (otherwise called virtual environments (VE), virtual private servers (VPS), guests, zones, etc.)—where each such partition looks and feels like a real server, from the point of view of its users.

Extended Descriptions of Selected Virtualization Software

The virtualization of the hardware for multiple OS sharing can be done through the following software products:

- Adaptive Domain Environment Operating Systems (Adeos) is a Hardware Abstraction Layer which may be loaded in the form of a kernel module in

Linux. The loading of a real-time kernel in the form of a module is permitted in this simultaneously, as Linux with a higher priority.

- Virtualization of the x86 for unmodified PC operating systems execution is provided by the parallels. The use of a lightweight hypervisor technology is done for enhancing the security and increasing the efficiency. The wide usage of parallels is simply due to its advantage of running Windows as a guest in the Mac OS X on the Apple-Intel architecture.
- QEMU (Quick EMU lator) is an emulator that is based on a virtual machine that enables the emulation of different guest CPU architectures on various different host platforms.
- Virtualization of the Windows based system is permitted by the Returnil Virtual System and for reconstruction of the virtual machine, just a restart is required.
- VirtualBox developed by Sun Microsystems, is an open source (GPL)/proprietary virtual machine. Virtualization of x86 is permitted by it and it also assists different host operating systems such as Windows, Linux, BSD and Solaris. VMware Workstation Virtual Machine Disk Format are also supported by it.
- Virtual machines for x86 which are capable of executing unmodified operating systems like Windows, Red Hat and SUSE are provided by Virtual Iron. It is an open source virtualization technology that executes native virtualization that sends near-native performance for x86 OS.

NOTES

Operating System as a Resource Manager

In every computer system, the basic vital resources are in the form of processors, storage, input / output devices, communication devices and data. A few functions of the OS include executing the user interface, hardware sharing amid the users wherein data sharing among the users is permitted. This avoids the users to interfere with each other, resources can be scheduled among the users, input/output can be facilitated, error recovery can be done, resource usage can be accounted, parallel operations may be facilitated, data can be organized for the purpose of secure and speedy access and network communications can also be handled.

Multiplexing of resources is included in two ways in the resource management — ‘in time’ and ‘in space’.

During the time multiplexing of a resource, various programs or users make use of that resource turn by turn such as a printer.

During the space multiplexing of a resource, the sharing of the resource, is done among them rather than taking turns. This means that every single program or user can make use of a part of the resource such as sharing main memory, hard disk etc.

Space multiplexing is another kind of multiplexing. In this, every customer gets a share of the resource instead of the customers taking turns. For instance, main memory is usually divided into various running programs due to which all can be resident simultaneously. For example, for the purpose of taking turns in using

NOTES

the CPU. If it is assumed that there is adequate memory present for holding more than one program, it is more proficient for holding various programs in the memory at the same time instead of giving everything to one of them, particularly if just a small fraction of the total is required. The hard disk is another resource which is space multiplexed. In various systems, files of various users can be held at the same time through a single disk. The allocation of disk space as well as keeping track of the users and the disk blocks is a basic task of the operating system resource management.

The generation of a link amid the material resources, the user as well as the applications like word processor, video game, etc., is the responsibility of the OS. When there is a requirement of a program to be accessed by a material resource, sending particular information to the peripheral device is not necessary and only the information conveying it to the relevant peripheral through its driver to the operating system is needed. In absence of any drivers, every program has to identify and consider the communication with all types of peripherals.

Therefore, the OS permits the 'dissociation' of programs and hardware, for the purpose of easing resource management and the user is provided with a simplified Man-Machine Interface (MMI) for overcoming the difficulty of the actual machine.

2.4.3 Types of Operating System

Operating systems can be classified on the basis of the purposes for which they are developed and the functions they support into nine classes listed as follows:

1. Early system
2. Batch processing system
3. Multiprogramming system
4. Time-shared system
5. Single user system
6. Multitasking system
7. Parallel system
8. Distributed system
9. Real-time system

1. Early system

Operating systems have evolved through a number of phases during the last half a century. The earliest electronic computer was developed in the mid-1940s and had no operating system at all. Machines (computers) were equipped with a console consisting of toggle switches and display lights. Eight or sixteen mechanical toggle switches arranged on the console were employed to enter 8 or 16 bits of instruction and data to the computer and other switches for reset, start and debug operations. There were no programming languages or translators for developing high-level language programs and converting them to machine understandable instructions. Programs were written in machine language instructions which the hardware or the CPU could decode and execute to do the work. Mainly the programmers were the users of the then computer systems.

NOTES

The scene improved in the 1950s. An input device known as punch card reader was developed to read programs recoded on a deck of punch cards. A card punching machine was used to record program instructions and data on punch cards. Magnetic tape memory was used as an intermediate storage device for programs and data read from punch card readers. Printers were also developed to record hardcopy output of programs. Assemblers and high-level programming languages were developed. Programs for linking object program and library functions and loading the linked machine code to memory were also developed. All these programs, viz., compilers, linker and loaders were either on punched cards or magnetic tapes. Users wishing to use the system had to sign-up manually to reserve the system use for a specified period say, one hour. A user wanting to execute his program had to set up card decks or mount magnetic tapes for the program, compilers, linker and loader. The output of execution was printed on a line printer. The system was operated in a serial fashion, one user at a time for the reserved time and after that, another user. However, this set up was very costly. There was a lot of time spent for setting up the needed devices like card readers/ magnetic tapes for source programs of user, compilers, library function, linkers and loaders manually for executing/ debugging a user program.

2. Batch processing system

To drastically reduce the setup time for executing a job and improving the system utilization, a batch processing program was developed in the mid-1950s. This helped to make the setup that was required for executing a job automatically and executing the programs (submitted for execution at the computer centre in the form of card decks) of different users one after another without user intervention. This batch processing system was called the first rudimentary operating system then known as resident monitor. A set of commands called job control language (JCL) for reading the programs using card readers/magnetic tapes and loading, executing and printing the output were developed. The commands of JCL were used to express the sequence of actions that the resident monitor must do to select and execute a job from the batch of jobs submitted by the different users.

With the introduction of resident monitors, the users were not permitted to directly execute their job. The system worked in two modes:

- (i) Privileged mode: Where all types of instructions could be executed.
- (ii) Non-privileged user mode: Where privileged instructions could not be executed.

Input/output (I/O) and some other instructions which are useful to control the execution of programs and to access sensitive resources are privileged ones. The total memory is divided into two regions, one for the operating system (monitor) and the other for loading user programs. The monitor is loaded in the OS memory and executes in privileged mode and always resides in memory; hence it is called the resident monitor. The monitor now controls the entire computer resources and their operations. As the I/O instructions are privileged, only the monitor can do the I/O. In case an user program prefers to do I/O, must request the monitor (OS) for the same, than it. In this case, the monitor takes back the control from the user to do the I/O in privileged mode for the user and gives control again to the user

NOTES

program. The users wishing to execute jobs on the computer submits them to the batch processing monitor. The monitor queues the submitted jobs and executed then one by one based on some strategy to execute. During the execution of user jobs, the system control is given to the user program, and when the user program completes its execution or comes across some errors, the monitor takes back the control. The monitor then takes the next job from the batch-queue and handover control to execute it, and so on. That is, the execution of the system worked in alternate periods of monitor (privileged) and user modes. If the user program executes any privileged instructions, an error condition occurs and the system control transfers to the privileged monitor mode.

Even though batch processing system reduced the setup times, the CPU utilization was still lesser. This was because there was a provision for loading only one program in the memory and ready for execution.

To improve the response time, a technique known as swapping was introduced to preempt the executing program and load a new high priority job for execution. The swapping technique writes back to the secondary storage, the memory image of the process that waits for slow I/O devices and loads the new process to replace the memory region of the saved process. Swapping not only improved the response time, but also the degree of multiprogramming at the expense of some additional I/O time.

Another major feature introduced to gain performance was the technique of SPOOLing (Simultaneous Peripheral Operation Online). In spooling, a high-speed device like a disk is placed between a running program and a low-speed device used for input/output such as a printer. This makes many virtual devices from a physical device and a program can do I/O to shared devices as if they are dedicated devices for individual program's use. Further, this arrangement drastically reduces the wait times of programs for slow I/O devices as programs can do I/O directly from and to high-speed disks.

High-speed secondary storage hard-disk drive was also developed for use in place of slower tape drives.

Batch processing system was widely used for information processing. Scientific computing and business data processing was the major applications of such computer systems. The need for organizing the data and programs of different users motivated the development of file technology. Files are abstractions of storage devices. Programmers and users maintained programs and their inputs and outputs on files with identifiable names.

Subsequently, for better protection and control, the computer systems were further equipped with features like mechanisms for protecting memory, and hardware timer which prevents a job from monopolizing the use of CPU time and interrupts.

3. Multiprogramming system

Jobs in a computer system consist of alternate cycles of I/O-bound and CPU-bound operations. In a computer system, I/O bound operations are done by special processors, called I/O processors. The CPU-bound operations are performed by

the main processor that is termed as the CPU of the system, which is in privileged monitor mode and on behalf of the user program (process) request, only needs to start the I/O operations, based on the I/O processors. The main CPU then continues executing the user programs in case there is some available process ready with all the required resources. When the I/O is completed, the interrupt mechanism informs the same to the monitor (OS). The interrupt mechanism switches the execution mode from the user to the monitor and begins executing from the interrupt handler routine that is part of the monitor (OS) code. The monitor then informs the completion of I/O to the concerned waiting user process and selects the next ready process (or the job or process whose I/O is just finished) for execution and transfers control to it. That is, the main CPU waits for the completion of I/O when there are no other jobs to be executed in the system. Therefore, in a uniprogramming system, at a time only one program is ready for execution, and if that program requests I/O during execution, then the CPU has to wait until that I/O is completed. Thus, as the I/O operation time is many times larger (100 or 1000 times) than the CPU instruction execution time, the CPU utilization is very low, normally less than the five per cent. In general, in a uniprogramming environment other resources like memory and I/O devices are underutilized.

Now you can improve resource utilization by making ready more than one program or jobs for execution. This requires the loading of more than one program in the memory. This needs partitioning the memory into more than two regions, one for OS and as many user memory partitions as there are user programs. By providing sufficient memory and by loading two or more jobs in the user memory regions there will be processes that are ready for execution. Therefore, you can allocate CPU to another process when one process is waiting for I/O completion. This is called multiprogramming or multitasking. That is, the CPU executes more than one program concurrently in a time-interleaved fashion. This also improves the memory and device utilization of the system as the idle time of memory regions and I/O devices are now less than as compared to the uniprogramming case.

With the introduction of multiprogramming, the OS needs to manage the memory and protect each user memory from other users in addition to protecting the operating system memory from users.

4. Time-shared system

A multiprogrammed batch system does not permit real-time interaction between users and computer as the user commands that are needed for executing jobs are prepared as scripts of job control language (JCL) and submitted to the batch system. As users are not permitted to submit the job script input and observe or take output directly, it took many days to debug and correct the mistakes in program development. The solution to this problem was the introduction of Interactive Time-shared Multiprogramming techniques. This enabled many users to interact with the computer system simultaneously, each one using a separate terminal keyboard and monitor connected to the system. Actually, each user is given a small time quantum (say, 100 milliseconds) to apply commands and receive responses from the computer system in a round robin fashion. So, there are 10 users, each will be served 100 milliseconds in every one second. Because of this fast switching of

NOTES

NOTES

execution among users, each one felt that the entire computer system time is available for his own use. This drastically improved the ease of use of computers and reduced the job processing time and program development time. With this interactive timesharing technology, a single computer system was made available for many people simultaneously for doing many different types of tasks.

The hardware of a computer system is normally very costly. In a multiuser system, as many users are sharing this costly hardware, the cost is shared among these users and accordingly, the resource utilization is also high. However, as the operating system has to switch between many users in a short time, there are some unproductive computations, called overheads computations that are done for job switching and associated work.

5. Single user system

A computer system in which only one user can work at a time is called a single user system. The familiar Intel processor based Windows OS Personal Computer (PC) is an example of single user system. Such a typical system has a single keyboard, mouse and monitor as I/O devices for the user to interact with the system. Users apply commands through the keyboard and mouse and the computer displays its responses (output) on the monitor. A user can do various tasks like preparation of a document or editing a program or writing a letter and printing it (using a printer). A user can simultaneously or concurrently execute many tasks or jobs in currently available personal computers. An important design issue of such interactive systems is the response time, which is the time taken by the computer to start producing output after a command has been entered. The response time must be within a predictable limit for user satisfaction and acceptability. When one user is working on the computer, the other users have to wait till the current user finishes his work and leaves. Examples of single user operating systems are MS-DOS, Windows 95, Windows NT, Windows 2000 and Windows Vista, each of which runs on Intel Processor based computers. Other examples are Macintosh OS X.x, Linux single user OS which are actually derivatives of UNIX operating system.

6. Multitasking system

Multitasking is concerned with a single user executing more than one program simultaneously. It may also mean a user executing his application as many concurrent processes. An operating system with multitasking ability runs many applications—for example, Microsoft Word, Microsoft Excel, printing documents, internet explorer, etc.—simultaneously/ concurrently. The operating system executes each for a small time slice in a round robin fashion so that the user cannot distinguish the switching of CPU among different applications. It appears to the user that they are running simultaneously.

The advantages of this system are as follows:

- (i) The ability of multitasking system that permits the user to run more than one task simultaneously, leads to increased productivity.
- (ii) Improves the system resource utilization, throughput and overall efficiency of the system

The disadvantages are as follows:

- (i) Increased overhead processing time
- (ii) Needs more resources like memory, CPU time and I/O devices for achieving user satisfaction

NOTES

7. Parallel system

A parallel system is the outcome of the thoughts concerned about the limits on the ultimate speed of computation achievable using the propagation speed of electronic signals. As you know, the speed of electronic signals is 10×10^{10} centimetre per second, and hence there is a limit on the physical size of the CPU chip when it is operated at speeds above 1 GHz (10^9 cycles per second). For example, a 10 GHz CPU chip cannot be more than 3-cm wide, as the signal travels only 3 cm distance in a clock period of 0.1 nanosecond. Moreover, the amount of heat generated becomes excessive when the size is small at such high clock speeds. Therefore, you have now reached the upper limit on the clock speed, and there is no way of increasing the processing speeds of application by increasing the clock speeds of the CPU further. The outcome of these ideas is the development of parallel processing systems with multiple CPUs/ computers There are two classes of parallel processing system which are as follows:

- (i) Tightly coupled bus-based shared memory multiprocessor system
- (ii) Loosely coupled distributed memory multiprocessor system

(i) Shared memory multiprocessor system

In this scheme, all processors share the same memory address space. Processors are interconnected with one another by using the common memory bus. If one processor writes a memory word, the change is visible to all other processors. At a time, only one processor can access the main memory. However, as each processor has its own private cache memory, the processor access will be mainly to this local cache. The processor needs to access the common shared memory only when there is a memory access outside the cache memory address space. This is called a cache miss. The cache will be designed in such a fashion that the number of misses will be normally less than 5 per cent. Thus, the presence of cache not only speeds up the memory access but also reduces the main memory access to about less than five per cent, reducing the chances of the bus contention out of access from different processors. Processes on different processors can easily communicate through shared memory. Figure 2.7 shows a tightly coupled bus-based shared memory multiprocessor architecture.

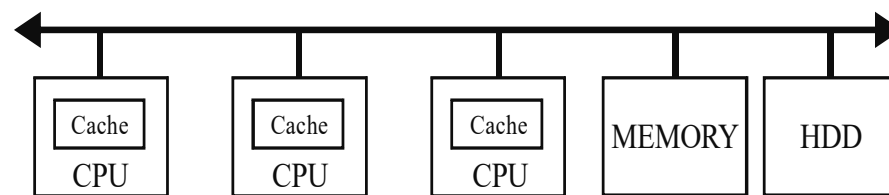


Fig. 2.7 Tightly Coupled Bus-Based Shared Memory Multiprocessor

NOTES

(ii) Distributed Memory Multiprocessor System

In this scheme, each processor has its own memory, cache and hard disk storage and they are interconnected through networks. Communication between the processes on different machines is carried out through message passing, which is normally slow due to network congestion when compared to shared memory communication. Figure 2.8 shows the structure of a distributed memory multiprocessor system.

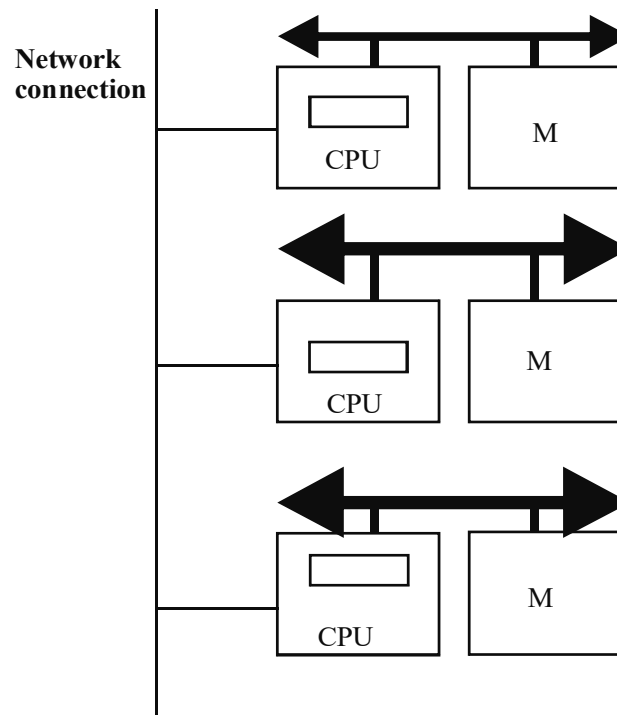


Fig. 2.8 Loosely Coupled Distributed Memory Multiprocessor

Executing your program on a distributed memory multiprocessor involves the organization of the program as a set of independent tasks that communicate with each other using messages. The partitioning of the job or program into independent tasks must be done in such a way as to avoid the communication requirements between different tasks running on different distributed memory computers. Therefore, the programmer must have the knowledge of the data processing need at various nodes of the distributed system.

Choosing to implement a program in shared memory versus distributed memory is usually based on the amount of information that must be shared by parallel tasks. Shared memory system might be better if the amount of shared information is large, as otherwise the distributed memory system communication cost will be much higher.

8. Distributed system

A distributed system is composed of a large number of computers connected by high-speed networks. The computers in the distributed system are independent, but they appear to the users as a large, single system. This is in contrast to the old

mainframe computers which are centralized, single processor systems with large memory, I/O devices and connected to many terminals (with no processing power), through which many people can work or run applications concurrently.

In distributed systems, users can work from any of the independent computers as terminals. The applications or programs run on any of the computers distributed even geographically distant places and possibly on the computer (intelligent terminal) from which the user entered a command to run the program. That is, people can work from any computers and the applications can run on any of the member computers in the system depending on the availability and workload. There is a mechanism built into the distributed OS to balance the load (CPU workload) on different independent computers of the system. A distributed system will have only a single file system, with all the files of a user accessible (as dictated by the permissions associated with each file) from any computers of the system with the same access permissions and using the same path name. Automatic teller machines (ATMs) of a bank are an example of a distributed system, which are useful for banking applications.

The advantages of a distributed system are as follows:

- (i) **Data sharing:** Allows data sharing access to common databases
- (ii) **Device sharing:** Allows many users to share expensive peripherals like high resolution printers, scanners and digital cameras
- (iii) **Communication:** Makes communication easier between users.
- (iv) **Even workload distribution** spreads the available workload over independent computers of the system in a cost-effective way.

The disadvantages of a distributed system are as follows:

- (i) **Delayed communication:** Network communication is delayed due to saturation or other problems
- (ii) **Security:** Needs tools and techniques for providing controlled access to sensitive data.

9. Real-time system

A real-time system (operating system) is the one which responds to real-time events or inputs within a specified time limit. The importance of real-time systems has increased especially in today's increased use of embedded applications. Real-time systems include satellite control systems, process control systems, control of robots, air-traffic control systems and autonomous land rover control systems. A real-time system must work with a guaranteed response time depending on the task; otherwise, the application might fail.

In a real-time operating system, some of the tasks are real-time and others are ordinary computing tasks. Ordinary computing tasks, like commands from interactive users, generally demand small response time as compared to the user interaction time. Faster response time keeps the users satisfied. Whereas, with a real-time task there always exists a deadline to start responding after an event and another deadline for the completion of the response. Real-time tasks are classified

NOTES

NOTES

as hard and soft. A hard real-time task must meet its deadline to avoid undesirable damage to the system and other catastrophic events. A soft real-time task also has a deadline, but failures to meet it will not cause a catastrophe or great losses.

Most real-time operating systems are designed with the objective of starting real-time tasks as rapidly as possible. This is possible only if the interrupt handling and task dispatching are fast. There should be some hard limit on the amount of execution time spent on an interrupt handler code, as the event may occur while a high priority interrupt handler is mid-way in its execution. The latency of response to events can be predictably controlled by carefully choosing the interrupt priority and limiting the amount of time that is needed by the CPU on interrupt handlers.

2.4.4 BIOS

BIOS is acronym for Basic Input/Output System, the built-in software that determines what a computer can do without accessing programs from a disk. On PCs, the BIOS contains all the code required to control the keyboard, display screen, disk drives, serial communications and a number of miscellaneous functions. The BIOS is typically placed in a ROM chip that comes with the computer and it is often called a ROM-BIOS. This ensures that the BIOS is always be available and is not damaged by disk failures. It also makes it possible for a computer to boot itself; this is because RAM is much faster than ROM. ROM-BIOS is normally located in an Electrically-Erasable Read-Only Memory (EEPROM) chip, which allows it to update through software control. This is commonly called flash BIOS. BIOS ROM is located in a socket on the motherboard and is relatively easy to locate because it is usually labeled with the name of the BIOS manufacturer. The BIOS set-up is accessed at start-up. When a PC is turned on, a short text message typically passes by very quickly on the screen indicating which key to press, usually it is the **[Del]** or **[F1]** function key. It also makes it possible for a computer to boot itself. These days, many computer manufacturers design their systems in way so that the BIOS is copied from ROM to RAM when each time the computer is booted. This is known as shadowing. Many modern PCs have flash BIOS, which means that the BIOS is recorded on a flash memory chip, which can be updated in future if necessary. The PC BIOS is fairly standardized, so all PCs are similar at this level. It varies on different BIOS versions. Additional DOS functions are usually added through software modules. This means users can upgrade to a newer version of DOS without changing the BIOS. PC- BIOS handles Plug-and-play (PnP) devices, which are known as PnP BIOS, or PnP-aware BIOS. These BIOS are always implemented with flash memory rather than ROM.

BIOS Set Up and Hardware Components

On start-up, the BIOS tests the system, i.e., it prepares the computer for operation based on the installed hardware and the configuration settings from the manufacturer and the user. For example, it initializes memory caches and assigns resources to all devices connected to the PCI bus. The BIOS searches for BIOS extensions on plug-in cards and sets up pointers which interrupt vectors in the main memory to access those routines. It then loads the operating system and passes control to it. The BIOS set-up is called the CMOS set-up or the CMOS RAM, because user

settings are initially held in a tiny, battery-backed CMOS memory bank that is part of the PC's real time clock circuit. Subsequently, more user configuration settings are generally stored in the BIOS flash memory.

Standard CMOS Features	Frequency/Voltage Control
Advanced BIOS Features	Load Fail-Safe Defaults
Advanced Chipset Features	Load Optimized Defaults
Integrated Peripherals	Set Supervisor Password
Power Management Setup	Set User Password
PnP/PCI Configurations	Save & Exit Setup
PC Health Status	Exit Without Saving
Esc : Quit	
F10 : Save & Exit Setup	
Virus Protection, Boot Sequence...	

NOTES

The BIOS settings control the PC to save the work in non-volatile memory so that they are preserved even when the machine is off. This is opposed to regular system memory. And it is cleared each time the users turn off the PC. A special type of memory is used to store this information, called CMOS memory. For this, a very small battery is used to trickle a small charge to it which makes sure that the data which it holds, is always preserved. These memories are very small, typically 64 bytes, and the batteries that they use typically last for years. This non-volatile memory is sometimes called Non-Volatile Random Access Memory (NVRAM). The system uses something called a CMOS checksum as an error-detecting code. Each time users change the BIOS settings, the checksum is generated by adding together all the bytes in the CMOS memory. And then it is stored as the lowest byte of the sum. Then, each time the system is booted, the system recomposes the checksum and compares it to the stored value. If they are different, then the system knows that the CMOS is corrupted somehow and gives an error as CMOS checksum error.

Hardware Component of BIOS

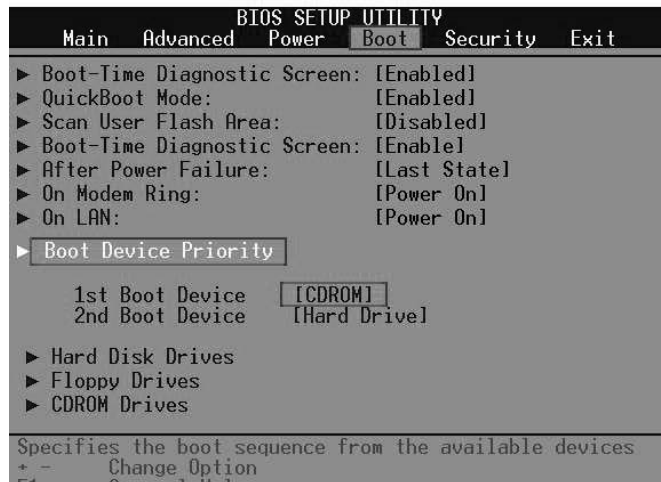
The main hardware component of the system BIOS is the system BIOS ROM itself. This is normally located in an Electrically-Erasable Read-Only Memory (EEPROM) chip, which allows it to update through software control. This is commonly called flash BIOS. Under normal circumstances, the BIOS ROM is permanent. If for some reason the BIOS ROM is corrupted due to an aborted flash update, then PC is left in a state so that it would not be rebooted. In this situation, BIOS ROM can be physically replaced.

BIOS Configuration

Before beginning the BIOS configuration, it is necessary to get into the BIOS set-up utility. The BIOS Set-up Utility is entered during the booting-up of the system. The first thing that users must do is to press the right key in order to get into BIOS, such as [F1], [F2], [Alt], [Ctrl] keys or other keys. During boot up, the BIOS entrance page flashes on the screen for a second or two. The distinct tasks of the BIOS set-up program are as follows:

NOTES

- It tests the computer's main components by running its power-on-self-test (POST) program to make sure that they are all functioning properly.
- It configures the main components that are part of the motherboard or that are attached to it so that the operating system knows what to do with them. The configuration role is essential to the operation of a computer, but it decreases all the time as the operating system continues to take over more and more of this role.
- It boots the operating system from the primary hard disk drive or from a bootable disk at start-up.
- It provides access to some of the computer's components and features such as the keyboard even when the operating system is running. The BIOS is programmed to run a small routine that handles the keyboard at the behest of the operating system. In this way, the BIOS code provides a uniform interface to the system in order that variable hardware can function without having to make the necessary changes to the operating system when the type of hardware changes. Windows 2000, Windows XP and Windows Vista only make a small number of such calls to the BIOS, but the Linux operating system makes no such calls while running, because it deals directly with the standard hardware it is likely to encounter.
- It uses device-driver software to deal with any non-standard or special hardware it encounters.



Each manufacturer has a different design and arrangement to its BIOS set-up utility. There are several approaches to BIOS settings. Some like to use BIOS settings to make their system run exactly the way they prefer. There is certainly merit to that approach. However, this is an initial configuration of BIOS, so it is wise to take a conservative approach. Many BIOS Utilities have a Default Section where we can set all default settings automatically without the need to attend to individual settings. BIOS Help is available after pressing the [Alt+H] key combination. The following features are available to enable or disable as per BIOS set up requirement:

Virus Program: It scans the hard drive at start-up and warns us of trouble.

Cache: It is enable and also, if available, enable ECC Error Checking.

Quick POST: It allows BIOS to boot up faster by skipping some tests by disabling or enabling.

Boot sequence: It controls the order in which bootable information is searched.

Swap floppy drives - This is usually disabled. It is enabled if users have two floppy drives and wish to assign the A and B drive letters interchangeably.

Fast A20 Gate: It is left at default.

Numlock: It is kept as enable.

CPU Serial Number: It is kept as disable.

Video BIOS Shadow: It is disabled by default.

The following rules can be followed for setting up standard CMOS:

- The system date and time should be correct.
- The floppy disk drive is set to 1.44MB, 3.5 inches.
- Video is set to VGA/EGA.
- In HALT state, select ON to stop the PC from repeatedly rebooting.
- Set IDE for auto-detection.

CMOS stands for Complementary Metal Oxide Semiconductor. CMOS technology is used to make semiconductors integrated circuits such as processors, chipset chips and DRAM. CMOS has the advantage of requiring very little power, compared to some other semiconductor technologies. This type of memory came to be called just CMOS. Ironically, with today's processors, they themselves are typically made entirely with CMOS technology. However, CMOS itself refers to the BIOS settings memory.

```
Virus Warning           [Disabled]
CPU Internal Cache      [Enabled]   Item Help
External Cache         [Enabled]
CPU L2 Cache ECC Checking [Enabled]
Processor Number Feature [Enabled]   Select Your Boot
Quick Power On Self Test [Enabled]   Device Priority
First Boot Device       [CDROM]
Second Boot Device      [Floppy]
Third Boot Device       [HDD-0]
Boot Other Device       [Enabled]
Swap Floppy Drive       [Disabled]
Boot Up NumLock Status  [On]
Gate A20 Option         [Fast]
Ata 66/100 IDE Cable Msg. [Enabled]
Typematic Rate Setting  [Disabled]
Security Option         [Setup]
OS Select For DRAM > 64MB [Non-OS2]
```

NOTES

NOTES

Check Your Progress

7. State any one of the goals of object-oriented languages.
8. What are the major objectives of an operating system?
9. When was the earliest electronic computer developed?
10. Give few examples of single user operating systems.
11. How can you access BIOS help?
12. What is the use of CMOS technology?

2.5 ANSWERS TO ‘CHECK YOUR PROGRESS’

1. A number system that uses only two digits, 0 and 1, is called the binary number system.
2. The methods used for converting a decimal number to its octal equivalent are the same as those used to convert decimal to binary. To convert a decimal number to octal, we progressively divide the decimal number by 8, writing down the remainders after each division. This process is continued until zero is obtained as the quotient, the first remainder being the LSD.
3. A computer cannot operate without any instructions and is based on a logical sequence of instructions in order to perform a function. These instructions are known as a ‘computer program’, and constitute the computer software.
4. Software quality can be affected by two categories of factors:
 - Factors that can be directly measured (e.g., defects per function point)
 - Factors that can be measured only indirectly (e.g., usability or maintainability)
5. A computer language essentially implies a language that is understandable to the computer. It is the computer’s native language. Computer languages serve the same purpose as human languages. They are a means of communication.
6. The computer can understand only a binary-based language. This is a combination of 0s and 1s. Instructions written using sequences of 0s and 1s constitute as are known as machine language.
7. One of the goals of object-oriented languages is the creation of flexible programs.
8. The three main objectives of an operating system are to:
 - Make the use of computer easy and convenient
 - Provide for efficient allocation and utilization of resources
 - Include the ability to evolve and adapt with new devices and hardware without too much problem

9. The earliest electronic computer was developed in the mid-1940s and had no operating system at all.
10. Examples of single user operating systems are MS DOS, Windows 95, Windows NT, Windows 2000 and Windows Vista, each of which runs on Intel Processor-based computers.
11. BIOS Help is available after pressing the [Alt+H] key combination.
12. CMOS technology is used to make semiconductors-integrated circuits, such as processors, chipset chips and DRAM.

NOTES

2.6 SUMMARY

- The decimal number system contains the numbers 0 to 9, whereas the binary system contains two numbers: 0 and 1.
- A number system that uses only two digits, 0 and 1, is called the binary number system. The binary number system is also called a base two system. The two symbols, 0 and 1, are known as bits (binary digits).
- Binary number system is used in digital systems. The octal and hexadecimal number systems are also used in digital systems and computers have efficient ways of representing binary quantities.
- A computer cannot operate without any instructions and is based on a logical sequence of instructions in order to perform a function. These instructions are known as a 'computer program', and constitute the computer software.
- Software requirements are the foundation for measuring the quality. If requirements are not properly specified there will be a lack of quality.
- FURPS is set of software quality factors that have been developed by Hewlett-Packard. FURPS stands for functionality, usability, reliability, performance and supportability.
- A computer language essentially implies a language that is understandable to the computer. It is the computer's native language. Computer languages serve the same purpose as human languages. They are a means of communication.
- The computer can understand only a binary-based language. This is a combination of 0s and 1s. Instructions written using sequences of 0s and 1s constitute as are known as machine language.
- FORTAN (FORmula TRANslation) was the first high-level language developed by John Backus at IBM in 1956.
- One of the goals of object-oriented languages is the creation of flexible programs.

NOTES

- Operating systems hide the idiosyncrasies of hardware by providing abstractions for ease of use. Abstraction hides the low-level details of the hardware and provides high-level user-friendly functions to use a hardware.
- The input/output devices of a computer include keyboard, monitor, mouse, pen, printers, scanners, secondary storage devices like hard disk, floppy disk, CD ROMs, etc.
- Operating systems have evolved through a number of phases during the last half a century. The earliest electronic computer was developed in the mid-1940s and had no operating system at all. Machines (computers) were equipped with a console consisting of toggles switches and display lights.
- Batch processing program was developed in the mid-1950s. This helped to make the setup that was required for executing a job automatically and executing the programs (submitted for execution at the computer centre in the form of card decks) of different users one after another without user intervention.
- BIOS is acronym for Basic Input/Output System, the built-in software that determines what a computer can do without accessing programs from a disk.
- On PCs, the BIOS contains all the code required to control the keyboard, display screen, disk drives, serial communications and a number of miscellaneous functions. The BIOS is typically placed in a ROM chip that comes with the computer and it is often called a ROM-BIOS.

2.7 KEY TERMS

- **Decimal Number System:** The number system that utilizes ten distinct digits, i.e., 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9.
- **Binary Number System:** A number system that uses only two digits, 0 and 1.
- **Efficiency:** This is the amount of computing resources and code required by a program to perform its function.
- **Correctness:** The level up to which a program satisfies its specification and fulfills the customer's objectives.
- **Reliability:** The level to which a program can be expected to perform its intended function without failure.
- **Integrity:** Efforts required for controlling the unauthorized access to software or data by persons.
- **C:** This language was developed by Dennis Ritchie of Bell Laboratories in order to implement the operating system UNIX.

- **Java:** Java is again an object-oriented language like the C++ but is a simplified version with extra features.
- **I/O Devices:** Refer to those devices that provide a means of communication between the computer and the external world.

NOTES

2.8 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is decimal number system?
2. Write the needs of software.
3. State about the ISO 9126 quality factors.
4. What are high-level languages?
5. Define the term object oriented language.
6. Differentiate between a network operating and distributed operating system.
7. What is the batch processing program/system?
8. Write a short note on MS-DOS.

Long-Answer Questions

1. Explain the various types of number system with examples.
2. Discuss about the conversion of binary, octal and hexadecimal system.
3. Describe the McCall's software quality factors giving appropriate example.
4. Briefly explain the various types of software with the help of examples.
5. Discuss briefly programming language and its types with the help appropriate examples.
6. What are the various services provided by an OS? Explain with the help of suitable examples.
7. Discuss briefly various types of operating system.
8. What do you understand by 'BIOS'? Explain its working and significance in personal computers.

2.9 FURTHER READING

Shepherd, Robert D. *Introduction to Computers and Technology*. EMC/Paradigm Publishing, 1998.

Kumar, A. Anand. *Fundamentals of Digital Circuits*. New Delhi: Prentice-Hall of India, 2003.

NOTES

Yarborough, John M. *Digital Logic Applications and Design*. Thomson Learning, 1996.

Silberschatz, Abraham, Peter Galvin and Greg Gagne. *Operating System Concepts*. John Wiley & Sons, 2004.

Tanenbaum, Andrew S. *Modern Operating Systems*. Prentice Hall, 2001.

Mano, M. Morris. *Computer System Architecture*. New Jersey: Prentice Hall Inc., 1993.

Peterson, J.L., Abraham Silberschatz and Peter B. Galvin. *Operating System Concepts*. New York: Addison Wesley, 1994.

UNIT 3 DOS, UNIX AND WINDOWS OPERATING SYSTEM

NOTES

Structure

- 3.0 Introduction
- 3.1 Objectives
- 3.2 Disk Operating System (DOS)
 - 3.2.1 DOS Command Types—Internal and External
- 3.3 Unix Operating System
 - 3.3.1 Architecture of UNIX
 - 3.3.2 Features of UNIX
 - 3.3.3 Working with Commands
- 3.4 Windows 95/98
- 3.5 Answers to ‘Check Your Progress’
- 3.6 Summary
- 3.7 Key Terms
- 3.8 Self-Assessment Questions and Exercises
- 3.9 Further Reading

3.0 INTRODUCTION

DOS has a command line, text based non-graphical user interface commonly referred to as character based user interface. Booting is the very first process after switching on the computer. A filename is divided in two parts, such as base name, i.e., the primary filename and the extension, i.e., indicating the type of file related to a certain format of file. Internal DOS commands, such as DIR, CD, MD, RD, etc., as well as external DOS commands, such as ATTRIB, RESTORE, BACKUP, XCOPY, etc. A file or directory is located with a relative pathname. An absolute pathname tells you the path of directories you must travel to get from the root of directory or file you want. A DOS printer will often include a DOS driver or other DOS printing software in order to be compatible with the older operating system.

In 1960, the Massachusetts Institute of Technology, AT&T Bell Laboratories and General Electric worked on an operating system, MULTiplexed Information and Computing Service (MULTICS). The MULTICS operating system was designed to run the GE-645 mainframe computers. The operating system was not very successful. Ken Thompson, who was a developer in the AT&T Bell Laboratories, developed a game called Space Travel for the GE-645 mainframe computers. The system was running at a very slow speed. Then, he wrote the game in the assembly language with the help of Dennis Ritchie.

In this unit, you will study about the Disk Operating System (DOS), directories, files, wild cards, DOS commands, redirection option in DOS, features of Unix operating system, multitasking, multi-user capability, Unix connections, components of the Unix operating system, hierarchical file system of Unix, basic directory commands, basic file commands, shell programming and windows 95/98.

3.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Understand the basics of Disk Operating System (DOS)
- Understand the directories, files and wild cards in DOS
- Discuss the DOS commands
- Describe the various features of Unix operating system
- Explain the shell programming
- Elaborate on windows 95/98

3.2 DISK OPERATING SYSTEM (DOS)

Microsoft Disk Operating System (MS DOS) is a single user, single tasking operating system. DOS has a command line, text based/non-graphical user interface commonly referred to as Character-based User Interface (CUI). When the computer is switched on, a small program checks all internal devices, electronic memory and peripherals. Once this process is completed, MS DOS is loaded.

The DOS prompt known as the command prompt looks like C:\> or D:\> where 'C', 'D' represent the hard drives of the computer system. All commands are typed at the DOS prompt. ENTER key is pressed to view the output of the typed command. If the command is correctly typed desired output would be displayed. Otherwise an error message (bad command or filename/invalid parameter) is displayed on the screen.

Limitations of MS DOS

1. It has a text based user interface where the commands have to be typed for each operation that the user wants to perform. The user is expected to remember the commands as well as their syntax.
2. It is a single user, single task operating system and the working is limited to one megabyte of memory. 640 kilobytes of the memory is used for the application program.
3. It does not allow using long file names. The user is restricted to eight character file names with three character extensions.

Below is the list of commonly used DOS commands. All DOS commands are case-insensitive, i.e., there is no difference whether you type COPY, copy or coPY. The same is true for the attributes, parameters and filenames. To view the complete description of each command type HELP followed by the command name, for example, HELP COPY. Alternatively help can also be displayed using <command>/? (COPY/?).

File Naming Conventions and Different Types of Files in DOS

A filename is a special type of string which is used for identifying a file stored on the file system of a computer.

A filename is divided in two parts: basename (the primary filename) and the extension (which indicates the type of file related to a certain format of file).

A filename includes one or more of the following components:

- Protocol or scheme: Method of accessing, for example, ftp, http, file, etc.
- Host or network ID: Host name, IP address, domain name, or the network name of LAN, for example, wikipedia.org, 207.142.131.206, \\MYCOMPUTER, SYS:, etc.
- Device or node: Port, socket, drive, root mountpoint, disc, volume, for example C:/, SYSLIB, etc.
- Directory or path: Directory tree, for example /usr/bin, \TEMP, [USR.LIB.SRC], etc.
- File: Base name of the file.
- Type (format or extension): Indicates the nature / type of the content of the file, for example, .txt, .exe, .dir, etc.
- Version: Revision number of the file.

NOTES

An extension name of a computer file is usually a three character addition after the file's name. This extension assists IBM compatible computers, such as those running Microsoft Windows, in identifying the program to relate the file with and correctly open the file.

Given below are the characters that cannot be used in filenames or directory names in most operating systems.

\, /, :, *, ?, ", <, > and |

System Files

The system files are used to run the system unit. These files are created and maintained by the operating system and are necessary for the system to perform necessary functions. System files are used by many commands and subroutines to perform operations. Before knowing the details of DOS system files, the concept of file is essential for beginners. On MS DOS systems, the system files are known as IO.SYS and MSDOS.SYS. A file is a collection of data that can be read from or written to. A file can be a program you create, text you write, data you acquire or a device you use. Commands, printers, terminals and application programs are all stored in files. This allows users to access diverse elements of the system in a uniform way and gives the operating system great flexibility. No format is implied when a file is created. Files are used for all I/O of information in this operating system. This standardizes access to both software and hardware. Input occurs when the content of a file is modified or written to. Output occurs when the content of one file is read or transferred to another file. For example, to create a hardcopy printout of a text file, the system reads the information from the text file and writes the data to the file representing the printer. Collections of files are stored in directories. These collections of files are often related to each other and storing them in a structure of directories keeps them organized. There are many ways to create, use and manipulate files. There are three basic types of files as summarized in Table 3.1.

Table 3.1 Types of Files and their Function

NOTES

<i>File Type</i>	<i>Function</i>
Regular	This file type is used to store data (text, binary and executable).
Directory	This file type is used to contain information used to access other files.
Special	This file type is used to define a First In, First Out (FIFO) file or a physical device.

All file types recognized by the system fall into one of these categories. However, the operating system uses many variations of these basic types. The modern DOS is distributed on 3-5 high density floppy disks. It comes with backup utilities and disk compression drivers. The files that goes into C:\DOS and its subdirectories are programs and utilities. The boot sector is a 512 byte record placed at the beginning of the C: drive when DOS was installed or placed there subsequently using the 'SYS C:' command. Two 'hidden' files are stored in the root directory of the C: drive. They do not show up in a DIR listing unless the /A switch is used. On IBM PC DOS systems, they are IO.SYS and IODOS .SYS. On MS DOS systems, they are called IO.SYS and MSDOS.SYS. These files are used to create the kernel of the DOS system. COMMAND.COM is the 'shell' or command interpreter. It also supports .BAT files. The user configuration files are CONFIG.SYS and AUTOEXEC.BAT. The reader is assumed to be familiar with these files. COMMAND.COM is initially stored in the C:\ root directory. The problem is that OS/2 and Windows NT have their own versions of COMMAND.COM. To avoid confusion, each COMMAND.COM should be stored in the subdirectory that belongs to its particular operating system. In normal use, this means that the DOS version should be in C:\DOS. IO.SYS locates MSDOS.SYS as follows. IO.SYS searches the root directory for MSDOS.SYS and makes a note of the starting cluster. IO.SYS loads the part of MSDOS.SYS located in that cluster into memory. IO.SYS checks the appropriate File Allocation Table (FAT) entry for a 'next in chain' link or 'end of chain' and continues to load the Cluster contents into memory until it reaches the 'end of chain' entry. IO.SYS instructs the CPU to execute the contents of memory where the whole MSDOS.SYS is located. OS boot process continues in a similar manner. To relocate it, two statements must be added to the user configuration files:

SHELL=C:\DOS\COMMAND.COM is added to CONFIG.SYS.

SET COMSPEC=C:\DOS\COMMAND.COM is added to AUTOEXEC.BAT.

The hidden files IBMBIO, IBMDOS, IO or MSDOS have names that do not conflict with each other or with any system file belonging to any other operating system. They can stay in the C:\ root directory no matter what gets added to the system. This means that the volatile part of the DOS system consists of the boot record, C:\CONFIG.SYS and C:\AUTOEXEC.BAT.

Putting DOS on a Floppy

To install a copy of the DOS system files on a floppy, simply issue the following command:

FORMAT A: /S

Recovering Damaged MS DOS System Files

If MS DOS system files have been corrupted during booting process then it can be recovered from MS DOS backup files. It is necessary to keep the exact version of MS DOS file. Type SYS C: command at command prompt. In MS DOS, the sectors of a DOS disk are divided into the groups which are listed in order of increasing logical sector numbers.

These groups are described in detail as follows:

Reserved Sectors

The first reserved sector, which is always logical sector 0, is also called the bootstrap sector. Table 3.2 shows the details of 512 byte bootstrap sectors.

Table 3.2 Details of 512 Byte Bootstrap Sectors

<i>Bytes</i>	<i>Function</i>
0-2	It refers to first instruction of bootstrap routine.
3-10	It refers to OEM name.
11-12	It represents the number of bytes per sector.
13	It represents the number of sectors per cluster.
14-15	It represents the number of reserved sectors.
16	It represents the number of copies of the file allocation table.
17-18	It represents the number of entries in root directory.
19-20	It refers to total number of sectors.
21	It refers to media descriptor byte.
22-23	It represents the number of sectors in each copy of file allocation table.
24-25	It represents the number of sectors per track.
26-27	It represents the number of sides.
28-29	It represents the number of hidden sectors.
30-509	It refers to bootstrap routine and partition information.
510	It represents the hexadecimal notation.
511	It also represents the hexadecimal notation.

The first instruction of the bootstrap routine is always a jump instruction that transfers control to the bootstrap routine in bytes 30-509. The Original Equipment Manufacturer (OEM) name is usually the version of DOS or the name of the utility that is used to format the disk, represented by eight American Standard Code for Information Interchange (ASCII) characters. Reserved sectors other than the bootstrap sector are not used by MS DOS. MS DOS disks have only one reserved sector but a larger number may be used to reserve space for a non-DOS partition or to cause DOS to avoid one or more bad sectors.

File Allocation Tables

The sectors immediately after the last reserved sector hold one or more copies of the file allocation table sometimes called the FAT. The number of copies is specified

NOTES

NOTES

by byte 16 in the bootstrap sector and the number of sectors in each copy is specified by bytes 22 and 23 in the bootstrap sector. MS DOS uses only the first copy of the file allocation table but it updates other copies, if any, to keep them identical to the first copy. Then if the first copy develops a bad sector, a special file recovery program can be used to retrieve the information from another copy. Most floppy and hard disks contain two copies of the FAT. Different versions of MS DOS use a file allocation table with 12-bit entries. Each group of three consecutive bytes contains two 12-bit entries arranged as follows:

- The first byte contains the eight least significant bits of the first entry.
- The four least significant bits of the second byte contain the four most significant bits of the first entry.
- The four most significant bits of the second byte contain the four least significant bits of the second entry.
- The third byte contains the eight most significant bits of the second entry.

FAT entries are numbered consecutively. Entry 0 (the first entry) contains the media descriptor byte (byte 21 in the bootstrap sector), padded at the more significant end with ones to make it a 12-bit or 16-bit value. Entry 1 contains either 12 or 16 ones. For file allocation purposes, the data area of the disk is divided into a number of clusters. Each cluster is the same size and consists of the number of consecutive sectors specified in byte 13 of the bootstrap sector. The clusters are numbered consecutively but it starts with cluster number 2 which lies at the beginning of the data area. Each entry in the file allocation table, except entries 0 and 1 is associated with the cluster with the same number as the entry. Numbers are arranged in sequential order because there are no clusters with number 0 or 1. Table 3.3 shows how the contents of each entry are interpreted:

Table 3.3 Hexadecimal Representation for 12-Bit and 16-Bit

<i>12-bit Representation</i>	<i>16-bit Representation</i>	<i>Function</i>
000	0000	The cluster is unassigned and available.
001	0001	It refers to invalid entry.
002-FEF	0002-FFEF	If cluster is assigned, the number of the next cluster is applied in the same file or subdirectory.
FF0-FF6	FFF0-FFF6	The cluster is reserved.
FF7	FFF7	This cluster contains a bad sector.
FF8-FFF	FFF8-FFFF	This cluster is the last cluster assigned to a file or subdirectory.

The 12-bit or 16-bit entries are used on the number of clusters specified in the data area. If the range from 002 to FEF is sufficient then all 12-bit entries are used otherwise 16-bit entries are used. Each subdirectory and each file that contains at least one byte of data is associated with a chain of file allocation table entries, one for each cluster. The directory entry for the file or subdirectory contains the number of the first cluster. The entry for each cluster other than the last cluster contains the number of the next corresponding cluster. The entry for the last cluster contains a value in the range FF8-FFF and FFF8-FFFF which marks the end of the chain.

NOTES

Root Directory

The root directory contains one 32-byte entry for each file in the root directory of the disk. The root directory is provided by the operating system and has a special name, for example in DOS systems the root directory is called '\'. The root directory is sometimes referred to simply as the root. Bootstrap sector is always chosen for the root directory sectors. In a computer file system, root directory is organized as a hierarchy or tree. The root directory includes all other directories. In MS DOS, the root directory has no name. It is simply represented by the special character that separates directories in a file system.

The Data Area and Hidden Sectors

The total number of sectors given in bytes 19 and 20 of the bootstrap sector includes all reserved sectors and all sectors in the root directory, the data area and all copies of the file allocation table. It does not include the hidden sectors. The size of the data area is not given directly, but it can be calculated from this information. Apparently it must be an exact multiple of the cluster size. The total number of sectors, plus the number hidden sectors given in bytes 28-29 of the bootstrap sectors, is the true total number of sectors in the partition. This is equal to the product of the number of sectors per track, the number of tracks per side and the number of sides. The number of tracks per side is not given but it can be easily calculated from this information.

Directory Structure and Protocol

File name and extension are assigned in the specific volume name. The volume name is case sensitive and only small letters are accepted. The name is always 11 characters long and is apparently also padded with blanks to fill out its field. When a subdirectory, file and volume label is created MS DOS writes the current time and date into the time and date fields. Bytes 28-31 (number of bytes in file) for a subdirectory entry are always zero. The length of a subdirectory is not recorded in its directory entry and is always contains at least one cluster. Its length can be determined only by examining the chain of file allocation table entries. Every entry in the root directory is considered as terminating entry if a disk is to be formatted. There cannot be two entries in the same directory with same name and extension unless they are deleted entries or one is a volume label and the other is not a volume label. If the creation of a new file or subdirectory would cause a conflict, MS DOS usually fails the operation. Table 3.4 shows the reserved bytes and their function.

Table 3.4 Reserved Bytes and their Function

NOTES

<i>Bytes</i>	<i>Function</i>
0-7	These bytes are reserved for file name or first eight characters of volume name.
8-10	These bytes are file extension or last three characters of volume name.
11	These bytes are called attribute byte.
12-21	These represent unused byte.
22-23	They represent time.
24-25	They represent date.
26-27	These bytes are reserved for number of first cluster.
28-31	They represent number of bytes in file or zero for subdirectory or volume label.

3.2.1 DOS Command Types—Internal and External

DOS commands are of two types known as internal and external. The internal commands do not require any special files for being executed and are brought into the computer’s memory as soon as the computer is switched on, for example COPY, DEL, DIR, TYPE, CD, MD, CLS, DATE, TIME and PATH. External commands are commonly external either because they require large requirements or are not commonly used commands.

Wildcard Characters

Wild card characters can be used in specifying filenames in DOS. There are two types of wild cards (?,*):

? It is used to represent any single character in the file name.

SYNTAX: C:\DIR BA?.TXT

Displays all the text files in the C: drive starting with ‘BA’ and ending with any single character.

Examples: BAT.TXT, BAG.TXT, BAR.TXT, BAD.TXT, etc.

* It is used to represent one or more characters in a file name.

SYNTAX: C:\DIR CON*.TXT

Displays all the text files in the C: drive starting with ‘CON’.

Examples: CONCEPT.TXT, CONCATENATE.TXT, CONTEMPT.TXT, CONSOLE.TXT, etc.

Internal Commands

The DOS internal commands are called internal because their instructions are a part of COMMAND.COM and executed at DOS command line interpreter. The internal commands reside in COMMAND.COM, which loads into memory when the computer system is started therefore these commands do not reside on disk. Internal commands are BREAK, CALL, CHCP, CHDIR(CD), CLS, COPY, CTTY, DATE, DEL (ERASE), DIR, ECHO, EXIT, MKDIR(MD), PATH, PAUSE, PROMPT, REM, RENAME (REN), RMDIR (RD), SET, TYPE, VER, VERIFY and VOL.

Command	Syntax	Explanation	Example	Notes
DIR	DIR <Drive/ Directory - Name Name>	Displays all the sub-directories and files of the specified drive/directory. It also shows the size of the files and the date and time they were last modified.	C:\>DIR D:	Displays all the contents (files and directories) of the D: drive.
DIR/P	DIR <Drive/ Directory -Name>/P	Displays the contents of directory one screen at a time and pauses until any other key is pressed to continue the display.	C:\>DIR DATA/P	Displays the contents of the 'DATA' directory by pausing the screen.
DIR/W	DIR <Drive/ Directory>/W	Displays the contents of the directory width wise. It omits file size, date and time of creation of file so that more files can be displayed at one time on the screen.	C:\>DIR DATA/W	Displays the contents of the 'DATA' directory width wise.
DIR/W/P	DIR <Drive/ Directory>/W/P	The Wide and Pause display option can be combined.	C:\>DIR DATA/W/P	Displays the contents of the 'DATA' directory width wise and by pausing the screen.
CD	CD <Directory- Name> CD\ -Directly takes to the root directory.	Displays the name of the current directory if no parameter is specified with the command. Changes the current directory to the specified directory.	C:\>CDDATA\ SUBDATA	Changes the current directory to 'DATA\SUBDATA'.
MD	MD <Drive/ Directory-Name>	Creates a new directory in the specified location.	C:\>MD 'HELLO'	Creates a directory named 'HELLO' in the C: Drive.
RD	RD <Directory- Name>	Removes the specified directory.	C:\>RD HELLO	To remove a directory first you should come to one level above the current directory and then remove command should be given. This command will delete the 'HELLO' directory from the C:drive.
DATE	DATE	Displays the system's current date and prompts to enter the new date.	C:\>DATE	Current date is Fri 05-09-2003 Enter new date <mm-dd-yy>:
TIME	TIME	Displays the current time and prompts the user to enter the new time.	C:\>TIME	Current time is 12:55:25.34p Enter new time:

NOTES

Modifying the path statement will set the files in Microsoft Windows environment which are required to run the program. For Windows API functions that manipulate file names can often be relative to the current directory while some APIs require a fully qualified path. A file name is relative to the current directory if it does not begin with one of the following:

- A Universal Naming Convention (UNC) name of any format which always start with two backslash characters (\\).
- A drive name with a backslash, for example 'C:\' or 'D:\'.
- A single backslash, for example '\directory' or '\file.txt'. This is also referred to as an absolute path.

A path is also said to be relative if it contains 'double dots' (. .), i.e., two periods represent together in one component of the path. This special specifier is used to denote the directory above the current directory which is known as the parent directory. Examples of this format are as follows:

- Expression '..\tmp.txt' specifies a file named tmp.txt located in the parent of the current directory.
- Expression '..\..\tmp.txt' specifies a file that is two directories above the current directory.
- Expression '..\tempdir\tmp.txt' specifies a file named tmp.txt located in a directory named tempdir that is a peer directory to the current directory.

Relative paths can combine both example types, for example 'C:..\tmp.txt'. This is useful because the system keeps track of the current drive along with the current directory of that drive. It also keeps track of the current directories in each of the different drive letters if more than one system is used regardless of which drive designator is set as the current drive. If a file name begins with only a disk

NOTES

designator but not the backslash after the colon, it is interpreted as a relative path to the current directory on the drive with the specified letter. MS DOS uses PATH= statement which resides in the autoexec.bat file. Additional information about the MS DOS path command is usable in Windows XP whereas additional information about the MS DOS command is found on set command page. The path is managed by Windows XP. Following steps are required to change the system environment variables:

- Select desktop and right click on 'My Computer' where you need to select Properties tab.
- In the System Properties Window, select Advanced tab.
- In the Advanced tab, select Environment Variables button.

In the Environment Variables tab the Path variable is highlighted in the Systems Variable section. This option can be modified if you select Edit button. Add or modify the path lines to access the computer. Path setting along with directory is separated by semicolon as follows:

C:\Program Files;C:\Winnt;C:\Winnt\System32

The function of operating system is to decide whether the path name for a file is absolute or relative to a working directory. Therefore, most relative directories are preceded by a period (.). A pathname prefixes the directory name or filename with a path and a route from one directory to another. Following are the types of path names followed by operating system.

External Commands

The external commands in DOS require certain special DOS files to get executed, for example FORMAT, CHKDSK, XCOPY, PRINT, ATIRIB, DISKCOPY, etc. You may not necessarily have all these files in your system. While copying these files onto your system, you must take care that they are of the same DOS version which is installed in your system. For example, messages like incorrect version may appear on your screen if the DOS version is different from the copied file version.

SYS Command

The SYS command in DOS is used to transfer the operating system files to another disk. The three files that are transferred are IO.SYS, MSDOS.SYS and COMMAND.COM. If you use IBM's version of DOS then they are named IBMIO.SYS, IBMDOS.SYS and COMMAND.COM. The syntax for SYS command is written as follows:

SYS [source] d:

The VOL command in DOS display the volume information about the selected drive and REN command is used to rename files and directories from the original name to a new name. DEL command is used to delete files from the computer.

NOTES

Command	Syntax	Explanation	Example	Notes
PROMPT	PROMPT [Text]	Changes the MS DOS command prompt to the specified text. If the command is typed without any parameters then the default prompt setting is restored.	D:\>PROMPT	Changes the prompt to the default setting.
PRINT	PRINT <Filename>	Prints a text file while other MS DOS commands are being used.	C:\DATA>PRINT TEMP.TXT	Prints 'TEMP.TXT' stored in the 'DATA' folder of the C: drive.
LABEL	LABEL	Makes, changes, or deletes the label of volume of a disk.	C:\>LABEL	Displays the current volume label and volume serial number. Also prompts to enter a new label.
MEM	MEM	Displays the amount of and free and used memory in your system.	C:\>MEM	Displays the total amount of memory, amount of used and free memory in the system.
MORE	MORE <Filename>	Displays output on the screen at a time for the text files. Useful in cases where the content of text file does not fit in a single screen.	C:\DATA> MORE TEMP.TXT	Breaks the contents of 'TEMP.TXT' in multiple screens. Subsequent screens can be viewed by pressing the 'enter' key.
ECHO	ECHO	Displays messages, or turns on or off the echoing command.	C:\>ECHO	Displays the current echo setting. ('OFF' or 'ON')
EDIT	EDIT	Starts MS DOS editor, which produces and changes ASCII files.	C:\>EDIT	Opens the MS DOS editor.
SET	SET	Displays, sets, or removes MS-DOS environment variables.	C:\>SET	Displays the settings for the current environment variables.
CHKDISK	CHKDISK <Drive-Name>	Checks a disk and gives the information like how many bytes have been used and how many are free and if any bad sectors are there on the disk.	C:\>CHKDISK A:	Checks A: drive and gives information about the disk.
SCANDISK	SCANDISK <Drive-Name>	Finds errors from a drive and fixes any problem it encounters.	SCANDISK A:	Scans the A: drive and repairs the disk if any problem is there, like damaged area or virus, etc.
DISKCOPY	DISKCOPY	Copies the content of one floppy disk to another.	DISKCOPY A:	Copies the entire content of one floppy to another.
FORMAT	FORMAT <Drive-Name>	Formats the specified drive.	C:\>Format A:	Floppy inserted in the A: drive will be formatted. Any information in the floppy A: will be erased.

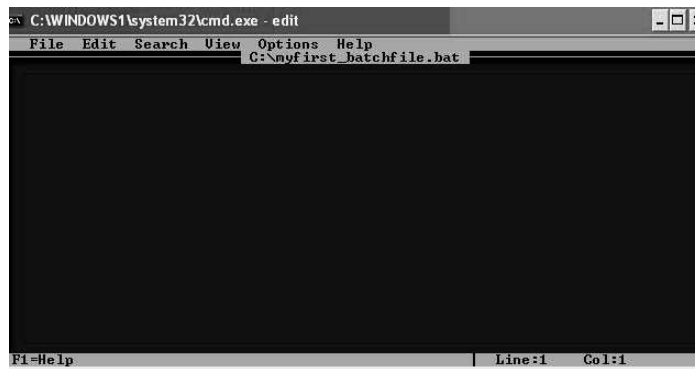
Simple File Operations

Command	Syntax	Explanation	Example	Notes
MOVE	MOVE <Source> <Destination>	Moves the file from one source to the specified destination. File will exist only at the specified destination.	C:\>DATA>MOVE TEST.TXT LETTER	Moves the file 'TEXT.TXT' from 'DATA' folder in the C: drive to the 'LETTER' folder in the C: drive.
COPY	COPY <Source> <Destination>	Creates a copy of the specified file and places it in the specified location, file will exist at the specified location as well as the source location.	C:\>DATA> MOVE HELLO.TXT LETTER	Creates a copy of HELLO.TXT in the LETTER folder of the C: drive.
DEL	DEL <Path><Filename>	Deletes the specified file present in the Specified path/location from the hard disk.	C:\>DATA>DEL Hi.TXT	Deletes the file 'Hi.TXT' located in the 'DATA' folder of the C: drive.
TYPE	TYPE <Filename>	Displays the contents of a text file.	C:\>DATA>TYPE TMP.TXT	Displays the contents of TMP.TXT.
ATTRIB	ATTRIB [+A -A] [+R - R][+H -H] [+S -S] <filename> + Sets an attribute + Clears an attribute A-Archive attribute R-Read only attribute H-Hiddenfile attribute S-System file attribute	Displays or changes file attributes.	C:\>ATTRIB+H+R FIRST.TXT	Sets the attributes of 'FIRST.TXT' as read only and hidden.
FC	FC <File1> <File2>	Compares two files or sets of files and displays the dissimilarities between them. Generally used to compare files with same names located in different directories.	C:\>FCDATA\ ONE.TXT INFO ONE.TXT	Compares the file 'ONE.TXT' located in DATA folder with 'ONE.TXT' located in INFO folder and displays the differences between them.
XCOPY	XCOPY <Source> <destination>	Copies files and subdirectories to the specified location.	C:\>XCOPY C:\DATA C:\INFO	Copies the entire contents of the 'DATA' folder to 'INFO' folder. If the 'DATA' folder contains 'ami' subdirectories then they will also be
BACKUP	BACKUP <Source> <Destination>	Copies the files onto diskettes or to other source.	C:\>BACKUP C:\DATA\TEMP.TXT C:\INFO	Copies 'TEMP.TXT' file in the 'DATA' folder into the 'INFO' folder.
RESTORE	RESTORE<Source> <Destination>	Reinstates files that were backed up by using the BACKUP command.	C:\>RESTORE C:\DATA\TEMP.TXT C:\INFO	Restores 'TEMP.TXT' file in the 'DATA' folder into the 'INFO' folder.

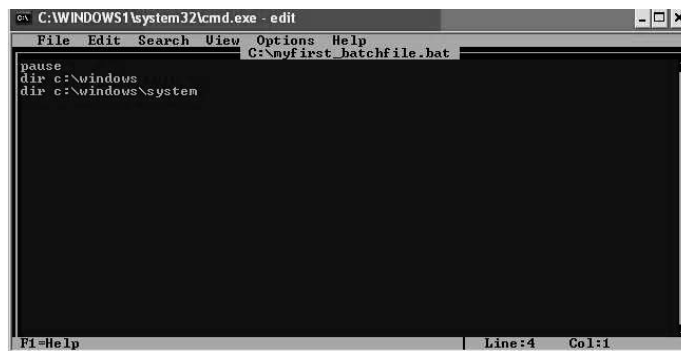
Using COPY CON and EDIT Command for File Creation/ Modification

NOTES

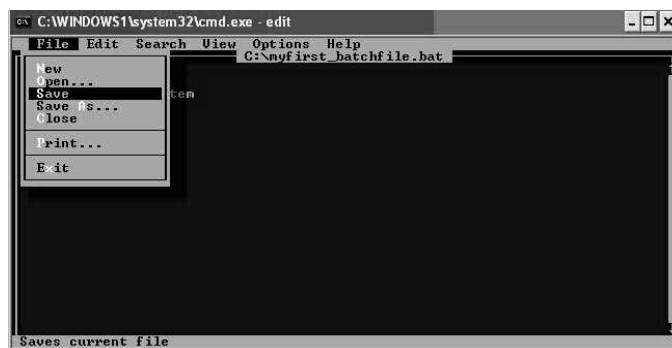
A user can create a file using the COPY CON command which creates the test.txt file. When you have completed creating the file, you can save and exit the file by pressing CTRL+Z which would create ^Z and then press enter key. COPY CON command can also be used to create new files. By copying from the keyboard console (COPY CON:) to the screen, files can be created and then saved to disk. The first filename you enter is referred to as the source file. The second filename you enter is referred to as the target file. If errors are encountered during the copying process, the COPY program will display error messages using these names. To create a batch file you need to go to DOS command prompt and type 'edit myfirst_batchfile.bat'. After pressing ENTER key, you will get the following screen where you type a batch of command:



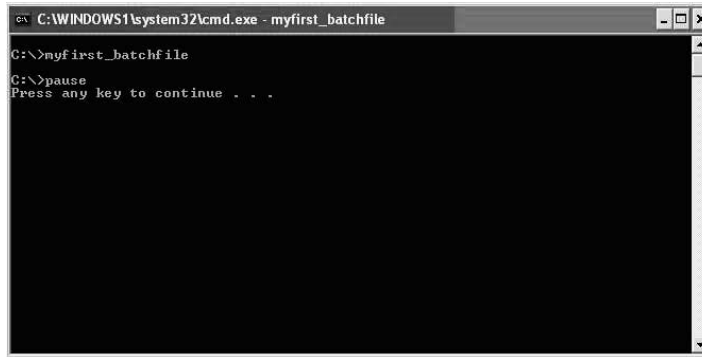
On the Edit screen, type the following command:



Save the changes by clicking on **File**→**Save** option and Exit from Edit screen environment.

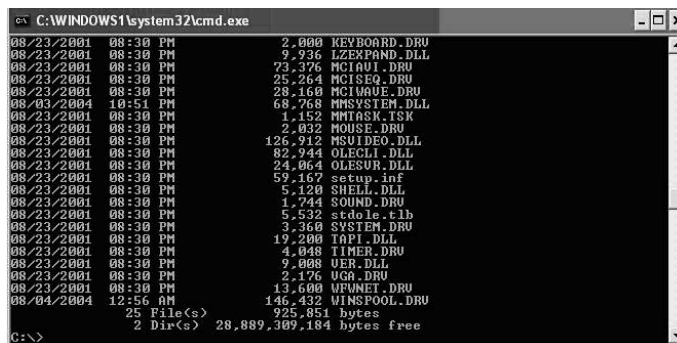


You can see the effect by typing the following command at DOS prompt as follows:

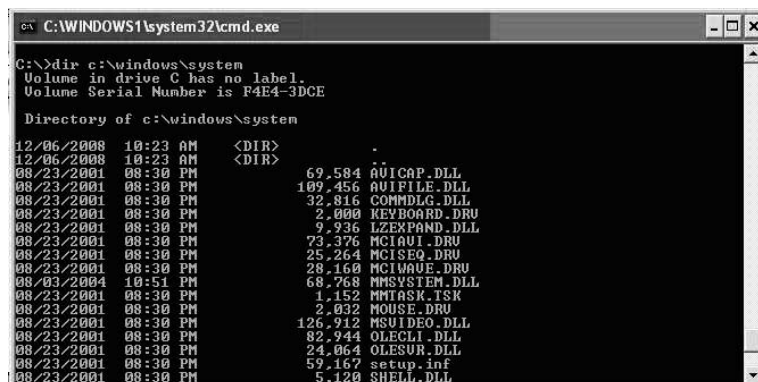


NOTES

After pressing any key, let say ENTER key, the next command 'dir c:\windows' is executed in the batch file program that displays all the directories along with files in '\windows' directory as listed below in the screen:

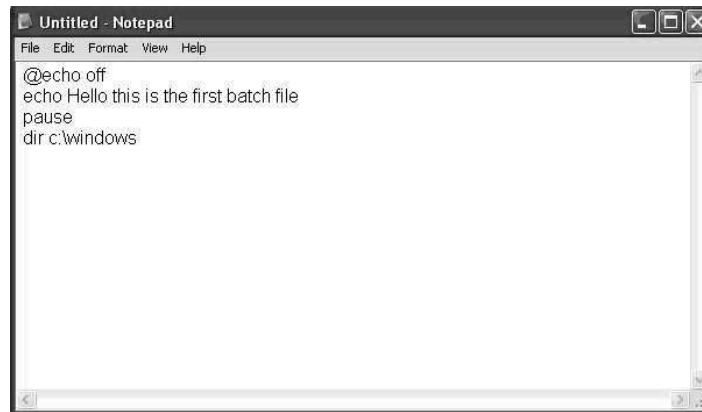


Similarly, the next command 'dir c:\windows\system' is executed in the batch file program which displays all the directories along with files in the \system subdirectory as listed below in the screen:



The Windows users also use the above DOS steps to create a batch file. You can write the batch file using DOS batch commands in either Notepad or WordPad applications with the .bat file name extension. The following steps are required to create a batch file in WordPad:

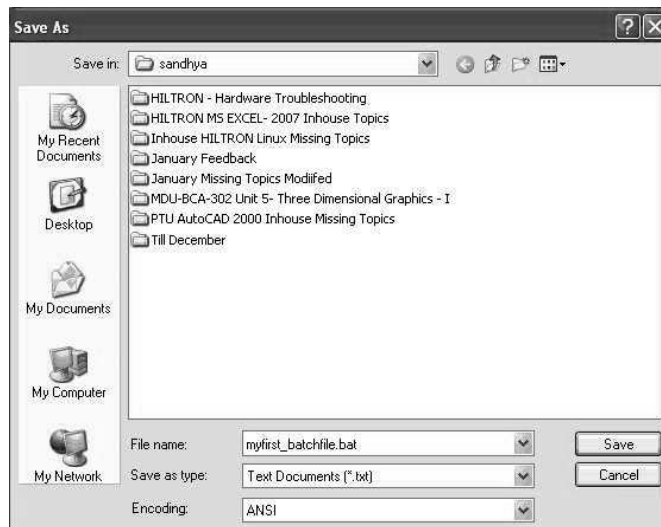
NOTES



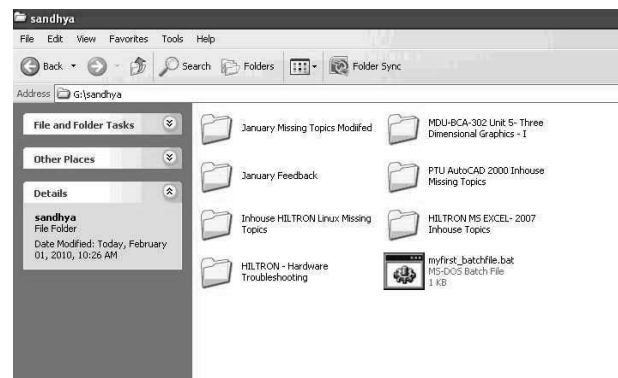
Open Notepad application and type the DOS commands.

In the Notepad, you need to type the following command:

Click on **File**→**Save As** option and type the file name 'myfirst_batchfile.bat' in 'File name:' bar as shown below:



Exit from Notepad application. Run the myfirst_batchfile by double clicking the file.



The very first statement of the batch file is being run as shown in the following screen:



Once the batch file has completed running it will close the Window automatically.

Using DOS Commands with Relative Pathnames

Relative pathname refers to the significance of '.' and '..'. A file or directory is located with a relative pathname. This concept basically gives the location relative to your working directory. Relative pathname can go through more than one directory level by naming the directories along the path. For example, if you are currently in the **users** directory the relative pathname to the **Student** directory then the relative pathname to the **MCA_Student** directory will reside is **Student/MCA_Student**. Neither pathname starts with a slash (/). Relative pathnames start at the working directory which is not the root directory. Therefore, a relative pathname never starts with a slash. In relative pathnames, you can go up the tree with double dot '..' for the parent directory. In either case, such as up or down, each level is separated by a slash (/).

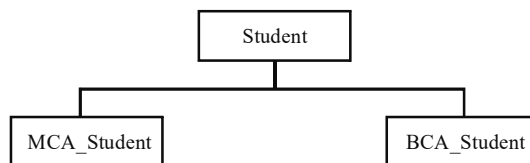


Fig. 3.1 Relative Pathname from MCA_Student to BCA_Student

Figure 3.1 displays that if working directory is **Student** then there are two pathnames for **MCA_Student** subdirectory as well as **BCA_Student**. Absolute and relative pathnames are totally interchangeable. Operating system, such as UNIX follows the path which has been specified to wherever it leads. If you choose relative pathname, the path starts from the working directory.

Absolute Pathname

The file system is used to organize its files and directories in an inverted tree structure with the root directory at the top. An absolute pathname tells you the path of directories you must travel to get from the root of directory or file you want. In a pathname, slashes (/) are put between the directory names. For example, **/users/Student** is an absolute pathname. It locates one and only one directory. Following are the objectives of absolute pathname:

NOTES

- The root is the first '/' directory.
- The directory 'users' represents a subdirectory of the root.
- The directory **MCA_Student** represents a subdirectory of **users**.

NOTES

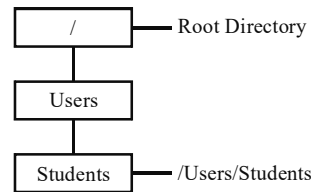


Fig. 3.2 Structure of Absolute Pathname

Figure 3.2, displays the structure of absolute pathname in which 'users' has a subdirectory named **Student**. Its absolute pathname is **/users/Students_BCA_Student**.

Using Internal Commands with Relative Pathname

A path, the general form of a filename or of a directory name, specifies a unique location in a file system. A path points to a file system location by following the directory tree hierarchy expressed in a string of characters in which path components are separated by a delimiting character to represent each directory. The delimiting character is most commonly the slash (/), the backslash character (\) or colon (:).

3.3 UNIX OPERATING SYSTEM

UNIX is a multi-user and multi-tasking operating system. It was developed in the early 1970s at the AT&T Bell Laboratories. There are several versions of UNIX which developed from time to time between 1970 and 2000.

An operating system has a user interface that is used to interact with the user and the command interpreter that helps in converting the user language into a high-level language understood by the computer. In UNIX, the command interpreter is not integrated with the operating system; instead it exists as a separate program. The UNIX operating system provides security for the UNIX users by giving each user his/her own login name and password to work on the UNIX computer.

Versions of UNIX

In 1970, UNIX, which was developed to work on the PDP-7 machines, was used on the PDP-11/20 machine for adding text formatting programs. A text formatting program, called roff and a text editor was added to the operating system. Roff was the first electronic publishing program with the capability of typesetting. The first manual of UNIX was printed on 3 November 1973. This was the first version of UNIX.

In 1973, it was decided to write UNIX in the C programming language. UNIX was made available for universities and commercial firms. The versions of UNIX were determined by the printing of manuals. Some extra features such as

increased development speed, and pipes were added in the 4th, 5th and 6th versions of UNIX.

The UNIX/32V was released for the Virtual Address eXtension (VAX) systems, which are the 32-bit computing architectural systems, in 1978. The version 7 UNIX was released in 1979.

In 1982, AT&T released UNIX System III based on the version 7 UNIX. Similarly, UNIX system V was introduced that contained the vi editor. AT&T added various features such as file locking, system administration and remote file system to the UNIX system V.

During 1993, the UNIX system V Release 4 was introduced and it was sold to Novell. Novell developed its own version, called UNIXWare. Novell sold all the certification rights of UNIXWare to the X/Open consortium. In 1996, X/Open merged with Open Software Foundation (OSF) and formed the Open Group.

In 2000, Novell sold the base code of UNIX to Santa Cruz Operation (SCO), which in turn sold it to the Caldera systems, also known as the SCO group. The Sun Microsystems released the OpenSolaris project, which was an open source project containing the system code based on the UNIX system V Release 4.

UNIX is available in various versions. Table 3.5 lists the various versions of the UNIX operating system.

Table 3.5 Versions of UNIX

Product	Company
Xenix	Microsoft Corporation
FreeBSD, NetBSD and OpenBSD	Berkeley Software Design
Solaris	Sun Microsystems
AIX	IBM
HP-UX	Hewlett-Packard/Packard
Digital UNIX	Digital Equipment Corporation
Tru64 UNIX	Digital Equipment Corporation
IRIX	Silicon Graphics
SCO Open Server	Santa Cruz Operation
SCO UNIXWare	Santa Cruz Operation
Linux	Red Hat, SuSE, Caldera and many other companies

The Xenix and the SCO Open Server products are discontinued. The SCO Open Server product is replaced by the SCO UNIXWare product. The FreeBSD,

NOTES

NOTES

NetBSD, and OpenBSD are commercial products started by the people who created BSD UNIX. The Linux is the clone of UNIX, gaining enormous popularity. The UNIX version that you are using is the FreeBSD product of Berkeley Software Design company.

Overview of Implementation of functions in UNIX Operating System

UNIX is a multi-user and multitasking operating system. In a multi-user environment, the computer can receive the commands from a number of end users to run programs, access files, and print documents simultaneously. Figure 3.3 shows a multi-user environment.

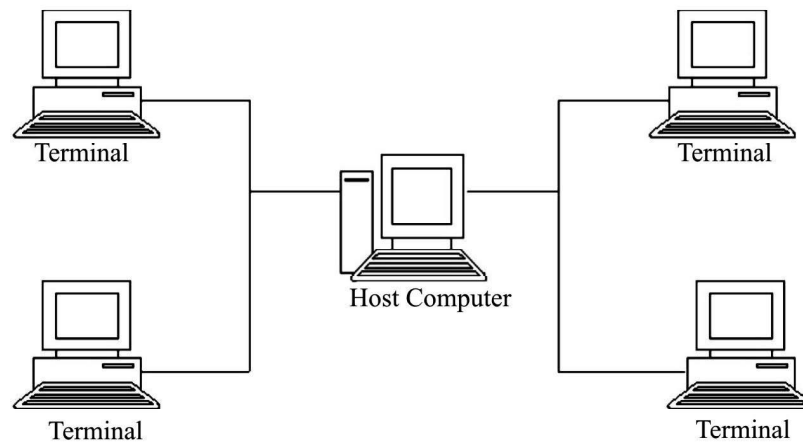


Fig. 3.3 Multi-User Environment

The host computer, which has a UNIX operating system, provides services to the terminals, such as file access services. Four terminals are connected to one host computer and all the terminals are sharing resources from the host computer.

Multitasking feature performs scheduling of work while one task is waiting for input and another task reads input from the hard disk. As a result, you can perform more than one task simultaneously.

3.3.1 Architecture of UNIX

An operating system includes a user interface used for interacting with user and the command interpreter that helps in converting the user language in a high-level language understood by the computer. In UNIX, the command interpreter is not integrated with operating system but as a separate program. The UNIX architecture is divided into different layers such as:

- **Hardware:** Provides the basic hardware interface modules such as memory management, process scheduling, interrupt handling, low-level device control, and process synchronization. These modules help in managing and scheduling the processes received by the user and send these processes to the kernel, the next layer of the UNIX architecture.
- **Kernel:** Controls the resources of the system and forms an interface between shell and hardware. You can interact with the kernel through system calls.

System call is a command given by the end user. There is only one kernel in a system. Programs, such as shell and vi editor interact with the kernel by invoking a well-defined set of system calls. The system calls enables the kernel to perform various operations for the calling programs and it also exchanges the data between the kernel and the program. System calls is also used for file manipulation, process control, information manipulation, and device manipulation.

NOTES

- **Shell:** Receives a command as an input and sends the command to the kernel for processing. Shell is a program that enables the end user to request the kernel to do the processing of a given command. There can be multiple shells in action in a system, one for each end user who is logged in.
- **Other program utilities:** Includes various end user applications such as desktop application, network application, and database application.

Figure 3.4 shows the UNIX system architecture.

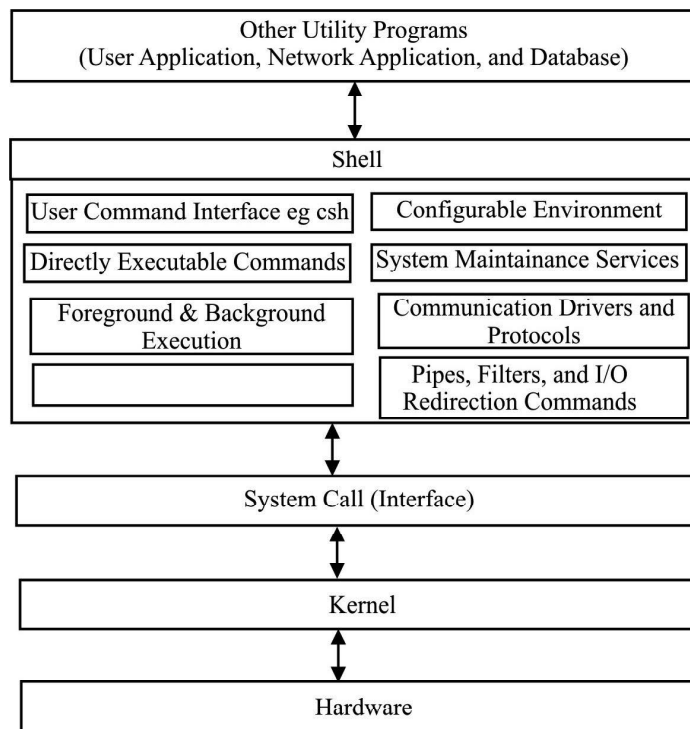


Fig. 3.4 The UNIX System Architecture

3.3.2 Features of UNIX

UNIX as an operating system that supports all features that are expected in an operating system. UNIX also provides additional features that are not supported in any other operating system. The general and additional features of UNIX operating system are:

- **File and process:** File and process are two entities that are supported by UNIX. A file contains information such as text, code, or directory structure that you need to save in the computer. The file is stored in the hard disk of

NOTES

the computer at a particular location, which can be easily remembered whereas a process is the name given to a file or a program that is currently running. UNIX provides various tools that enable you to control a process, change the sequence of the process and kill the process.

- **Multi-user system:** UNIX supports multitasking system as the kernel is designed to handle multiple processes. A single user can run multiple processes simultaneously. For example, an end user can print a file and edit another file simultaneously. The kernel handles the multiple processes as foreground and background processes. The current process runs in the foreground and the other processes run in the background. The foreground and background processes can be switched according to the requirements. This multitasking feature is an advantage for the programmers, as they do not have to close the editor and run the program; this can be done simultaneously.
- **UNIX toolkit:** The UNIX toolkit provides various tools that enable you to perform different tasks in UNIX as the kernel alone cannot perform every task. The tools that are included in the UNIX toolkit are:
 - o General purpose tools such as vi editor
 - o Text manipulation utilizes filters that are used to retrieve the output from two or more commands simultaneously
 - o Compiler and interpreter
 - o Network administration and system tools such as mailx and pine
- **Pattern matching:** UNIX supports pattern matching feature that enables you to retrieve the output according to the required pattern. Pattern matching in UNIX can be implemented using a special characters such as * known as metacharacter. The * character denotes multiple characters. For example, if you need to retrieve all the files whose name start with A, you can give the argument for the list command as A*, instead of specifying all the names of the files.
- **Programming facility:** UNIX provides a programming facility known as shell that is developed specifically for programmers and not for the users. The shell programming includes all programming features such as variables, loops and control structures that help you to create a shell script.

Logging on with UNIX

The UNIX operating system provides security for the UNIX users. As multiple users can use the UNIX computer, each user is provided with his/her own login name and password to work in the UNIX computer. The system administrator is the person who is responsible for creating a login name and password for the end-users. To enter the login name and password, UNIX provides a login prompt, which is:

```
FreeBSD/i386 (bsd.belmaksit.com) (tty0)
login:
```

The above login prompt is preceded with the UNIX version installed. You need to enter the login name in front of the login: prompt. After entering the login

name press Enter from the keyboard, which prompts you to enter the password that was created by the system administrator. Password is the secret code that should not be known to others. For example, if you need to login with the login name, Steve, enter the string at the prompt and press Enter button. Pressing the Enter button prompts you to enter the password. The password string will not be visible on the screen. The login screen to login with username Steve is:

```
FreeBSD/i386 (bsd.belmaksit.com) (ttyp0)
login: Steve
password:
```

Enter the password string and press Enter button to login. If the login and password strings are entered correctly then the UNIX provides you with the welcome screen along with the last login time information. The welcome screen is:

```
Last login: Sun Oct 19 11:30:59 from 192.168.0.123
Copyright (c) 1980, 1983, 1986, 1988, 1990, 1991,
1993, 1994
The Regents of the University of California.
All rights reserved.
FreeBSD 4.8-RELEASE (GENERIC) #0: Thu Apr 3
10:53:38 GMT 2003
Welcome to FreeBSD!
$
```

After logging on with UNIX, the system prompts you with the dollar (\$) prompt with the blinking cursor.

3.3.3 Working with Commands

The UNIX system is based on commands, which means that each action in UNIX is performed using commands. UNIX uses command line interface (CLI) to run the commands. The CLI interface is not supported by Microsoft Windows operating system as each task in Windows is performed by clicking a button. The UNIX commands are designed to ensure that maximum amount of output is retrieved from the typed command.

Command Structure

Before using the commands, you must be aware of the command structure that implies how to use the commands. The UNIX command structure is divided into two parts: the command name and the argument used along with the command. The command and argument are separated using a space or a tab to enable the command interpreter to interpret these commands and arguments as a word. The arguments used along with the commands are categorized as:

- Options
- Expressions
- Filenames
- Instructions

NOTES

NOTES

The syntax for using any command is:

```
Command [options|expression|filenames|instructions]
```

The options argument is the special argument as the options used with a particular command are predefined. The option argument uses a minus sign (-) in front of the option to distinguish it from the filename argument. An example of the use of the option argument is:

```
ls -l
```

where `ls` is the name of the command and `-l` is the option used with the `ls` command. The `-` sign is used with the option argument. Multiple options in a command are used with a single minus sign. An example of the use of multiple minus sign is:

```
ls -l -a -t
```

where `-l`, `-a`, and `-t` are options used with the `ls` command. This command can also be used as:

```
ls -lat
```

The filename argument is used with a command to retrieve the input from a particular file. The filename argument must be at the end of the command after specifying all the options. An example of the use of the filename argument is:

```
ls -lat sample
```

where `-lat` are the options of the `ls` command and `sample` is the name of the file. The expressions argument is not predefined; you can use any word as an expression along with command. An example of the use of the expression argument is:

```
cat hello
```

where `cat` is a command that is used to display the arguments specified in front of it and `hello` is the expression used with the `cat` command.

UNIX provides certain commands that do not require arguments. For example, the `pwd` command that prints the name of the current working directory does not accept any argument.

UNIX also provides certain commands that may or may not require arguments, which means that means providing arguments with these commands is optional. For example, `ls` is a command that is used with and without arguments, according to the requirement.

In UNIX, certain characters such as pipe sign (`|`), `<`, and `>` operators when used with the commands are not treated as arguments. These characters have some special meaning in UNIX. For example, pipe sign is used when you need to run multiple commands simultaneously.

Working with Keyboard Commands

UNIX provides various keyboard commands that help you to interact with the UNIX system. Table 3.6 lists the various types of keyboard commands.

Table 3.6 Keyboard Commands

Commands	Function
Ctrl+h	Enables you to delete a character.
Ctrl+c or Delete	Interrupts the current command.
Ctrl+d	Terminates the current login session or a command that requires an input from the keyboard.
Ctrl+s	Stops the scrolling of the screen output and locks the keyboard.
Ctrl_q	Unlocks the keyboard and resumes the scrolling of the screen output.
Ctrl+u	Terminates command line without running the command.
Ctrl+\	Terminates the running command but creates a core file containing the memory image of the program.
Ctrl+z	Returns the shell prompt and suspends the process.
Ctrl+j	Alternative command for Enter.
Sty sane	Restores terminal to normal status.

NOTES

Using the Simple Commands

The UNIX commands are programs that are stored under the /bin directory. When a command is executed, UNIX searches that command in the /bin directory for which you need to specify the path of the /bin directory in the PATH global variable. These commands are known as external commands. An example of external command is the ls command, which is stored under the /bin directory. UNIX also provides commands that are not stored anywhere in the directory and run independently. These commands are known as internal commands. An example of internal command is the echo command, which is not been searched under any directory. The simple commands that help you to start working in UNIX are:

- ls
- echo
- type

The ls Command

The ls command is used to display all the file and directories of the current directory. An example of the use of the ls command is:

```
$ ls
archive.tar.gz      guide2.html      guide4
gimp-1.2.5          guide1           hello
gimp-1.2.5.tar.gz  guide2           jack
gude1.html          guide3
```


The `ls` command has several command options that display additional file and directory information. Table 3.7 describes the options of `ls` commands.

Table 3.7 The `ls` Command

NOTES

Options	Description
<code>-x</code>	Displays the list in a multicolumn format.
<code>-F</code>	Displays symbols along with files and directories that helps you identifying the type of file. For example, <code>*</code> symbol for the executable files, <code>/</code> symbol for the directories, and <code>@</code> symbol for the symbolic links.
<code>-a</code>	Displays filename that starts with a dot (<code>.</code>) including <code>.</code> and <code>..</code> .
<code>-R</code>	Displays the recursive list, which means filenames of the current directories and sub directories.
<code>-r</code>	Displays the filenames in the reverse order, that is in descending order.
<code>-l</code>	Displays one filename in a separate line.
<code>-l</code>	Displays list of files along with its attributes such as file permissions, type, size, and filename.
<code>-d dirname</code>	Displays the directory name specified along with this option.
<code>-t</code>	Displays the sorted filenames according to the last modification time.
<code>-lt</code>	Displays the list according to the last modification time.
<code>-u</code>	Displays the sorted filenames according to the last modification time.
<code>-I</code>	Displays inode number.

The command to use the `ls` command with `-l` option is:

```
$ ls -l
```

The output of the above command is:

```
archive.tar.gz
gimp-1.2.5
gimp-1.2.5.tar.gz
gude1.html
gude2.html
guide1
guide2
guide3
guide4
hello
jack
$
```

The command to use the `ls` command with `-l` option is:

```
$ ls -l
```

The output of this command is:

```
total 14714
-rw-r-r- 1 jack jack 169 Oct 11 15:12 archive.tar.gz
drwxr-xr-x 20 1000 1000 2048 Oct 11 15:08 gimp-1.2.5
-rw-r-r- 1 jack jack 142 Oct 11 13:40 gude1.html
-rw-r-r- 1 jack jack 150 Oct 11 13:40 gude2.html
-rw-r-r- 1 jack jack 0 Oct 11 15:10 guide1
-rw-r-r- 1 jack jack 0 Oct 11 15:10 guide2
-rw-r-r- 1 jack jack 0 Oct 11 15:10 guide3
-rw-r-r- 1 jack jack 0 Oct 11 15:11 guide4
drwxr-xr-x 2 jack jack 512 Oct 11 13:35 hello
-rw-r-r- 1 jack jack 0 Oct 11 12:41 jack
$
```

NOTES

The preceding commands display the list of filenames along with their attributes. Table 3.8 shows the attributes of a long list of commands.

Table 3.8 *Attributes of a Long List of Commands*

Options	Description
<code>-x</code>	Displays the list in a multicolumn format.
<code>-F</code>	Displays symbols along with files and directories that helps you identifying the type of file. For example, <code>*</code> symbol for the executable files, <code>/</code> symbol for the directories, and <code>@</code> symbol for the symbolic links.
<code>-a</code>	Displays filename that starts with a dot (<code>.</code>) including <code>.</code> and <code>..</code>
<code>-R</code>	Displays the recursive list, which means filenames of the current directories and sub directories.
<code>-r</code>	Displays the filenames in the reverse order, that is in descending order.
<code>-l</code>	Displays one filename in a separate line.
<code>-l</code>	Displays list of files along with its attributes such as file permissions, type, size, and filename.
<code>-d dirname</code>	Displays the directory name specified along with this option.
<code>-t</code>	Displays the sorted filenames according to the last modification time.
<code>-lt</code>	Displays the list according to the last modification time.
<code>-u</code>	Displays the sorted filenames according to the last modification time.
<code>-I</code>	Displays inode number.

The command to use the `-F` option that displays the difference between the executable files and directories is:

```
$ ls -F
```

NOTES

The output of this command is:

```
File.d      file5      guide1*
file.d      final      guide4*
file.inst   fun        hello/
file.ph     funny     jack
file1       gimp-1.2.5/  salary
file2       gimp-1.2.5.tar.gz  table
```

In this command, the appended slash (/) character denotes that it is a directory and the appended asterisk (*) character denotes that it is an executable file.

The `ls` command does not list the directories or files that start with the character dot (.). The `-a` option is used to list the files and directories that starts with the dot (.) character. The command to use the `-a` option with `ls` command is:

```
$ ls -a
```

The output of this command is:

```
.file.d      gimp-1.2.5
..file.inst   gimp-1.2.5.tar.gz
.Xauthority  file.ph      gudel.html
.cshrc       .login      .login_conf
.mail_aliases .mailrc     .profile
.rhosts      hello       jack
.shrc        File.d      final
File.new     fun         table
archive.tar.gz
$
```

You can use multiple options in a single command. For example, if you want to list the files and folders that start with dot (.) character along with the long listing, then the command is:

```
$ ls -al
```

```
drwxr-xr-x  4 jack      jack  1024 Oct 17 16:10 .
drwxr-xr-x  4 root      wheel  512 Oct 11 18:00 ..
-rw-----  1 jack     jack    0 Oct 11 15:01 .Xauthority
-rw-r--r--  1 jack     jack   771 Oct 11 12:40 .cshrc
-rw-r--r--  1 jack     jack   255 Oct 11 12:40 .login
-rw-r--r--  1 jack     jack   165 Oct 11 12:40 .login_conf
-rw-----  1 jack     jack   371  Oct 11 12:40
.mail_aliases
-rw-r--r--  1 jack     jack   331 Oct 11 12:40 .mailrc
-rw-r--r--  1 jack     jack   801 Oct 11 12:40 .profile
-rw-----  1 jack     jack   276 Oct 11 12:40 .rhosts
-rw-r--r--  1 jack     jack   852 Oct 11 12:40 .shrc
-rw-r--r--  1 jack     jack    0 Oct 16 11:48 File.d
```

```
-rw-r--r- 1 jack  jack  90 Oct 15 17:38 File.new
-rw-r--r- 1 jack  jack  169  Oct  11  15:12
archive.tar.gz
-rw-r--r- 1 jack  jack  222 Oct 16 11:50 file.d
-rw-r--r- 1 jack  jack  10 Oct 15 16:54 file.inst
-rw-r--r- 1 jack  jack  85 Oct 16 17:17 file.ph
$
```

NOTES

The echo Command

The echo command displays the arguments specified along with the command on the standard output. A blank line is displayed if the echo command is executed without any arguments. An example of the use of the echo command is:

```
$ echo "First Message"
First Message
$
```

In this command, the 'First Message' argument passed with the echo command is displayed on the screen.

The value of a particular variable can also be displayed using the echo command. Variable is the name given to a memory space to store values. To display the value of the variable, you need to provide dollar (\$) character in front of the variable name. The command to display the value of a variable is:

```
$ var=10
$ echo $var
10
$
```

In this example, var is name of the variable and 10 is the value assigned to the var variable.

The type Command

The commands in UNIX are stored under a directory structure. You can display the path of a particular command using the type command. An example of the use of the type command is:

```
$ type less
less is /usr/bin/less
$
```

This example specifies that the less command is stored under the /usr/bin directory.

Getting Help for the Commands

UNIX provides various commands that enable you to retrieve online help on the commands that are used in UNIX. The commands that are used for retrieving online help are:

- Man
- Whatis
- apropos

NOTES

The man Command

UNIX provides the `man` command, which helps in accessing the online documentation or manual pages of a particular command. The `man` command is used to retrieve the information on a particular command such as the process to run the command, the arguments used with the command and the explanation of each argument of the command.

The `man` page of a command includes various compulsory and optional sections. Every command does not have the same sections. The compulsory sections that are included in the `man` page are:

- **Name:** Provides a one-line introduction to the command.
- **Synopsis:** Displays the syntax of a command. If the command argument is enclosed with the rectangular brackets (`[]`), it means that the argument is optional. The ellipsis that means three dots represents that there can be more instances of a preceding word. For example, `[file...]` argument represents that more than one filename as an argument can be used with the command.
- **Description:** Provides a detail description of a command.

For example, the command to read the `man` page of the `echo` command that displays arguments on the standard output is:

```
$ man echo
ECHO(1) General Commands Manual      ECHO(1)
NAME
echo - write arguments to the standard output
SYNOPSIS
    echo [-n] [string ...]
```

The `echo` utility writes any specified operands, separated by single blank (` `) characters and followed by a newline (` \n `) character, to the standard output.

`-n` Do not print the trailing newline character. This may also be achieved by appending ` \c ` to the end of the string, as is done by iBCS2 compatible systems. Some shells may provide a builtin `echo` command, which is similar or identical to this utility.

Diagnostics

The `echo` utility exits 0 on success, and `>0` if an error occurs.

Standards

The `echo` utility is expected to be IEEE Std 1003.2 (‘POSIX.2’) compatible.

The commands in UNIX are divided into various sections. When a command is specified as an argument with the `man` command, the `man` command starts searching the specified command. There is a possibility that a particular command can be included in two sections, for example, the `password` command. In such cases, you need to specify the section number as an argument along with the name

of the command because the `man` command stops the search even if the command is included in the other section. Table 3.9 lists the sections in which the UNIX commands are divided.

Table 3.9 UNIX Commands

Section	Description
1	User programs
2	Kernel's system call
3	Library functions
4	Administrative file formats
5	Miscellaneous
6	Games
7	Special files (in/dev)
8	Administration commands

NOTES

The `whatis` Command

The `whatis` command provides one-liners for a command. The `whatis` command does not search man pages; it searches database that is separate from the man pages. An example to use the `whatis` command is as follows:

```
$ whatis man
catman(1)          - preformat man pages
groff_man(7)       - groff 'man' macros to support
generation of man pages
man(1)            - format and display the on-
line manual pages
manpath(1)        - determine user's search path
for man pages
pod2man(1)        - translate embedded Perl pod
directives into man pages
```

The `apropos` Command

The `apropos` command is used to search the commands and files related to the passed argument. This command does not give detailed information but provides you with information that the particular keyword is a file or a command. An example of the use of the `apropos` command is:

```
$ apropos http
```

The output of this command is:

```
accf_http(9)      - buffer incoming connections
until a certain complete
```

NOTES

```
HTTP requests arrive
fetchMakeURL(3), fetchParseURL(3),
fetchFreeURL(3), fetchXGetURL(3), fetchGetURL(3),
fetchPutURL(3), fetchStatURL(3),
fetchListURL(3), fetchXGet(3), fetchGet(3),
fetchPut(3), fetchStat(3), fetchList(3),
fetchXGetFile(3), fetchGetFile(3),
fetchPutFile(3), fetchStatFile(3),
fetchListFile(3), fetchXGetHTTP(3), fetchGetHTTP(3),
fetchPutHTTP(3), fetchStatHTTP(3),
fetchListHTTP(3), fetchXGetFTP(3),
fetchGetFTP(3), fetchPutFTP(3), fetchStatFTP(3),
fetchListFTP(3) file transfer functions
CGI::Carp(3)      CGI routines for writing to the
HTTPD (or other) error log
$
```

The man command with `-k` argument is the alternative of the `apropos` command. You can use either of the two commands to search the specified keyword. An example of the used of the `man -k` command is:

```
$ man -k awk
```

The output of this command is:

```
a2p(1)          - Awk to Perl translator
awk(1)          - pattern-directed scanning and
processing language
gawk(1)         - pattern scanning and processing
language
English(3)     - use nice English (or awk) names for
ugly punctuation variables
$
```

Introducing Files in UNIX System

The operating system provides the filesystem to organize a data by grouping of files on the disk. The UNIX file system has a hierarchical structure of directories and stores the files in these directories. The filesystem begins with the directory called root. The root directory is denoted by `/` and it contains a file called UNIX, which is the UNIX kernel file. The sub-directories that are present in the root directory are `bin`, `lib`, `usr`, etc, `tmp`, and `dev`. Table 3.10 describes the sub-directories present in the root directory.

Table 3.10 Built-in Directories of UNIX

Directory	Contains
Bin	The bin directory contains executable files for most of the UNIX command.
Lib	The lib directory contains all the library functions provided by UNIX. The programs written in UNIX make use of the library function in lib directory.
Dev	The dev directory contains files that control various input/output devices, like terminals, disk drives, and printer. Each device has a separate file in the UNIX.
Etc	Binary executable files are required for system administration.
Tmp	It contains temporary files created by UNIX users.
Usr	Home directory of all users.
/usr/bin	Additional binary executable file.

NOTES

The benefits of the filesystem are as follows:

- The UNIX file system has a hierarchical file structure. Files can grow automatically.
- Files have access permissions.
- All devices are implemented as files.

There are various types of files in UNIX, which are:

- **Ordinary files:** The files created by the end user come under this category of files, such as all the data files, program files, object files and executable files. An end user can make changes to such files.
- **Directory files:** These files are created automatically when the directory exists. This file has the same name as the directory name and it contains information about the files under the directory. For example: /home/tom, that means tom is a directory under /home, which contains information on all the files and directories under the directory, tom.
- **Special files:** Special files are associated with input/output devices. That is found in standard UNIX directories, such as/dev and/etc. The end-user has no right to alter this file.

NOTES

Figure 3.5 shows UNIX File types.

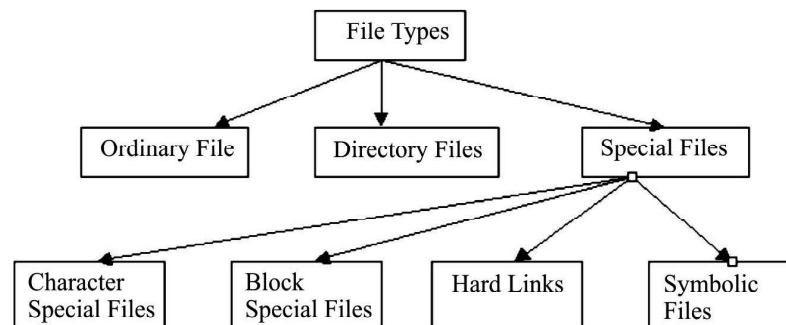


Fig. 3.5 UNIX File Types

UNIX supports different types of special files:

- **Character special files:** Reads and writes data one character at a time, such as modem.
- **Block special files:** Accesses a block of data at a time. A block of data can be either 512 or 1024 bytes. The kernel collects the data in a memory for block memory and makes it available for the end-user. This makes input/output operation faster; hard disk is an example of block device.
- **Hard links:** Allows a single file to have multiple names. Hard links are only for a file, not for a directory. They create a direct link to an inode. Each file in a filesystem has an inode which it identifies by an inode number. You can do hard link to only that file which is in the same file-system.
- **Symbolic links:** Are similar to hard links except that symbolic links work across different filesystems.

Filename

A filename allows you to access a file for doing different operations, such as reading, writing and running a file. The characteristics for naming a file in UNIX operating system are:

- File name should be up to 255 characters long. This restriction is only on file name, not on the path name.
- File name can contain any character except ('/') forward slash.
- File name in UNIX is case-sensitive.
- File name should not have components and extension. You can use as many dots as you need, up to 255 characters.
- File name should not have a blank or a tab.

Working with Wildcards

The shell facilitates you to perform an operation on a set of files without having to specify all the names of the files on which the operation is to be performed. This is

possible with certain special characters called wildcard characters. The shell interprets wildcard characters as a specific pattern of characters. Wildcards are used to enable several files with related names to be manipulated by a single command. On a UNIX system, the shell is responsible for examining the end user input and converting the wildcard expression to a list of file names, which are then passed on to a selected command. Table 3.11 describes the wildcards understood by UNIX shells.

NOTES

Table 3.11 Wildcard Characters

Wildcards	Description
*	Matches any sequence of character, matches none, one character or a string of more than one character.
?	Matches any single character.
[]	Matches exactly one of a specified set of characters.
[!..]	Matches any character not included in the list.

The command to use the wildcard character , * is:

```
$ ls file*
```

This command displays all the file names starting with file and followed by none or any number of characters. The output of this command is:

```
file.d      file1      file4      file7
file.inst   file2      file5      file8
file.ph     file3      file6
$
```

The command to use the ? wildcard is:

```
$ ls file*.*
```

This command displays all the filenames starting with file, followed by none or any number of characters with any single character extension. The output of this command is:

```
file.d      file1.c
file.c      file2.d
$
```

The command to use the wildcard [] is:

```
$ ls file[1 - 3]
```

This command displays the filename starting with file and ending with 1, 2, or 3. The output of this command is:

```
file1
file2
file3
$
```

The command to use the wildcard, [!...] is:

```
$ ls file[!1 - 3]
```

This command displays the filename starting with file and ending with any character other than 1, 2, or 3. The output of this command is:

```
file4  
file5  
file6  
file7file8  
$
```

NOTES

UNIX Kernel

UNIX kernel is a software, which controls resources to execute a program by allocating hardware and software resources to the program. The various examples of hardware and software resources are allocating CPU time, memory, access rights for reading and writing to the disk drives and access the network.

Kernel is a collection of routines written in C language that are loaded into the memory when the system is booted. It controls the resources of the system and forms an interface between the shell and the hardware. You can interact with the kernel through system calls. There is only one kernel in a system. The kernel provides the file system, CPU scheduling, memory management, and other operating system functions through system calls. Programs, such as shell and vi editor interact with the kernel by invoking a well-defined set of system calls. The system calls enables the kernel to perform various operations for the calling programs and it also exchange the data between the kernel and the program. System calls are also used for file manipulation process control, information manipulation, and device manipulation. Figure 3.6 shows the kernel of the UNIX operating subsystem.

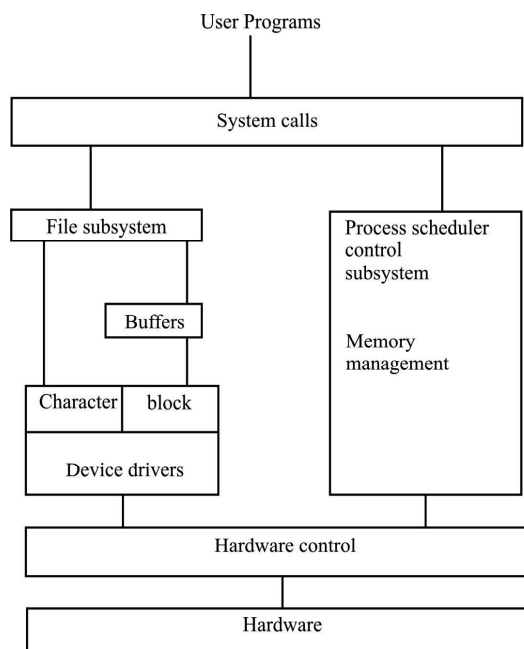


Fig. 3.6 Kernel of UNIX Operating System

UNIX kernel is divided into four main subsystems which are:

- **Process management:** It defines composition of processes, methods used by the system to switch and share the resources between processes at the time of execution. A process can be defined as a program which is under execution. In order to execute, a process requires certain resources such as CPU, I/O devices and memory.
- **Memory management:** It manages several processes in memory at the same time. Memory management helps improve the utilization of the CPU and speed of its response to execute a process. Memory management in UNIX is accomplished through paging and swapping.
- **I/O management:** It manages and controls I/O operations and I/O devices. I/O devices such as monitor and processor are used to perform I/O operations in a computer. I/O management controls device allocation, buffering, spooling, caching, file-system space allocation, error handling, failure recovery, access control to files and I/O scheduling.
- **File management:** File management defines creation and deletion of files, manipulation of files, how to get a backup of the file on stable storage media and how files can be protected from external modification. It also defines sharing of files among multiple processes, users and computers and protection of file.

NOTES

Types of UNIX Kernels

The kernel acts as a mediator between the programs and the hardware. It uses the memory management for all running programs and ensures that they all get a fair share of the processor's cycle. UNIX can be configured on two types of kernels, modular and monolithic.

Modular kernels include the code for memory management support and module support. The rest of the code is compiled as modules and loaded at runtime to handle the communication between kernel and hardware. When a device needs to be used, you can instruct the kernel to load that code into the memory; once the code is loaded, the hardware can be accessed and used.

Reduced memory overhead and flexibility are the main advantages of modular kernels. All the drivers are compiled as modules in starting, so it is not necessary to recompile the kernel every time a new hardware is installed on your system. The loading time of the driver is a disadvantage.

Monolithic kernels include all the necessary drivers for each hardware device that will be installed on the computer. If a new hardware is installed in the future, you will have to add it either at compile time or recompile the kernel. Advantages of using a monolithic kernel are:

- Reduced compile time for the entire kernel
- Making bootable kernels for diskless workstation
- Lack of initial RAM device at boot time

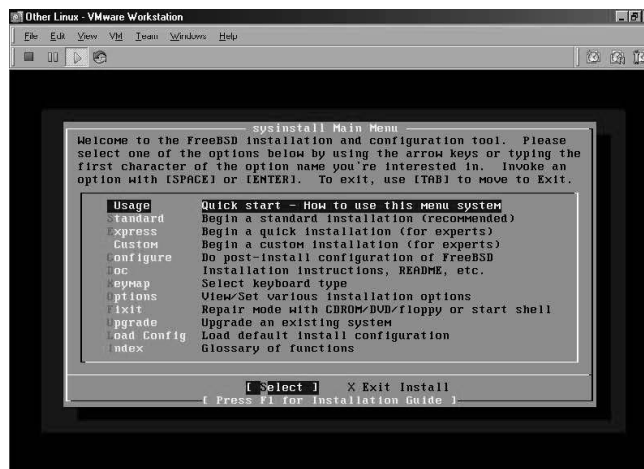
NOTES

Configuring and Compiling UNIX Kernel

Compilation of kernel is done when you install a new kernel or upgrade an existing kernel. The process of compiling the kernel is handled by the `sysinstall` utility. The `sysinstall` utility is the installation application provided by the FreeBSD version of UNIX. The `sysinstall` utility is console based and allows you to navigate using the arrow keys, Enter, Space and Tab keys. The steps to configure and compiling UNIX kernel are:

1. Execute the following command on command prompt to open `sysinstall` main menu.
`testserver# sysinstall`

The `sysinstall` Main Menu Page appears, as shown in the following screen:



2. Select a configuration that allows you to configure FreeBSD kernel. The FreeBSD Configuration Menu page appears, as shown in the following screen:

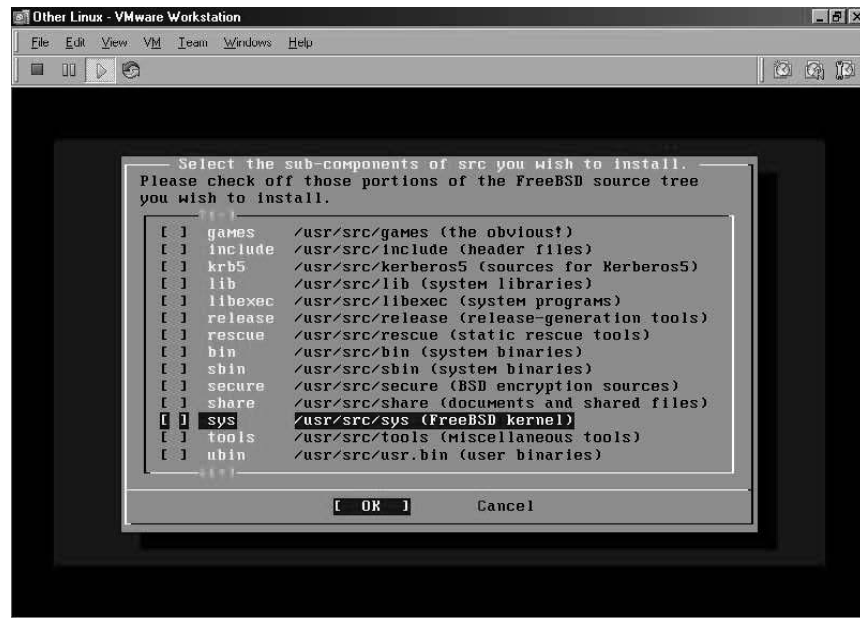


3. Select Distribution, which allows you to select dependencies required to install FreeBSD. The Select the distribution you wish to install; page appears, as shown in the following screen:



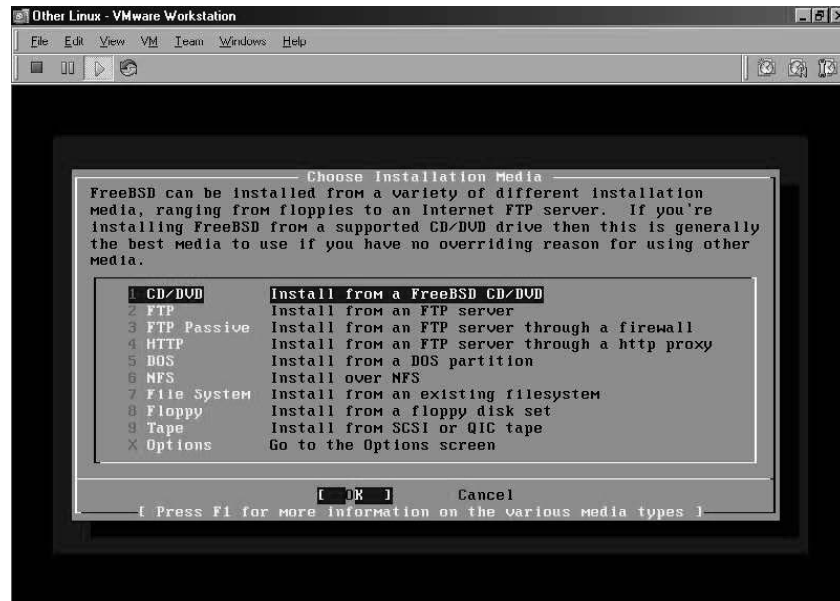
NOTES

4. Select src to open directory structures of the FreeBSD kernel sources such as kernel and system binaries. The Select the sub-components of src you wish to install page appears, as shown in the following screen:



5. Select sys to open the FreeBSD Configuration Menu.
6. Select Exit and enter OK, to open the Choose Installation Media page, as shown in the following screen:

NOTES



7. Select CD/DVD to select CD/DVD as the source for compiling the FreeBSD kernel. The source files for the kernel are copied to the UNIX machine. Go out of the `sysinstall` utility.
8. Execute the following command on command prompt to change current directory to the directory in which `GENERIC` file is stored. The `GENERIC` is a configuration file from which the kernel boots.

```
testserver# cd /usr/src/sys/i386/conf
```
9. Execute the following command on command prompt to copy `GENERIC` configuration file into `NEWKERNEL`.

```
testserver# cp GENERIC NEWKERNEL
```

Note: *Newkernel* is the user-defined name of the configuration file.
10. Execute the following command on command prompt to execute a set of Practical Extraction and Report Language (PERL) scripts to gather all the source code files and checks for kernel dependencies.

```
testserver# make depend
```
11. Execute the following command on command prompt to compile the kernel along with device drivers, ethernet, Universal Serial Bus (USB) and supported file systems.

```
testserver# make
```
12. Execute the following command on command prompt to install the new kernel.

```
testserver# make install
```

Process Management Functions and UNIX

A process is defined as an instance of a program in execution. A process is born when you start executing a program and remains active as long as the program executes. The process dies when the execution of a program is complete. The

process generally has the same name as the name of the executing program. The UNIX operating system enables you to run multiple processes simultaneously. The kernel of the operating system manages all the processes and assigns them priorities for execution. The kernel stores the attributes of the processes in process table. The attributes are also known as characteristics of the process. The two attributes of a process are:

- **Process-ID (PID):** This identifies a process and is allotted to a process by the kernel.
- **Parent PID (PPID):** This identifies the parent of a process.

Shell Process

The kernel of the UNIX operating system sets up the shell process when you log on to the operating system. The command you enter from the command prompt is treated as the shell process. The kernel of the operating system kills the shell process when you log out from the system. The shell process is assigned a PID number and a shell variable, \$\$ stores the PID number. The syntax of the command to see the PID number of the shell process is:

```
echo $$
```

An example showing the PID number of the current process is:

```
$ echo $$
265
$
```

This example shows that PID of the current process is 265. The PID of the shell process does not remain constant, it changes when you log onto the system the next time.

The process which generates another process, is known as the parent process and the process that is born, is known as child process. A process can have only one parent but it can have more than one child. When you enter any command in the command prompt, the shell process creates another process. For example, when you type the command:

```
$ cat file8
```

the shell creates the process `cat`. The shell process becomes the parent of the `cat` process and the `cat` process is known as the child process. When you type more than one command in the command prompt, the shell creates more than one process. For example:

```
$ cat file8 | grep 'Manager'
```

Here, the shell creates the process `cat` and the process `grep`. In this case, the shell process is the parent of the `cat` process and the `grep` process. When the execution of a process is completed, it dies. If a parent process dies, then its children processes automatically die.

A parent processes waits for the children process so that it can generate other processes. When you enter a command in the command prompt, a process is created and the shell process is the parent of this process of this process. The shell process waits for this child process to die. When the child process dies, the

NOTES

NOTES

shell process returns the command prompt so that it can create a new process when you enter a new command. There are some processes, which do not wait for their children processes to die and keep on generating new processes. An example of such a process is the `init` process.

Using `ps` Command

The `ps` (Process Status) is used to display the attributes of the process. The syntax of the `ps` command is:

```
ps
```

The output of the `ps` command displays the PID of the processes, the terminal with which the processes are associated, and the cumulative time consumed by the processor since the process started. An example showing the output of the `ps` command is:

```
$ ps
  PID TT  STAT      TIME COMMAND
   144 p0  S+      0:00.01 -sh (sh)
   146 p1  S        0:00.01 -sh (sh)
   194 p1  R+      0:00.00 ps
$
```

You can use `u` option with the `ps` command to display the activities of the currently logged-in user. The syntax of the `ps` command with `u` option is:

```
Ps -u <login-name>
```

An example showing the output of the `ps` command with `u` option is:

```
$ ps -u jack
  USER  PID %CPU %MEM    VSZ   RSS  TT  STAT  STARTED
  TIME COMMAND
  jack   144  0.0  0.1   636   440  p0  I+    2:34PM
  0:00.02 -sh (sh)
  jack   146  0.0  0.1   636   440  p1  S     2:34PM
  0:00.01 -sh (sh)
  jack   214  0.0  0.1   388   248  p1  R+    3:38PM
  0:00.00 ps -u
```

The `ps` command with the `a` option shows the processes of all the users. The syntax of the `ps` command with the `a` option is:

```
ps -a
```

An example showing the output of the `ps` command is:

```
$ ps -a
  PID TT  STAT      TIME COMMAND
   141 p0  Is       0:00.02 login [pam] (login)
   144 p0  I+       0:00.03 -sh (sh)
   143 p1  Is       0:00.02 login [pam] (login)
   146 p1  S        0:00.01 -sh (sh)
   291 p1  R+       0:00.00 ps -a
   132 v0  Is+      0:00.00 /usr/libexec/getty Pc ttyv0
```

```
133 v1 Is+ 0:00.00 /usr/libexec/getty Pc ttyv1
134 v2 Is+ 0:00.00 /usr/libexec/getty Pc ttyv2
135 v3 Is+ 0:00.00 /usr/libexec/getty Pc ttyv3
136 v4 Is+ 0:00.00 /usr/libexec/getty Pc ttyv4
137 v5 Is+ 0:00.00 /usr/libexec/getty Pc ttyv5
138 v6 Is+ 0:00.00 /usr/libexec/getty Pc ttyv6
139 v7 Is+ 0:00.00 /usr/libexec/getty Pc ttyv7
$
```

NOTES

System Processes

The system processes are the processes that the kernel of the UNIX operating system initiates and these processes are not associated with any user. You can use `ps` command with `e` option to display the list of all the system processes. The syntax of `ps` command with `e` option is:

```
ps -e
```

An example showing the output of the `ps` command with `e` option is:

```
$ ps -e
PID TT STAT TIME COMMAND
144 p0 I+ 0:00.03 TERM=xterm PATH=/sbin:/bin:/usr/sbin:/usr/bin:/usr/ga
146 p1 S 0:00.01 TERM=xterm PATH=/sbin:/bin:/usr/sbin:/usr/bin:/usr/ga
305 p1 R+ 0:00.00 USER=jack MAIL=/var/mail/jack HOME=/home/jack PAGER=m
$
```

Processes and Jobs

The processes in the UNIX operating system are created using various system calls. The system calls are the routines in the kernel of the operating system and responsible for the process creation. A collection of processes having common purpose is known as a job.

Creating Processes

The UNIX operating system uses three system calls for creating a process. The system calls are:

1. **Fork:** It generates a copy of the process that invokes the system call. The generated copy of the process has different PID from that of the calling process and the generated process is known as the child process.
2. **Exec:** It is used for overwriting the child process with the program that is to be executed. After overwriting the program, the new process has the same ID as that of the child process.
3. **Wait:** It is used to wait for the child process to die and when the execution of the child process is complete, the parent process notes the exit status of the child and the child process disappears from the operating system. The parent process now can create other processes.

NOTES

- The child process has a different PID from that of its parent process, but it inherits some properties from the parent process. The properties that are inherited from the parent process are:
- UID of the parent process. The UID is associated with the user.
- The user that runs the program. The UID represents the User-ID and it is assigned to the user to access the UNIX operating system.
- GID of the parent process. The GID is also associated with the user that runs the program. The GID represents Group-ID and it is the group name of the user. The system administrator allots the group name to the user when it creates the user account.
- Environment variables such as HOME and PATH.
- Descriptors of the files that a parent process opens. The descriptors are represented by integers.
- Current directory that runs the process.

Running Jobs in the Background

The UNIX OS is a multitasking operating system, which enables you to run more than one job simultaneously. When you run many jobs, one of the job runs in the foreground and other jobs run in the background. Using the operator &, you can run jobs in the background. When you use &, the shell does not wait for the child process to die. The shell returns the command prompt and allows you to enter another command. For example:

```
$ sort -o file8 file8 &  
$
```

The output of this command shows that sorting of the file8 continues in the background and the shell returns the command prompt so that you enter the new command. For example, you enter the command to see the contents of file6:

```
$cat file6
```

The output of this command is:

```
Employee      designation   address  
Sunil Roy     Manager      Delhi  
Kishore roy   Technician   Mumbai  
[1] Done      sort -o file8 file8
```

The output of this command shows that a message is displayed about the job running in the background. The message indicates that the job running in the background is completed. You can see the sorted contents of the file8 by executing the command:

```
$cat file8
```

The output of this command is:

```
Arun Mehta    Assistant    Mumbai      8000  
Arun Saxena   Assistant    Delhi        6900  
Dev Sharma    Manager      Mumbai       7000  
Kishore Roy   Technician   Delhi        5000
```

Neha Sinha	Manager	Delhi	5500
Nita Singh	Manager	Agra	9200
Ram Kishore	Manager	Agra	6000
Sohan Sharma	Manager	Mumbai	9000
Sunil Roy	Manager	Bangalore	8500

NOTES

The background execution of the job has the advantage that you can run the important jobs in the foreground while other less important jobs run in the background.

The job in the background stops running when you log out. The no hangup (nohup) command allows the job to run in the background even when you log out. The syntax of the nohup command is:

```
nohup <command> &
```

An example showing the output of the nohup command is:

```
$ nohup grep 'Manager' file8  
appending output to nohup.out
```

In this command, the sorted output is written to the file nohup. Out as output file is not specified. Using nohup command, you can log out even if job is not completed and the job continues to run in the background.

In order to run more than one command, you need to use nohup command before each command. For example:

```
$ nohup sort file8 & nohup grep 'Manager' file8 &  
$ appending output to nohup.out  
appending output to nohup.out  
[1] Done          nohup sort file8  
[2] Done          nohup grep Manager file8  
$
```

In this example, the sort and grep commands run in the background.

Executing Low Priority Jobs

The UNIX operating system manages the utilization of CPU by assigning priorities to various processes. You can reduce the priority of the job using nice command. The lower the nice value of a process, higher is the priority of the process. By default, the nice value of a process is increased by 10 units, which means that priority of the process is reduced by 10 units. You can also specify a value to reduce the priority of a process. For example, the command to reduce the priority of the date command by 5 units is:

```
$ nice -5 date &
```

The user cannot increment the priority of a process. The priority of a process can be incremented only by the super user where the super user refers to the system administrator.

Using Signals to Kill the Processes

The UNIX system communicates with the running processes using various signals. The signals are represented by certain numbers. The signals are also represented

by names having the prefix SIG. You can terminate a process running for a very long time by pressing the interrupt key. When you press the interrupt key, the signal SIGINT is sent to the process.

NOTES

You can also terminate or kill a process using kill command. The kill command takes PID number of the process to be killed as its argument. The syntax of the kill command is:

```
kill <argument>
```

An example of the kill command is:

```
$ kill 130
```

This command will kill the process having 130 as PID number.

You can kill two or more processes by specifying their PIDs in the same statement. For example:

```
$ PID 130 141
```

The command to list all the signals on your UNIX system is:

```
$ kill -l
```

```
hup int quit ill trap abrt emt fpe kill bus segv  
sys pipe alrm term urg
```

```
stop tstp cont chld ttin ttou io xcpu xfsz vtalm  
prof winch info usr1 usr2
```

```
$
```

The output of this command shows the names of all the signals available on your UNIX system.

The command to kill the shell is:

```
$kill -hup PID
```

In this command, PID is the PID number of the shell.

Job Control

You can create a job by running a sequence of commands simultaneously. The job control involves manipulation of jobs. The job control activities involve:

- Sending a job to the background
- Transferring a job from background to the foreground
- Finding the jobs that are currently active
- Suspending the job running in the foreground
- Killing the job

If you invoke a command in the command prompt and the shell does not return the command prompt immediately, you can suspend the job by pressing Ctrl-Z key. For example:

```
$ cat> file10
```

```
Employee Designation Salary
```

```
^Z[1] + Suspended cat >file10
```

```
$
```

This example shows that the shell displays the message [1] + suspended which indicates that the current job is suspended.

You can send the currently suspended job to the background and run another job in the foreground. The command to send the currently suspended job in the background is:

```
$ bg
```

After pushing the currently running job in the background, you can start a new job in the foreground or background. For example:

```
$ sort -o file8 file8 &  
$ sort -o file2 file2 &  
[5] Done sort -o file8 file8
```

This example shows the commands for running the jobs of sorting file 8 and file 2 in background.

You can display the status of various jobs using the following command:

```
$ jobs
```

The output of this command is:

```
$ jobs  
[1] Stopped (tty input) cat >file10  
[2] Stopped (tty input) cat >file11  
[3] - Stopped (tty output) man bg  
[4] + Stopped (tty input) cat >file10  
[6] Done sort -o file2 file2
```

You can bring a job in the background to the foreground using fg command. The fg command used with job number brings the specified job in the foreground. When you use fg without any argument, it brings the most recent job in the foreground. For example:

```
$ fg  
cat >file10
```

You can kill a suspended job in the background using kill command. An example showing the output of the kill command is:

```
$ cat > hello  
[Ctrl-Z]  
suspended  
$bg  
[1] cat > hello &  
[1] + suspended (tty input) cat > hello  
$ kill %1  
[1] Terminated  
$
```

This example shows that the suspended job [1] in the background is terminated using kill command.

NOTES

NOTES

Job Scheduling

You can schedule the execution of various jobs using `at` and `batch` commands. Using `at` command, you can specify a particular time for the execution of a job. An example showing `at` command is:

```
$ at 16:05
at > hello.sh
```

The output of this command is:

```
Job 1 will be executed using /bin/sh
```

In this example command, the job 1, which is `hello.sh`, will be executed at 16:05.

Using `batch` command, you can schedule a job to execute later. With this command, you do not specify the time for the execution of a job and the job is executed whenever the system load is less or system is free. An example showing the output of the `batch` command is:

```
$ batch < hello.sh
Job 1 will be executed using /bin/sh.
```

You can determine the time taken by a particular operation using `time` command. The output of the `time` command shows real time, user time and the system time. The real time represents the time duration between the invocation and the termination of the command. The user time represents the time of the execution of the command and the system time represents the time that the kernel takes to work for the execution of the command. The syntax for using `time` command is:

```
time command
```

An example showing the output of the `time` command is:

```
$ time man echo
FreeBSD 4.8 July 22, 1993      FreeBSD 4.8
5.20 real    0.00 user    0.01 sys
```

Process Abstraction

The process abstraction treats a process as a sequence of instructions that the Central Processing Unit (CPU) executes. It describes the interaction of the process with the operating system and the hardware.

Virtual Address Space

The virtual address space of a process is defined as the set of memory locations that a process can access. The virtual address space consists of:

- **Environment variables and command line arguments:** It is the first entry in the virtual address space of a process.
- **Stack:** It is a Last In First Out (LIFO) data structure that stores all the arguments and local variables of various functions. The data stored in a stack is manipulated by `PUSH` and `POP` operations. The `PUSH` operation is used to store data and `POP` operation is used to retrieve data.

- **Heap:** It allocates memory to a program during run-time using various functions such as `malloc`, `sbrk`, and `calloc`.
- **Initialized data segments:** It includes various constants, static variables, and global variables.
- **Uninitialized data segments:** It consists of header of the program and is created at run-time.
- **Text segment:** It consists of the instructions for execution.

Figure 3.7 shows the organization of virtual address space of a process.

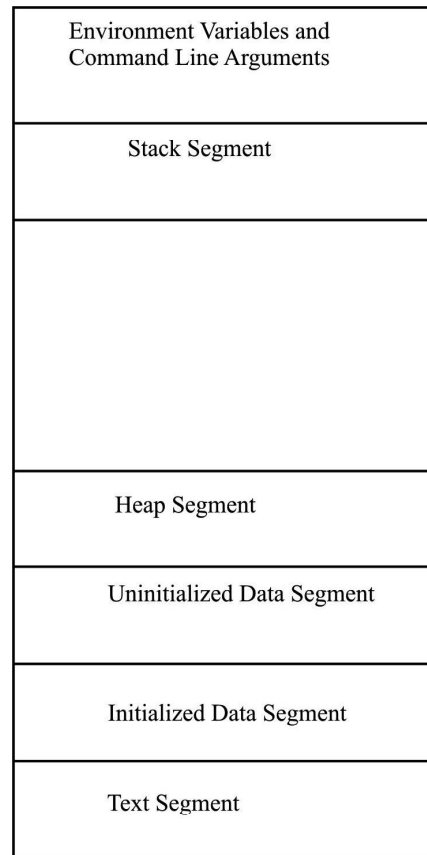


Fig. 3.7 Organization of Virtual Address Space of a Process

Process Modes

A running process can be in user mode or kernel mode. A process runs in the user mode until it does not invoke any system call. When process invokes any system call, it switches into the kernel mode. The process runs in the kernel and uses the data structures of the kernel.

Context Switching

The context of a process refers to the environment of the process during run time. The context of a process includes:

- **Hardware Context:** Includes registers such as stack pointer and program counter that store the state of the running process. The program counter

NOTES

NOTES

stores the address of the next instruction to be fetched. The stack pointer points to the address of the element on the top of the stack.

- **Software context:** Includes address space and data structures for storing the data, such as UID, PID, and the table of file descriptors.

The Process States

A process can be in one of the several states. The various states of a process are:

- **Created State:** A process is in created state when it is first created using the fork system call.
- **Ready-to-run state:** In this state, the process is ready to run after the scheduler schedules the process.
- **Running state:** A running process can be either in the kernel mode or the user mode.
- **Sleeping state:** A process in the kernel mode enters into the sleeping state when it waits for some other process to complete.
- **Zombie:** A process enters the zombie state when it dies and its address space is deallocated. The process remains in the zombie state until its parent receives the message about its state.

The `proc` and `u` area Structure

The kernel of the UNIX operating system consists of the `proc` and `u` area structure to maintain control information of various processes. The `proc` structure maintains information about a process even when the process is not in the running state. The `u` area maintains the control information of a process in the running state. The control information that `proc` structure maintains includes:

- PID of the process
- PPID of the process
- State of the process, which may be one of the various states, such as ready-to-run, running, created, sleeping, suspended and zombie
- Scheduling priority of the process, which determines which process will run next
- Pending signals mask that consists of the list of signals ending against a process

The control information that `u` area structure maintains includes:

- UID of the process
- Information regarding CPU usage
- Information regarding the current system call, which consists of the arguments and the return values
- Information regarding memory management, which consists of information size of data text and stack segments
- File descriptor table

The `u` area structure consists of only the information that the kernel requires when a process is in the running state.

Check Your Progress

1. Name the system files which are found in MS DOS system.
2. State one restriction of the volume name assigned in operating system.
3. What are the two types of DOS commands?
4. Where do internal DOS commands reside?
5. Define the term UNIX.

NOTES

3.4 WINDOWS 95/98

The first version of Windows 3.0 was released by Microsoft in 1990. It was a graphical interface based package and not a complete operating system because it required DOS to be installed first on the computer and only after that could it be loaded and used. With the launch of Windows 3.11, huge improvements in terms of usability and performance were seen because the user did not had to remember complex DOS commands which work on a single application at a time or suffer from the limited use of input devices, such as a mouse or trackball.

Windows 95

Windows 95 was a graphical user interface released by Microsoft Corporation in 1995. It had significant improvements over the earlier version of operating system distributed by Microsoft under the name of Windows 3.11. In addition to the complete change in the user interface, there were a number of important internal modifications made to the core of the operating system. Windows 95, also known as Windows version 4.0 during its development phase, was one of the most successful operating systems of that time. Windows 95 operated independently of MS DOS rather than in conjunction with it and reduced the use of MS DOS to only a boot loader for Windows 95.

Windows 98 SE

The Windows 98 SE (Second Edition) is an improved and enhanced version of Windows 98. It includes new versions of Microsoft applications as compared to Windows 98, to improve user experience and stability of the operating system. Some of the new or improved elements of this operating system are inclusion of Internet connection sharing, Windows Driver Model (WDM) for modems, wake on LAN, Internet Explorer 5.0, integrated support for DVD ROM drivers, bug free Windows, Microsoft Plus!, support for Web TV and updates for other Microsoft programs, such as NetMeeting, MicroSoft Network (MSN), Microsoft Wallet, Windows media player, etc.

Windows Millennium Edition

Windows ME was released on 14 September, 2000, targeted especially at home PC users. This OS was in continuation to Windows 98 with restricted access to

NOTES

real mode MS DOS shell to improve functionality. Among other changes, Windows ME incorporated an improved look and feel to the user interface and a system restore option of going back to a previous state of the machine. The key features of this operating system were that it had upgraded version of Microsoft products, such as Internet Explorer 5.5, Windows media player 7, system restore options, applications to easily connect with digital cameras and scanners, Windows Movie, Improved Generic support for Universal Serial Bus (USB) interface and shell extension of ZIP files into the Windows Explorer.

Windows NT

Microsoft released this version of Windows in 1993. It increased ease of use and simplified management. It used the Windows 95 interface and included advanced network support, and trouble free and better access to the Internet and corporate intranets. With the intent of designing it as an operating system capable of supporting high level language and at the same time processor independent and supporting a multiuser and multiprocessing environment, Windows NT had high acceptance in both the home user and professional user markets.

Windows 2000

Microsoft released this version of Windows in 2000. It was an upgrade from Windows NT 4.0 and was designed with the aim of replacing Windows 95, Windows 98 and Windows NT on all business desktops and laptops. This version was easy to use, Internet compatible and supported mobile computing. It made hardware installation much easier by including support for a range of new plug and play devices, including advanced networking and wireless products, infrared and USB devices. The main features of Microsoft Windows 2000 were Dump capabilities, wherein the operating system gave its users the option of dumping either a part of the memory or the entire contents into a file on the hard drive which helped in saving critical information in case of a system failure, Microsoft Management Control (MMC) to control the access to administrative tools, and system settings and Recovery Console to support the distributed file system.

Windows 2003

Windows 2003 was released by Microsoft on 24 April 2003. This OS was designed and developed over various functional parts of Windows 2000 and Windows XP. It featured better stability, compatibility and security than Windows 2000 and XP. It improved performance of the system by taking advantage of the recent hardware developments, redesigning the system interface and developing better services. Major updates of Microsoft in-house applications and services, such as networking, Web server, compatibility with Windows NT, etc., were released with this OS.

Windows XP

Windows XP was first released on 25 October 2001 and since then over 600 million copies have sold worldwide. It is a successor to both Windows 2000 and Windows ME and was the first OS aimed at home users built on the Windows

NT kernel and architecture. Due to the integration of multiple technologies from various operating systems, it gained wide popularity among home and business desktop, notebooks and media centre users. As acknowledged by most Windows XP users as well as Microsoft Corporation, this version of Windows is the most stable and efficient OS released by Microsoft yet.

Windows Vista

The most recent in the line of Microsoft Windows personal computer operating system, Windows Vista was developed to succeed Windows XP. Microsoft started the development of Windows Vista five months after releasing Windows XP and work continued till November, 2006 when Microsoft announced its completion, ending the longest development cycle of an operating system. Since the original idea of building Longhorn from Windows XP's code was scrapped, it was built on Windows 2003 SP1, with several developments including all new graphical interface named Windows Aero, refined and faster search capabilities, an array of new tools, such as Windows DVD Maker, integrated Windows Media Centre in the Vista Home Premium and Vista Ultimate Editions, print, audio, display subsystems and redesigned networking.

Windows CE

The Windows Embedded Compact (CE) is an operating system optimized for devices with minimum hardware resources, such as embedded devices and handhelds. It integrates advance and reliable real time capabilities with Windows technology. The kernel of this OS is just a trimmed down version of desktop Windows but in fact it is a brand new kernel which can run on less than a megabyte of memory. Besides the advantage of performing on a minimum specification, it is also an OS which satisfies the prerequisites of a real time operating system. Another distinct feature of Windows CE is that it was made available in a source code form to several hardware manufactures so that they could modify the OS to adjust with their hardware and also to the general public. Since Windows CE was developed as a component based and embedded operating system, it has been used as a basis in the development of several mobile operating systems, such as AutoPC, PocketPC, Windows Mobile, Smartphone, etc., and also embedded into games consoles such as Microsoft Xbox.

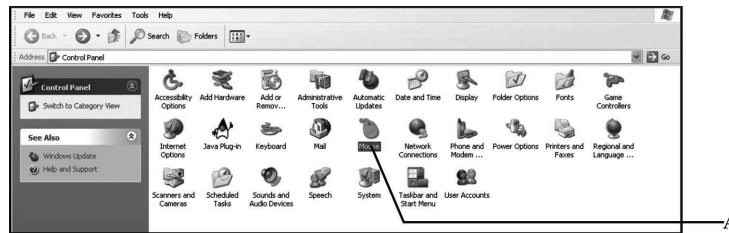
Mouse and its Events for Left and Right Buttons

A mouse is a small hand held device through which you can quickly and easily navigate different parts of the screen. You can customize various aspects of the mouse to make it work the way you prefer. You can control the mouse pointer speed, shape, button configuration, double click speed, etc. You can even swap the left and right buttons. If you frequently lose track of where your mouse pointer is, you can choose to have a mouse trail effect which will make the mouse pointer more visible. To change any of the above mentioned mouse settings:

NOTES

A Choose **Mouse** option from the **Control Panel**.

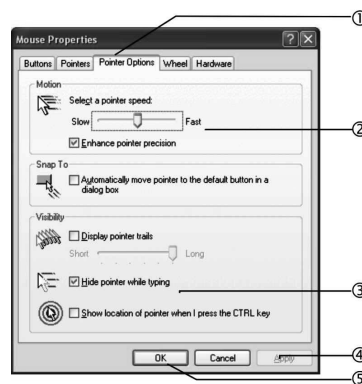
NOTES



Motion Configuration

You can also customize the motion configuration of the mouse. To do so:

- ① Click on the **Pointer Options** tab.
- ② Click on this button to increase or decrease the pointer speed.
- ③ You can also click on the **Hide pointer while typing** option to hide the pointer while you are typing.
- ④ Click on the **Apply** button to apply the changes you have made.
- ⑤ Click on the **OK** button to continue.



Click, Double Click and Drag and Drop

- **Clicking:** Move the mouse pointer to the position you want, and then press the mouse left button once.
- **Double clicking:** Move the mouse pointer to the position you want, then press the mouse left button twice in rapid succession. Alternatively, if your mouse has been so installed, you can also click the middle button once, instead of pressing the left button twice.
- **Dragging:** Move the mouse pointer to the right place, highlight the desired text by left clicking once in the beginning and releasing the left button at the end, then click once again in the highlighted block, keep the mouse left button pressed and roll the mouse to move the block to a new location by releasing the left button at the desired location.

Initially, using the mouse may be difficult (double clicking and dragging). However, learning to use it is worth the trouble as many of the advanced tools in modern programs are only operated with the mouse. In any case, most operations

are much simpler and faster to execute through the mouse rather than through the keyboard.

How Does Windows 98 Work?

Microsoft released the next version of Windows in 1998. Like its predecessor, Windows 98 supported a hybrid 16/32 bit file access system and better graphical user interface. It is often referred to as an operating system that 'Works Better, Plays Better'. Code named 'Memphis' during its development stage, Windows 98 integrated Internet Explorer (IE) into the user's desktop to allow its users get a global view of technologies over the World Wide Web (WWW) and enable easy access to it.

Some of the main features of this operating system are as follows:

- New features added to enable easy access to Internet Related Information. Introduction to Push technology which delivers news and other Internet related content directly to websites.
- Multiple displays supported using several Visual Display Units (VDUs) simultaneously to increase the capacity of the desktop and support running of different programs on separate monitors.
- Universal serial bus made computers easier to use with advanced plug and play capabilities. It allowed adding devices to your computer without having to restart each time a device is added to the computer.
- Accessibility wizard made it easier for people with disabilities to operate a computer without installing any special software.
- Help on an extensive and easy to use self help system was provided in its interface.
- In Windows 98, new interfaces, such as USB, Digital Versatile Disk (DVD) and advanced configuration and power interface were also supported.

Windows User Interface







The Windows User Interface (WUI) supports GUI. WUI standard facilitates icon oriented interface; therefore instead of typing the commands at command prompt user can avail the resources with the help of menus, icons, pictorial images, thumbnails, pointing images which can be selected by pointing the mouse on touch pad (laptop) on the screen. These standards make you able to navigate the corresponding Windows. The standard phase of GUI is a type of checklist which reminds the developers of the interface of the tasks that are needed to create usable systems. For example, setting of command buttons in the user interface is very critical task to be decided in this phase. This standard is considered to be the building code. This design ensures a consistent look and feel of GUI designing. It includes data gathering to design the GUI. It includes users, task flow environments and system requirements. This phase also includes standard screen prototype. In this phase, designer is able to flip the various menus, color, standard keys, buttons, graphic drilldown, etc. This design standard also determines the standard screen size offering 800×600 pixel settings to get the appropriate amount of information

NOTES

NOTES

on the screen, font decisions, color choices, color consistency, meaning of color decision, designing and choosing the graphics, tooltip, icons, charts or graphs, help tips, etc. Above standards reduce the development time and maintenance costs but improve usability of interface designing. Table 3.12 shows a list of buttons, actions and keyboard equivalent for command button settings.

Table 3.12 Command Button Settings for WUI

<i>Buttons</i>	<i>Action</i>	<i>Keyboard Equivalent</i>
	This button makes changes and closes the Window	ENTER Key
	This button does not make changes and closes the Window	ESC Key
	This button closes the Window when changes cannot be made or are cancelled	C
	This button is used to reset the defaults and leaves Window open	R
	This button is used to make changes and leaves Window open	A
	This button is used to open online help	H

The menus, check boxes, buttons of various kinds are all controls and refer to visual components of the GUI. They are characterized from the users' viewpoint. They also consist of various user friendly features and exhibit certain standard behaviors.

Table 3.13 summarizes the components of the Windows User Interface.

Table 3.13 Components of the Windows User Interface

<i>Components</i>	<i>Functions</i>
Windows and Messages	The Windows section describes the elements of an application with a Windows based graphical user interface.
Data Exchange	The data exchange section describes methods of exchanging data.
High DPI	This section contains information about writing Dots Per Inch (DPI) aware Win32 applications.
Resources	The resources section describes Window resources, such as icons, cursor and dialog boxes.
User Input	The user input section describes how user input is captured and used.

Desktop

The first thing that you see when you start Windows XP is the desktop. The desktop hosts icons for programs and shortcuts to programs and files.

Icon

Icons are graphical representations of the programs and files stored on the desktop.

Selecting an Icon

In order to give a command, you have to select the particular icon. Selecting an icon can be done by moving the mouse pointer over it and pressing the left mouse button once. Each time you select a particular icon, the color of the icon changes, indicating that the icon is selected.



NOTES



Moving an Icon

You can move icons around to rearrange your files on the disk or for your own convenience.

Icons can be moved on screen by simply holding the left mouse button down when you click on the icon, dragging the mouse to the desired position on the screen and then releasing the mouse button.

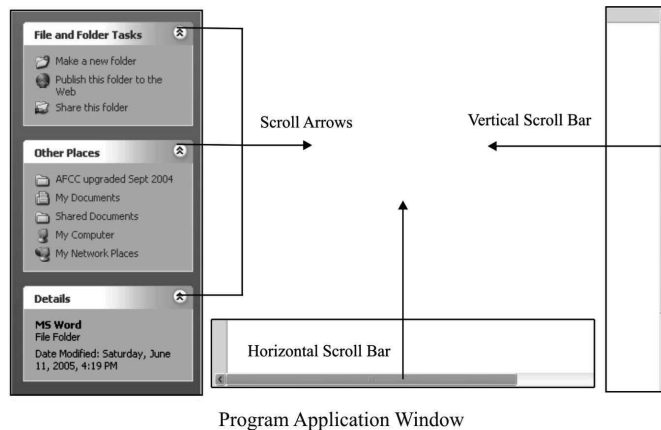
Typically, you would be simultaneously running more than one application or task. Each application will run in a separate window. Every open Window (folder or an application Window) has a separate button that appears on the Taskbar. You can make any one of these windows 'active' at a time. To do so you will need to click on the respective button on the Taskbar.



When the contents of the window do not fit in one screen, you will see boxes with arrows at the top and bottom. These boxes are called scroll bars. You can click on the arrows in the scroll bars to browse the hidden part of the screen. The blue rectangular indicators on the scroll bar tell you the position of the screen.

Scroll bars can be vertical as well as horizontal. The horizontal scroll bar is the one that moves the Window on the horizontal axis, while the vertical scroll bar moves the window on the vertical axis.

NOTES

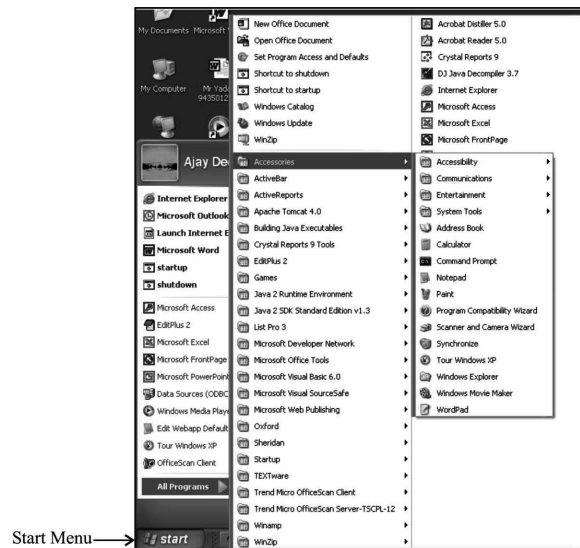


Taskbar and Title Bar

A taskbar is a bar displayed on a full edge of a GUI desktop which is used to launch and monitor running applications. The title bar contains icons at top right to the Window, such as maximize, minimize and close buttons. Menus are very useful parts of the Windows system. They consist of various commands and subcommands for default and other applications and programs stored on your computer.

Start Menu

Windows XP also has a new **Start menu** for enabling faster access to commonly used programs and common system areas, such as My Computer, Control Panel and Search.



The Start menu and the Start button are user interface elements in Microsoft Windows and serve as the central launching point for all applications. By default, the Start button is always visible in the lower left hand corner of the screen. It exhibits the Windows logo and the word 'start'.

For choosing a program from the Start menu click on the Start menu option on the Taskbar. Once the Start Menu is expanded you can drag the mouse to the right and switch over to the submenus.

Switching between Programs

When you have more than one program running, you can switch between these to make any one of them current. You can switch between two programs by simply clicking on the desired program button on the **Taskbar** as shown below. You will notice that the active program window button will be darker in color.



Alternatively, you can perform the same task by using a combination of keys on your keyboard, by pressing the **Tab** key while holding the **Alt** key. On pressing this combination you will see a box around the icons for the running programs. Each time you press the **Tab** key, the box moves to the next program. As soon as you release the **Tab** key the highlighted icon window will open. In case of the buttons shown, the window for 'My Computer' will open.

NOTES

Check Your Progress

6. State about the system call.
7. Why is the man command used?
8. Define UNIX kernel.
9. What is window 95?
10. State about the window user interface.

3.5 ANSWERS TO 'CHECK YOUR PROGRESS'

1. On MS DOS systems, the system files found are known as IO.SYS and MSDOS.SYS.
2. The volume name is case sensitive and only small letters are accepted.
3. DOS commands are of two types known as internal and external.
4. The internal commands reside in COMMAND.COM, which loads into memory when the computer system is started therefore these commands do not reside on disk.
5. UNIX is a multi-user and multitasking operating system.
6. System call is a command given by the end user.
7. The man command is used to retrieve the information on a particular command such as the process run the command, the arguments used with the command and the explanation of each argument of the command.
8. UNIX kernel is software that controls resources to execute a program by allocating hardware and software resources to the program.
9. Windows 95 was a graphical user interface released by Microsoft Corporation in 1995. It had significant improvements over the earlier version of operating system distributed by Microsoft under the name of Windows

NOTES

- 3.11. In addition to the complete change in the user interface, there were a number of important internal modifications made to the core of the operating system.
10. The Windows User Interface (WUI) supports GUI. WUI standard facilitates icon oriented interface; therefore instead of typing the commands at command prompt user can avail the resources with the help of menus, icons, pictorial images, thumbnails, pointing images which can be selected by pointing the mouse on touch pad (laptop) on the screen.

3.6 SUMMARY

- Microsoft Disk Operating System (MS DOS) is a single user, single tasking operating system. DOS has a command line, text based/non-graphical user interface commonly referred to as Character-based User Interface (CUI).
- A filename is a special type of string which is used for identifying a file stored on the file system of a computer. A filename is divided in two parts: basename (the primary filename) and the extension (which indicates the type of file related to a certain format of file).
- The system files are used to run the system unit. These files are created and maintained by the operating system and are necessary for the system to perform necessary functions. System files are used by many commands and subroutines to perform operations.
- If MS DOS system files have been corrupted during booting process then it can be recovered from MS DOS backup files. It is necessary to keep the exact version of MS DOS file.
- In MS DOS, the root directory has no name. It is simply represented by the special character that separates in a file system.
- DOS commands are of two types known as internal and external. The internal commands do not require any special files for being executed and are brought into the computer's memory as soon as the computer is switched on. External commands are commonly external either because they require large requirements or are not commonly used commands.
- In 1970, an operating system named UNiplexed Information and Computing System (UNICS) was developed, which allowed two users to work simultaneously. The name was then changed from UNICS to UNIX.
- UNIX as an operating system was developed in the early 1970s at AT&T Bell Laboratories.
- UNIX is a multi-user and multitasking operating system and can receive the commands from a number of end users to run programs, access files and print documents, simultaneously.
- In UNIX, the command interpreter is not integrated with the operating system but exists as a separate command.

- The UNIX operating system provides security to its users. Each user is provided with his/her own login name and password to work on the UNIX computer.
- The UNIX system is based on commands, which means that each action in UNIX is performed using commands. UNIX uses the command line interface (CLI) to run the commands.
- The UNIX file system has a hierarchical structure of directories and stores the files in these directories. The file system begins with a directory called root directory.
- UNIX kernel is software, which controls the resources to execute a program by allocating hardware and software resources to the program. The various examples of hardware and software resources are allocating the CPU time, memory, access write for reading and writing to the disk drives and for accessing the network.
- The first version of Windows 3.0 was released by Microsoft in 1990. It was a graphical interface based package and not a complete operating system because it required DOS to be installed first on the computer and only after that could it be loaded and used.
- With the launch of Windows 3.11, huge improvements in terms of usability and performance were seen because the user did not have to remember complex DOS commands which work on a single application at a time or suffer from the limited use of input devices, such as a mouse or trackball.
- Subsequent versions of Windows are Windows 95, Windows 98 SE, Windows Millennium Edition, Windows NT, Windows 2000, Windows XP, Windows Vista and Windows CE.
- A mouse is a small hand held device through which you can quickly and easily navigate different parts of the screen.
- Windows 98 integrated Internet Explorer (IE) into the user's desktop to allow its users get a global view of technologies over the World Wide Web (WWW) and enable easy access to it.
- The Windows User Interface supports GUI. WUI standard facilitates icon oriented interface.
- The first thing that you see when you start Windows XP is the desktop.
- Icons are graphical representations of the programs and files stored on the desktop.
- A taskbar is a bar displayed on a full edge of a GUI desktop which is used to launch and monitor running applications.

NOTES

3.7 KEY TERMS

- **Microsoft Disk Operating System (MS DOS):** It is a single user, single tasking operating system.

NOTES

- **Virtual Address space:** It refers to the set of memory locations that a process can access.
- **Kernel:** It controls the resources of the system and forms an interface between the shell and the hardware.
- **Shell:** It is a program that enables the end user to request the kernel to do the processing of a given command.
- **Windows 2003:** This operating system was designed and developed over various functional parts of Windows 2000 and Windows XP.

3.8 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What are the limitations of MS DOS?
2. Write the function of DATE command in DOS.
3. Name the different layers that the UNIX architecture is divided into.
4. Write a short note on how to log on with UNIX.
5. What is an echo command?
6. What are the four main subsystems of a UNIX kernel?

Long-Answer Questions

1. Discuss about the disk operating system and its types with the help of appropriate examples.
2. Explain some of the internal DOS commands and the importance of wildcard characters used in displaying more information in files.
3. Describe the different versions of UNIX operating system.
4. Explain the various commands that UNIX uses for retrieving online help.
5. Enumerate the steps required to configure and compile a UNIX kernel.
6. Describe the organization of virtual address space of process.

3.9 FURTHER READING

Stuart, Brian L. *Principles of Operating Systems: Design & Applications*. Boston: Thompson Learning.

Das, Sumitabha. *UNIX Concepts and Applications*. New Delhi: Tata Mcgraw Hill.

Stuart, Brian L. *Principles of Operating Systems: Design & Applications*. Boston: Thompson Learning.

Stallings, William *Operating Systems: Internals and Design Principles*. New Delhi: Pearson Education Inc.

UNIT 4 DBMS AND COMMUNICATION SYSTEMS

NOTES

Structure

- 4.0 Introduction
- 4.1 Objectives
- 4.2 Database Management Systems
 - 4.2.1 Data Processing
 - 4.2.2 Users of a DBMS
 - 4.2.3 Benefits and Limitations of the Database Approach
 - 4.2.4 Structure of a DBMS
 - 4.2.5 Database Administration
 - 4.2.6 DBMS Engine
 - 4.2.7 Data Models
- 4.3 Basic Communication System
 - 4.3.1 Channel Characteristics
 - 4.3.2 Modes of Transmission
 - 4.3.3 Data Transmission Systems/Technique
 - 4.3.4 Communication Parameters
 - 4.3.5 Error Detection
- 4.4 Computer Virus
 - 4.4.1 New Technology and Viruses
 - 4.4.2 Classification of Viruses
 - 4.4.3 Data Attacks
- 4.5 Internet
 - 4.5.1 Uniform Resource Locator (URL)
 - 4.5.2 Domain Name System (DNS)
 - 4.5.3 Web Browsers
 - 4.5.4 Internet Service Provider (ISP)
 - 4.5.5 Internet Services
 - 4.5.6 Internet Requirements
 - 4.5.7 Intranet
- 4.6 Answers to 'Check Your Progress'
- 4.7 Summary
- 4.8 Key Terms
- 4.9 Self-Assessment Questions and Exercises
- 4.10 Further Reading

4.0 INTRODUCTION

Database means a place where data can be stored in a structured manner. It is a shared collection or batch of data that is logically related. It has various objectives varying from organization to organization. Data is an asset for any enterprise or entity. The basic objective of a database is to provide security, safety and storage to data. In a database management system (DBMS), all files are integrated into one system, making data management more efficient by providing centralized control on the operational data. A DBMS is requested to carry out various operations, such as insert, delete, update and retrieval, on the database by the user. There are several parts of a DBMS to carry out the requested operations on the database and provide required data to the users.

NOTES

Data communication aims at the transfer and maintenance of data during the process, however, not the actual generation of information at the source and receiver. Analog and digital signals facilitate communications through signals. An analog is represented as a continuous electromagnetic waveform and digital communications represents information in binary form through a series of discrete pulses.

Internet is a 'network of networks' that links computer systems around the world. It allows communication across networks, i.e., communication can take place between one network and another. The ISP allows you to set the personal WWW homepage on the Web server. The Web server runs a script or an application program, if requests arrive. The various applications related to the Internet such as WWW, HTTP, HTML, etc. WAP is designed to provide Internet and advanced telephony services that are independent of manufacturer, vendor and technology.

In this unit, you will study about the database management system, structure of DBMS, DBMS engine, database administrator, Database users, data models, communication systems, data transmission, methods of transmission, error detection, communication parameters, modes of transmission, communication channels, computer virus internet, uniform resource, web browsers, domain name service, internet service provider, internet requirements, internet talking and intranet.

4.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the basics of database management system
- Describe the structure of DBMS
- Explain the significance of DBMS engine and database administrator
- Discuss the database users and data models
- Explain about the communication systems
- Understand the methods of transmission and error detection
- Describe the communication channels
- Define the viruses and worms
- Explain the significance of internet and uniform resource locator
- Discuss the web browsers and domain name system (DNS)
- Describe the internet service provider and internet requirements
- Explain the concept of internet talking and intranet

4.2 DATABASE MANAGEMENT SYSTEMS

4.2.1 Data Processing

With a database management system (DBMS), the scenario is totally different. Programs do not deal with stored data by their locations but they are provided with a software by a DBMS. This software allows application programs to deal

with data field names irrespective of the location of the fields within the records, the location of the records within a file and the file within a device. In a DBMS, all files are integrated into one system, thus making data management more efficient by providing centralized control on the operational data. Database management systems are not only used in the commercial applications, but also in many of the scientific/engineering applications.

Database

Database means a place where data can be stored in a structured manner. It is a shared collection or batch of data that is logically related, along with their descriptions designed to meet the information requirements of an organization. A database is a complex data structure. It is stored in a system of mutually dependent files containing the following information:

- The set of data available to the user, the so-called ‘end-user data’. This is the real data, which can be read, connected and modified by the user (if he has the corresponding rights).
- The so-called ‘metadata’ (the data describing the end-user data). Here, the properties (e.g., their type) and the relative relations of the end-user data are described.

DBMS

DBMS, which is short for database management systems, is a software system that allows users to not only define and create a database but also maintain it and control its access. It can be called a collection of interrelated data (usually called database) and a collection or set of programs that helps in accessing, updating and managing that data (which forms part of a database management system).

The primary benefit of using a DBMS is to impose a logical and structured organization on data. A DBMS provides simple mechanisms for processing huge volumes of data because it is optimized for operations of this type. The two basic operations performed by any DBMS are as follows:

- (i) Management of data in the database
- (ii) Management of users associated with the database

Management of the data means specifying how data will be stored, structured and accessed in the database. This includes the following:

- **Defining:** Specifying data types and structures and constraints for data to be stored
- **Constructing:** Storing data in a storage medium
- **Manipulating:** Querying, updating and generating reports
- **Sharing:** Allowing multiple users and programs to access data simultaneously.

Further, the DBMS must offer safety and security of the information stored, in case unauthorized access is attempted or the system crashes. If data is required to be shared among many users, the system must ensure that possible anomalous results are avoided.

NOTES

NOTES

The management of database users means managing the users in such a way that they are able to perform any desired operation on the database. A DBMS also ensures that a user cannot perform any operation for which he is not authorized.

In short, a DBMS is a collection of programs performing all necessary actions associated with a database. There are many DBMS available in the market, such as Access, dBase, FileMaker Pro, Foxpro, ORACLE, DB2, Ingress, Informix, Sybase, etc.

A database application is a program or a set of programs that interacts with the database at some point in its execution. It is used for performing certain operations on data stored in the database. These operations include inserting data into a database, extracting data from a database based on certain conditions, updating data in a database and producing data as output on any device such as a screen, a disk or a printer.

A database system is a collection of application programs that interacts with the database along with DBMS (and sometimes the users who use the system). Database systems are designed in a manner that facilitates the management of huge bodies of information.

A database clearly separates the physical storage of data from its use by an application program to attain program–data interdependence. While using a database system, the user or programmer is unaware of the details of how the data is stored. Data can be changed or updated without making any difference to the other components of the system.

DBMS Services

A DBMS is mainly used for data or database management. The processes of database management and data management are complementary. The responsibility of data management includes data and its structure as well as the integration of data and processes. On the other hand, database management covers the security, physical implementation and maintenance of the physical databases. It is the responsibility of database management to manage and enforce the enterprise's policies related to individual databases. Almost all the additional utilities and services outlined here are provided by most database management systems:

- **Enforcement of integrity:** It is necessary for the data values stored in a database to be consistent in a certain way. The balance of a bank account, for instance, may never be below a specific amount, say Rs 1000. Integrity can be maintained through centralized control of database. It also allows the database administrator (DBA) to define validation procedures that need to be carried out whenever an attempt is made to update data, that is, modify, create or delete.
- **Transaction management:** At times, a single logical unit of work is formed by numerous operations on the database. Transaction is an action used to perform some manipulation on data stored in the database. A DBMS is responsible for supporting all the required operations on the database; it also manages the execution of transactions so that only the authorized and allowed actions are performed. The execution of transactions requires ACID

properties (**A**tomicity, **C**onsistency, **I**solation, and **D**urability). Either all operations of a transaction will be executed or none of the operations will take effect (*atomicity*). As a result of a transaction, data records are accurate (*consistency*). When two or more transactions run concurrently, their effects must be isolated from one another. If a transaction has completed its operations, its effect should not be lost even if the system fails immediately after the transaction completes (*durability*). In case of failure, abandoning the partial transaction and re-applying it becomes necessary. Also, in the event of failure, the database should be able to restore itself to a consistent state.

- **Backup and recovery:** To ensure the restoration of the previous state in case of loss of data or the logical or physical corruption of data, a DBMS keeps back-up copies of the database. It keeps a log of all the operations performed in the database so that the database can be restored up to the last consistent state after the system's failure. In such a case, following the logs, operations can be re-done to make the database up-to-date.
- **Security management:** Security management by a DBMS includes the following activities:
 - Preventing unauthorized database users from accessing the database
 - Preventing unauthorized users from accessing a part of the database information or manipulating data. This control is normally done using sub-schemas and user views or by applying access rights.
 - Protecting data to prevent unauthorized users from reading or understanding the content of the database. Data encryption is used for protecting information stored on a disk as well as information exchanged on a network.
- **Concurrency control:** Simultaneous accessing of a single database by multiple users and/or programs is possible. Some of the major issues addressed by concurrent access to data include the following:
 - o A wrong view of the database state by one user while the database is being updated by another user.
 - o Updation by multiple users, concurrently, may lead to an inconsistent state or result.
 - o DBMS must ensure the avoidance of such concurrent anomalies.
- **Storage management:** Data has to be externally stored on a high-speed, random access device. The users do not have to worry about where data is stored and in which manner. This is because they can rely on the DBMS to do that. These details are simplified by the DBMS. The effectiveness of a DBMS can be measured on the basis of its efficiency and speed in storing and retrieving data. Efficiency can be measured in two ways—space and speed. Under normal circumstances, you can have either one of them, not both. *Database manager* is a component of a DBMS which addresses such problems. All related storages and retrievals are dealt with by this component.

NOTES

- **Restructuring:** Any change in the database schema may result in inconsistency in the database. Restructuring is a utility for transferring the old database into a new instantaneous database under a new schema.

NOTES

Table 4.1 File Processing System versus DBMS

DBMS	File processing system
It is a shared set of data that is logically related. It describes the data designed to meet the information requirements of an organization.	It is a set of application programs that serves the end-users in various ways, such as by producing reports. Each program defines and manages its own data.
Redundancies and inconsistencies in data are reduced due to single file formats and duplication of data is eliminated.	Redundancies and inconsistencies in data exist due to single file formats and duplication of data.
Data is easily accessed due to standard query procedures.	Data cannot be easily accessed as special application programs are needed to access data.
Isolation/retrieval of the required data is possible due to common file format and there are provisions to retrieve data easily.	Data isolation is difficult due to different file formats, and also because new application programs have to be written.
It provides program and data independence.	There is a dependency between application program and data, because, definition of data is embedded in the application program rather than stored separately.
Integrity constraints, whether new or old, can be enforced as per need. Data integrity is determined on the data itself.	Introduction of integrity constraints is tedious and new application programs have to be written in order to determine data integrity in each application.
Atomicity of updates is possible.	Atomicity of updates may not be maintained.
Several users can access data at the same time, i.e. concurrently without problems.	Concurrent accesses may cause problems such as inconsistencies.
Security features can be enabled in a DBMS easily.	It may be difficult to enforce security features.
Cost of the system depends on application.	It involves high cost as additional hardware is needed and conversion of the existing one is required.
A DBMS is usually a large piece of software adding to an overhead.	The size of the software is small as compared to the DBMS.

Basics of a DBMS

A DBMS must have some means of representing the data in a way that a user can easily understand. A DBMS provides users with the conceptual representation of data. The system hides certain details regarding data storage and maintenance and data is retrieved efficiently. This is performed by defining the levels of abstraction at which the database may be viewed.

Logical and physical view of data

Separating the logical and physical structures of data clearly is one of the main features of the database approach. The term ‘logical structure’ indicates the manner in which the programmers view it, whereas the term ‘physical structure’ refers to the manner in which data is actually stored on the storage medium.

A logical view of data expresses the way a user thinks about data. Usually, it is expressed in terms of a *data model*.

A physical view of data refers to the way data is handled at a low level, i.e., its storage and retrieval. Specifically stated, it is expressed in terms of specific locations on storage devices and techniques used to access it.

A set of logical constructs that can help describe the structure of a database, that is, its data types, constraints and relationships, is referred to as a data model. It is also a set of basic operations that specify updates and retrievals on the database. A data model is used to refer to a set of general principles for handling data (Tsitchizris and Lochovsky, 1982). The set of principles that defines a data model may be divided into the following three major parts:

- (i) Data definition: A set of principles concerned with how data is structured
- (ii) Data manipulation: A set of principles concerned with how data is operated upon
- (iii) Data integrity: A set of principles concerned with determining which states are valid for a database

Schema and subschema

The overall description of a database is called *database schema*, which is specified during database design and is not expected to be changed frequently. The values of a data item can be fitted into a framework. A database schema includes such information as:

- Characteristics of data items
- Logical structure and relationship among those data items
- Format for storage representation
- Integrity parameters, authorization and backup policies

A **subschemas** is its proper subset designed to support 'views' belonging to different classes of users in order to hide or protect information. It refers to the same view as schema but for the data types and record types which are used in a particular application or by a particular user.

Data dictionary

A data dictionary can be treated as a special file, which stores the information about the organization and usage of data contained in the database. This information is called metadata (which means data about data). It is sometimes termed as system catalog as it documents the data in the database. In a DBMS, the detailed structure and organization of each file are stored in the system catalog. The two terms, system catalog and data dictionary, are used interchangeably. A system catalog is a repository that integrates metadata. A data dictionary is a repository that manages metadata. It is a part of the system catalog that is generated for each database. A data dictionary can function in a variety of ways, as follows:

- **Active (Integrated):** This is always consistent with the current structure and definition, maintained automatically by the system itself.
- **Passive (Non-integrated):** It is used only for documentation purpose and is not used by the DBMS software. It is simply a self-contained application and a set of files used for documenting the data processing environment. It is not consistent and managed by users of the system and modified whenever the structure of the database is changed.

NOTES

Database languages

These languages are used to define and query a database.

NOTES

Data definition language (DDL): A database scheme is specified by a set of definitions, which are expressed by a special language called DDL. DDL allows the creation and deletion of structures of database objects and provides facilities for defining and altering defined physical data structures. CREATE, DROP and ALTER are the most frequently used DDL statements. The definition also includes any constraints that are set of rules to be maintained for the integrity of a database.

A DDL statement

```
CREATE TABLE EMPLOYEE
(FNAME   VARCHAR (15) ,
LNAME   VARCHAR (15) ,
ECODE   CHAR (5) PRIMARY KEY,
DATE_JOIN DATE,
SEX     CHAR,
SALARY  NUMBER (10,2) ,
DNO     VARCHAR (5) REFERENCES DEPARTMENT (DNUMBER)) ;
```

In most DBMS, the DDL also defines user views and sometimes, storage structures. In other DBMS, separate languages such as View Definition Language (VDL), Storage Definition Language (SDL), etc., may exist for specifying views and storage structures.

In databases where there is a separation between the conceptual and internal schemas, the DDL is used to specify the conceptual schema, and the SDL is used to specify the internal schema.

An SDL statement in ORACLE

```
CREATE TABLESPACE payroll
DATAFILE 'C:/ACTS/payroll.tsp' SIZE 10M
DEFAULT STORAGE (
INITIAL 10K
NEXT 50K
MAXEXTENTS 999
PCTINCREASE 10) ;
```

For true three-schema architecture, View Definition Language (VDL) is used to specify the user views and their mappings to the conceptual schema.

An Example of a VDL statement is as follows:

```
CREATE VIEW sales AS SELECT * FROM employee WHERE dno
= 'D04' ;
```

However, in most DBMS, the DDL is used to specify both the conceptual and external schemas.

Data manipulation language (DML): Once the schemas are compiled and the database is populated with data, users need to manipulate the database. DML is a language that allows users to access as well as manipulate data. Retrieving data

from the database, inserting new data into the database and deleting or modifying the existing data are the activities that comprise data manipulation. A query refers to a statement in the DML that is used for data retrieval from the database. A query language is a subset of the DML, used to pose a query. However, the terms DML and query language are used synonymously.

Example of DML statements are as follows:

```
SELECT ECODE, ENAME, DNO, SEX FROM EMPLOYEE;
DELETE FROM EMPLOYEE WHERE ECODE = 'E01';
UPDATE EMPLOYEE SET DNO = 'D04' WHERE ECODE='E03';
```

DML can be used in an interactive mode or embedded in conventional programming languages, such as Assembler, COBOL, C, C++ Pascal or P/LI. Whenever DML statements are embedded in a general-purpose programming language, that language is called the host language and the DML is called the data sublanguage.

There are two types of DML, such as:

- (i) **Low-level or procedural DML:** This requires a user to specify what data is needed and how to get it. Examples are SQL, QUEL.
- (ii) **High-level or Non-procedural DML:** Here, the user is required to specify the data needed without specifying the manner of retrieval, for example, datalog, QBE, etc.

In most existing DBMS, the external view of data is defined outside the application program or interactive session. Data is manipulated by procedure calls to subroutines provided by a DBMS or through preprocessor statements. A uniform collection of constructs forming part of the user's programming environment, is used to define and manipulate, in an integrated environment.

Note: In most DBMS, VDL, DDL and DML are not considered separate languages but a comprehensive integrated language for conceptual schema definition, view definition and data manipulation. Storage definition is kept separate to fine-tune the performance, usually done by the DBA staff.

An example of a comprehensive language is SQL, which represents a VDL, DDL, DML as well as statements for constraint specification, etc.

As mentioned earlier, when DML commands are included in a general-purpose programming language, that programming language is called the host language and the DML is called data sublanguage.

Architecture of a DBMS

Database management systems are very complex systems. To understand general database concepts and the structure and capabilities of a DBMS, it is useful to examine the architecture of a typical DBMS.

There are two different ways to interpret the architecture of a typical DBMS — the logical DBMS architecture dealing with the way data is stored and presented to users and the physical DBMS architecture concerned with the software components that make up a DBMS.

NOTES

Logical DBMS architecture

The logical architecture of a DBMS is known as the three-level architecture. It was suggested by ANSI/SPARC (American National Standards Institute/ Standards Planning and Requirements Committee). Refer Figure 4.1.

NOTES

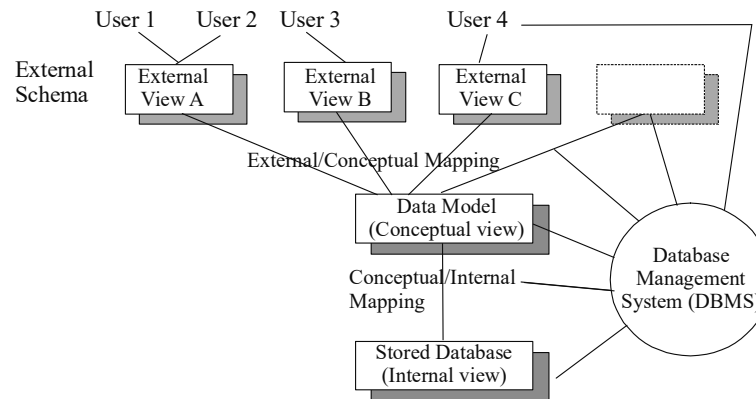


Fig. 4.1 Logical Architecture of a DBMS by ANSI/SPARC

The logical architecture describes how users perceive data in a database. A DBMS provides the user with an abstract view of data. Users can access and manipulate data without worrying about where it is located or how it is actually stored and maintained. This is done by defining levels of abstraction. There are three levels of abstraction, as follows:

- (i) **The physical or internal level** is the lowest level of abstraction. It describes how data is actually stored on physical media. It provides a low-level description of the physical database. The internal level is supported by an internal schema, which explains the physical storage arrangement of the database. The internal schema uses the physical data model, which has full information of the data storage, the access path for the database and how the data can be retrieved or inserted in the database. A data model is described as a collection of abstract tools used to describe the data, data relationship, data semantics and consistency constraints.
- (ii) **The logical or conceptual level** is the next higher level of abstraction. It describes the types of data stored and how they are interrelated. The conceptual level is supported by a conceptual schema which describes the entire database for different users who access the database. The conceptual schema hides the information of the physical storage arrangements and concentrates basically on entities, relationships and constraints.
- (iii) **The external or view level** is the lowest level of abstraction as seen by a user. It provides a window for a conceptual view. This allows the users to see only the data which is of interest to them. In other words, this level of abstraction describes only a part of the entire database or a subset of the database.

Figure 4.2 shows the logical architecture for a typical DBMS.

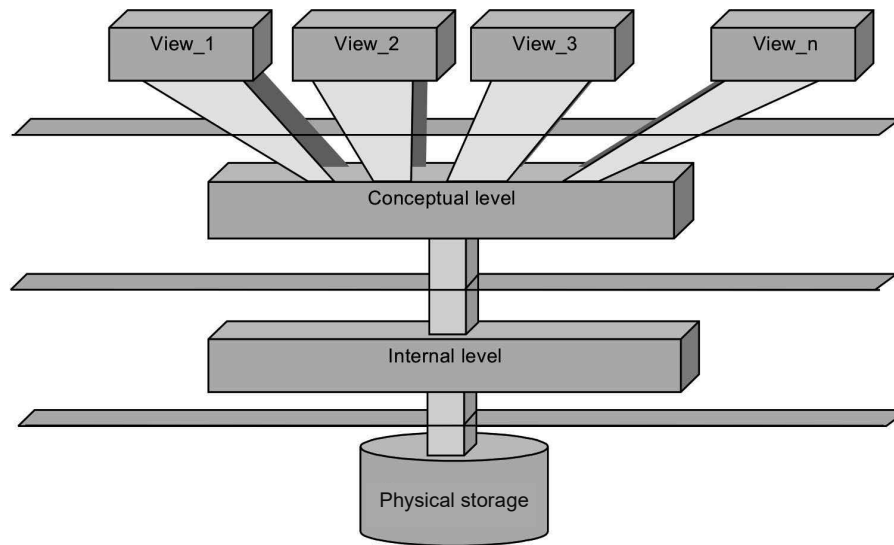


Fig. 4.2 Logical Architecture of a DBMS

The three-level database architecture allows a clear separation of data representation as the users see it and the physical data structure layout. This separation of different views of data is flexible and adaptable. This flexibility and adaptability is known as data independence.

Since a schema defines each view, there exist several schemas in the database partitioned according to the levels of abstraction. The internal view is expressed by the internal schema, which contains the definition of the stored record, the method of representing the data fields and the access aids used. The conceptual schema defines this conceptual view. There is only one conceptual schema per database. Each external view is described by means of a schema called an external schema or a subschema. Any number of external schemas can be defined and they can overlap each other.

Mapping between the levels

Mapping is the transformation of requests and results between different levels of abstraction. Mapping can be *conceptual/internal* or *conceptual/external* mapping.

The conceptual/internal mapping exists between the conceptual and internal levels. It defines the correspondence between the records and the fields of the conceptual view and the files and data structures of the internal view.

If a modification is made to the structure of the stored database, then accordingly, a change must be made in the conceptual/ internal mapping to ensure that the view from the conceptual level remains as it is. In other words, if the physical structure of the database gets modified, the DBMS has knowledge of these modifications and continues to provide the same logical view as before the changes. This is called physical data independence.

The conceptual/external mapping exists between the external and conceptual levels. This defines the correspondence between a particular external view and

NOTES

NOTES

the conceptual view. If the structure of the database is modified at the conceptual level, then the external/conceptual mapping must also change accordingly so that the view from the external level remains constant. This mapping provides logical data independence for the database.

Two types of data independence, that is, logical data independence and physical data independence can be defined with respect to the three-level architecture. The ability to modify the conceptual scheme without modifying the external schemas or application programs is called logical data interdependence. At this level, changes are usually made to improve performance.

The ability to modify the internal schema without changing the conceptual schemas or external schemas is known as physical data interdependence.

If the conceptual view is separated from the internal view, it allows the provision of a logical description of the database without any need for specifying physical structures.

Modifications to the internal schema may be required because some physical files needed reorganization. This is usually done when the logical database structure is modified. Sometimes, it is needed to make a change in the logical structure of the data.

Physical DBMS architecture

The related and interconnected software components of a DBMS are described by the physical architecture. At an extremely basic level, the physical DBMS architecture can be split into two parts: the *back end* and the *front end*. The management of the physical database and provision of relevant support and mappings for the internal, external and conceptual levels are the responsibilities of the back end. In addition, the back end is also responsible for other advantages of a DBMS, such as access control, security and integrity.

The front end is really just any application that runs on top of the DBMS and acts as a user interface. These may be applications provided by the DBMS vendor, the user or a third party.

The back end can be further divided into the functional software components, which are as follows:

DML precompiler: It converts DML statements embedded in an application program to normal procedure calls in a host language. Through the DML precompiler, DML commands and application programs written in the host language are separated. DML commands are sent to the DML interpreter for translating it into object codes for database access and the rest of the program is sent to the compiler of the host language. Object codes for the commands and the rest of the program are combined together through linking and sent to the DBMS engine (also called database manager) for execution. The precompiler must interact with the query processor in order to generate the appropriate code.

DML compiler: It translates DML statements into low-level instructions that a query processor understands.

DDL interpreter or compiler: It interprets DDL statements and records definitions into the data dictionary.

Query evaluation engine: It executes low-level instructions generated by the DML compiler.

It mainly deals with solving all problems related to queries and query processing. It helps the database system simplify and facilitate access to data.

Database manager: It is a program module providing an interface between low-level data (stored in the database) and the application programs and queries which are submitted to the system.

The functions of the database manager include:

- Efficient storage, retrieval and updation of data
- Interaction with the file manager
- Ensuring a state of consistency in the database irrespective of system failures
- Maintenance of integrity controls and user authority for data accession

File manager: It manages allocation of disk space and data structures used to represent information on a disk.

In addition, several data structures are required for the physical system implementation.

- **Data file:** Data are stored in **data files** as database.
- **Data dictionary:** Actually, this is a critical element in the DBMS. The results of the compilation of DDL statements are a set of tables that is stored in a special file called data dictionary, which documents data in a database. A data dictionary contains metadata (data about data). Metadata is data about the storage details of a database.
- **Indexes:** To improve the performance of a DBMS, a set of access aids in the form of indexes are usually provided in the database systems. An index is a data structure that helps access data through a given part of their value. It provides fast access to data. There are several indexes and for implementing indexes, several techniques are used. However, each technique is specific to a particular database application. A DBMS provides commands to build, maintain and destroy such indexes.
- **Statistical data file:** The query processor uses statistical data to optimize queries.
- **Log file:** Each log record comprises the values for database items before and after a modification and it can be utilized for the purpose of recovery.

4.2.2 Users of a DBMS

Those who use a DBMS include the following:

- **Database designer:** Before implementing a database and populating the database with data, it is required to identify the data to be stored in the database and the appropriate structure to represent and store the data. A database designer does these tasks before the database is actually

NOTES

NOTES

implemented and populated with data. The database designer analyses and integrates all external views from different groups of users and their data and processing requirements. It is the responsibility of the database designer to identify the integrity constraints during the database design.

- **Data administrator (DA):** A DA is a non-technical person responsible for managing the data resource of an organization and deciding on policies.
- **Application programmer:** An application programmer is a computer professional who writes application programs through which a user can interact with the system. Application programmers use several tools, such as RAD (Rapid Application Development) tools, to construct forms and reports without application programs. Sometimes, they interact with the system through DML calls, which are embedded in a program written in host languages. This category of database users has a clear idea of the structure of the database and knows clearly about the needs of the organizations.
- **Database administrator (DBA):** A DBA is the person (or a group of people) responsible for overall control of the database system. Once a database is installed and is functioning properly in a production environment of an organization, the database administrator takes over the charge and performs specific DBA-related activities, including:
 - Database maintenance
 - Database backup
 - Grant of rights to database users
 - Managing print jobs
 - Ensuring quality of service to all users

Routine works of a DBA include the following:

- Acting as a liaison with users to ensure that the data they require is available and to write the necessary external schemas and conceptual/external mapping (again using DDL)
- Responding to modifications in requirements, monitoring performance and modifying details related to storage and access, thereby organizing the system in a manner that will provide the best performance for the organization.
- **End-users:** They are the people who use the DBMS for querying, updating and generating reports. They can be classified into the following categories:
 - **Naive users** who interact with the system by invoking one of the permanent application programs that have been written previously by the application programmer. Tellers in a bank, reservation clerks for airlines, hotels, etc., are the native users. The main interface that a native user uses is a form interface using GUI.
 - **Sophisticated users** who interact with the system without writing a program. They form their requests using a database language, especially

DML. Other tools they use are OLAP and data mining tools, which help them summarize data in different patterns, occasionally accessing the database using database query language.

- **Casual users** who access the database occasionally and have different needs each time. they use a sophisticated query language.

NOTES

Interactions among DBMS users

A DBA can interact with the database designer during the database design phase so that he has a clear idea of the database structure for easy reference in future. This helps the DBA perform different tasks related to the database structure. The DBA also interacts with the application programmers during the application development process and provides his services for a better designing of applications. End-users also interact with the system using application programs and other tools as specified in the description above.

Objectives of a Database

A database has various objectives varying from organization to organization. Data is an asset for any enterprise or entity. The basic objective of a database is to provide security, safety and storage to data. It is an easy tool for entering the data. It facilitates the organization of data as per the requirements of the users. Integrity, sharing, availability and retrieval of data are some other objectives of a database. Databases, have made it easier for the users to handle various data. The primary objectives of a database may be discussed as follows:

(i) Availability of Data in a Database

Making available the data in an organization's database is an important objective of any database. DBMS makes available the data as well as the DBMS required to deliver it. Availability means bringing the data of an organization to the users. The system that manages data resources of the organization should be easily accessible to the people working for it. It makes the required data available whenever and wherever it is required. Availability functions include defining and creating a database, getting data in and out of a database, etc. These refer to the direct functions of a DBMS. A DBMS needs to accommodate diversity in the data stored. Most of the data of an organization is traditionally handled in accounting systems that are needed in analysing and predicting financial data. A DBMS needs to be able to handle greater diversity in the data stored, including subjective data, fragmentary marketing intelligence data, uncertain forecasts and aggregated data, as well as factual marketing, manufacturing, personnel and accounting data.

(ii) Evolvability of Data in a Database

Evolvability is the ability of the DBMS to respond to the users' changing needs and advancing technology. It refers to the system characteristic that increases the future availability of the data resources. It is different from expandability or extensibility, which implies extending or adding to the system. However, it covers expansion or contraction, both of which may occur as the system tries to change according to the changes in needs and desires of the using environment. Evolvability

NOTES

leads to the gradual unfolding, development and growth of a system so as to better meet the needs of the using environment. In the contemporary technological environment, such changes are externally administered. In the future, however, such changes may be brought automatically within the system, involving adaptive behaviour.

(iii) Shareability of Data in a Database

Shareability, i.e., the ability to share data resources, is another important objective of a DBMS.

A database facilitates shareability in the following manner:

- It serves distinctly for the types of users with varying skill levels.
- It handles different user views of the same stored data.
- It combines interrelated data and sets the database standards.
- It controls concurrent updates maintaining data integrity.
- It coordinates restart and recovery operations across multiple users.
- It represents different people and processes using the same database virtually at the same time.

(iv) Adaptability of Data in a Database

Adaptability refers to a more advanced form of evolvability in which built-in algorithms enable a system to change itself. It involves purposive, self-organizing, or self-controlling behaviour, i.e., self-regulation toward the sole criterion of success: either accessed by ultimate survival or long-term survival that can be taken as successful maintenance of data. A system representing an adaptive behaviour actively seeks a specific state or goal by changing itself whenever there is any change in itself or in its environment.

(v) Virtualization of Data in a Database

The function of the data storage system and DBMS is to provide resource virtualization to the users. Both are assembled by gateway services allowing an access to data. The various databases services, such as caching, replication, transformation and transaction, work for the optimization of data location. Data location refers to databases, SQL and RDBMS servers that are used for sending requests to the clients. The data transport services are the low-level transport protocols that help in communication between servers and users. Data transformation allows basic operations on data sets, such as searching, file testing, inserting, updating, modifying, deleting and formatting conversions of records as well as database structure.

(vi) Integrity of Data in a Database

The database integrity is required for regular and complex database. The integrity is needed to ask queries because all the data in the database is complete and consistent. Destructions, errors and improper disclosures need to be anticipated and there should be explicit mechanisms to handle them. The three primary facets of database integrity are as follows:

- Maintaining the quality of the database
- Ensuring the privacy of the database
- Protecting the existence of the database

In developing a DBMS, the accountant's concept of internal control is practically ignored. Computer specialists need such concepts so that they are able to improve the database integrity and enhance management confidence.

(vii) Security and Protection of Data in a Database

A database provides a security option using which only registered users can access the data. The various protocols, such as virtual private network (VPN) or client-server communication tunneled over secure socket layer (SSL), are assembled with DBMS/RDBMS servers so that if somebody wants to hack the system of databases, the security protocols prevent it or malicious programs. The database can be maintained safely by a set of instructions that are written in the configuration section of database servers. Firewall installed with a database is also resilient to the obstacles from foreign accessing of the database across network. For achieving this, the database system looks for a distributed technology that is flexible and is able to perform the prediction analysis of the specific organization. Sometimes, central implication of sharing the databases is needed to compromise between conflicting user needs, such as the establishment of a data structure and corresponding storage structure. This mechanism helps to keep and maintain proper backup.

4.2.3 Benefits and Limitations of the Database Approach

DBMS has its share of advantages and disadvantages which will be discussed in this section.

Benefits of a Database

A DBMS has the following three main features which allow it to provide a number of advantages for data management:

- (i) Centralized data management
- (ii) Data independence
- (iii) Data integration

The following are the advantages of a DBMS:

Centralized control on data and data source: In a DBMS, all files are integrated into one system thus making data management more efficient by providing centralized control on the operational data. By providing centralized control of data, a DBMS provides a number of advantages, including reducing redundancy, avoiding inconsistencies, sharing of data, giving better service to users, enforcing standards, etc.

Data consistency: Minimal data redundancy means improved data consistency. This improves data quality.

Data independence: Data independence can be defined as immunity of applications to any change in the physical representation and access technique. The provision of data independence is a major objective for database systems. If

NOTES

NOTES

the data is well designed, the user can access different combinations of the same data for query and report purposes.

Data integration: Since related data is stored in one single database, enforcing data integrity is much easier. In a DBMS, the integrity rules can be enforced with minimum programming in the application programs.

Data sharing: Related data can be shared across programs since data is stored in a centralized manner. This provides improved data sharing, data accessibility and responsiveness. Even new applications can be developed to operate against the same data.

Access control: A DBMS should provide facilities for protecting the security and privacy of data from unauthorized users.

Ease of application development: A DBMS provides increased productivity for application development. The application programmer is not required to build the functions that handle issues such as data integrity, security and concurrent access. All that a programmer is required to do is implement the application business rules. This makes the development of application more convenient. It is also easier to add functional modules than it is in file-based systems.

Reduced maintenance: Though data in a database can be shared by many applications and systems, maintenance is minimal and easy due to the centralized nature of the system.

Limitations of a Database

A database provides a centralized storage of data. The users need to access this data from different locations. A database provides the facility of online access to the authorized users. The accessibility of database by a huge number of users involves the risk of data manipulation and becomes the disadvantage of the database management system. A database management system is vulnerable in the following areas:

- **Data integrity:** A large number of users can access the database through the Internet. It becomes difficult to maintain the integrity of data with an increase in the volume of users of database. Data integrity becomes vulnerable when multiple users try to update data at the same time.
- **Data quality:** As data is accessible by remote users, it increases the chances of reduced data quality. The remote users can change, manipulate or damage the data. Adequate controls are needed to secure the data from manipulation.
- **Data security:** In a centralized database, data is available to remote users. It increases the chances of data abuse. To reduce the chances of unauthorized users accessing important information, it becomes necessary to take administrative and technical measures.
- **Enterprise vulnerability:** The centralization of all the enterprise information in the database makes the database an indispensable resource for the organization. The security of the central database becomes a cumbersome task for the organization as the survival of the organization may depend on the security of the database.

4.2.4 Structure of a DBMS

A DBMS is a boundary between the user and the database. The DBMS is requested to carry out various operations, such as insert, delete, update and retrieval operations on the database by the user. The various parts of a DBMS carry out the requested operations on the database and provide required data to the users. These parts are shown in Figure 4.3.

NOTES

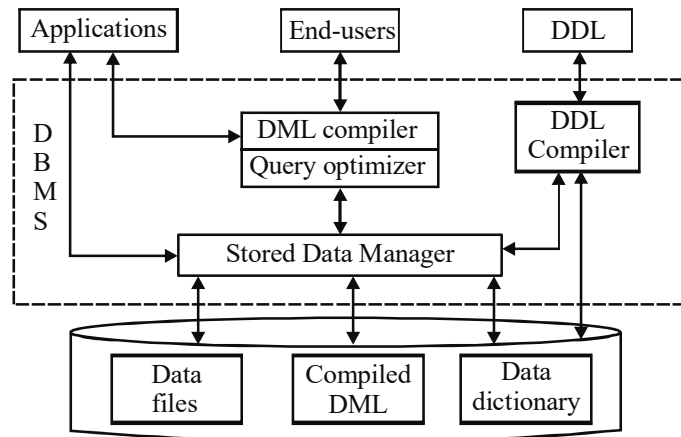


Fig. 4.3 Structure of a DBMS

The parts of a DBMS mentioned in Figure 4.3 are briefly discussed as follows:

DDL compiler: DDL compiler takes care of the processing of the schema definitions defined in the DDL. It involves metadata information, such as the name of the files, data items, storage details of each file, mapping information and constraints, etc.

DML compiler and query optimizer: DML is a language by which user accesses or manipulates the data model. Data manipulation comprises insertion of new data into the database, deletion or modification of existing data and retrieval of data from the database. The DML commands like insert, update, delete, retrieve from the application program are transferred to the DML compiler for its assembling into object code for database access. The object code is then optimized in the most favourable way to execute a query by the query optimizer and then transfer it to the data manager.

Retrieval of data is done by using a query. A query is a statement requesting retrieval of desired data from the database. Databases use a query language that is a subset of the DML. Here, terms DML and query language are used synonymously. Basically, there are two types of DML: *procedural* DML, in which user specifies data that is needed and the procedure to get it, and *non-procedural* DML, which does not specify any procedure to get data from database. It simply specifies what data is needed.

Data manager: Data manager is the central component of the DBMS software that can also be termed as the database control system.

NOTES

The main functions of data manager are as follows:

- Converting operations in user's queries obtained from application programs or combination of DML compiler and query optimizer, which is known as query processor, from logical view of the user to physical file system
- Controlling access to DBMS information stored on the disk
- Controlling and handling buffers in main memory
- Enforcing restrictions to sustain steadiness and reliability of the data.
- Synchronizing the simultaneous processes carried out by the concurrent users
- Controlling the backup and recovery operations

Data dictionary: Data dictionary is stored information about description of data in the database. It is essential for control of data integrity, proper operation of database and its accuracy. It keeps information relating to:

- *Data:* Data items are table names, names of table attributes, length of each attribute and number of records in a table.
- *Relationships between data items and transactions that reference them:* This tells the way transactions may be affected in case changes are made in certain data definitions.
- *Data constraints which mean domain of that data:* For example, if months are numbered from 1 to 12, any number below 1 and above 12 is invalid in the month field.
- *Physical design of database:* Physical design comprises structure for storage, access paths, files and record sizes.
- *Authorization for access to the database:* This is means for controlling use of the database. A database may contain sensitive information related to business or state secrets. This requires control on user access. This ensures only authorized access for the users. Database contains information about users. Different users have different privileges.
- *Statistics that contain information on frequency of transactions and queries.*

Importance of data dictionary: Data dictionary is important due to the following reasons:

- It provides better control of database over the information system and understanding of the user about the system.
- It is good for documentation of the database design process. It stores documentation pertaining to the result of every design decisions in the design phase.
- It aids in providing views on the database definitions.
- It is of great help in producing a report containing data elements that are used in programs.

- It helps data independence since any change in structures of the database does not affect the application program.

Data files: These keep data as records.

Compiled DML: After conversion of high level Queries into low level commands for file access and is called compiled DML.

End-users: The purpose of database is fulfilled when it finds an end-user. The popularity of a product, whether a programming language or a database system, lies in its use. End-users play a significant role in development of a database management system. User may be a single person having a desktop or a laptop. It may be a work group, small or big. It may be a government agency which wants to streamline the system to give speedy services to public or provide information for the public benefit. The end-user may be a research laboratory, a financial institution, a medical healthcare unit, etc.

NOTES

4.2.5 Database Administration

Database administration (DBA) installs, configures, troubleshoots and maintains a database system and also separates the function of resource management from database services and technologies, for managing data. The database administrator is responsible for monitoring and optimizing the performance of a system through disk optimization, index tuning, etc. It is also responsible for understanding the complete scenario of business lexicon where it is translated into a logical data model. DBA is administered to analyse the design phase that collects data for the design, development, testing and operational phases. DBA sets user privileges in the database environment, which affects two areas — Internet-enabled databases and storage of procedural logic in the DBMS — which are incorporated with user-defined functions and stored procedures. A comprehensive approach is used with database administration which helps the system programmers to develop a DBA interface that explains its functioning. Sometimes, the DBA functions, such as planning, coordinating and implementing the security measures to safeguard the computer databases, are used by the system programmers when more than one DBMS product is used. In such a case, the DBA performs completes the following tasks:

- Identifies and catalogs the required data for business users
- Produces conceptual and logical data models to depict accurately the relationship among various business processes of the organization
- Sets data policies for the organization
- Identifies data owners, such as administrative staff and local users
- Sets standards for central use of data
- Focusses on the DBMS and physical databases
- Opposes metadata but deals with data

For example, a value represents number 5. But, without metadata, it cannot be specifically defined. It can be analysed as follows:

- It can represent a month day, the 5th month of the year.
- It can be the 5th day of any month.

- It can be an age of 5 years.
- It can be the number of books.

Analysing with metadata, value 5 can lead to the following questions or queries:

NOTES

- Is 5 a smaller number or larger number than any other number?
- What is its domain?
- What is its data type?
- Is it an integer or a decimal number?
- Is it the mantle of modular and data planner?

These queries can be solved by DBA. DBA is incorporated to create an appropriate physical database that is transformed by a logical data model. It is prepared to recover data in a reusable point. Basically, DBA is involved with the actual implementation of databases. It can also modify complex databases. In the database world, DBA is responsible for the following complex tasks:

(i) Designing a Physical and Conceptual Schema

DBA interacts with system users to understand what type of data is to be used. It designs the conceptual schema to determine the type of relationships, such as 1:1, 1: M, M:1 or M:M, which will be maintained between various tables and attributes in databases. The physical schema determines how databases will be stored. Basically, it is used for external schema which is created by additional views.

(ii) Tuning a Database

The DBA is responsible for modifying the database to ensure and register adequate performances as per the changing requirement. It can alter data to be imported from Sequential Query Language (SQL) and Comma Separated Values (CSV) and also to export data to various formats, such as SQL, CSV, Xtensible Markup Language (XML), Portable Document Format (PDF), Word and Excel. It is featured to successfully tune with multiple database servers. Creating PDF graphics to database layout, transforming stored data using a set of predefined functions, such as Binary Large Objects (BLOB) data or download link, are maintained by DBA.

Making data available to and recovering data users.

(iii) Data Availability and Recovery for Users

In case of any system failure, the user will not be disturbed and can continue to work with uncorrupted data because DBA provides the restored data as a consistent state. Implementing the procedures to take the backup of data at regular intervals is possible through a DBA because it maintains the log of the system activities, which facilitates recovery from system crisis.

(iv) Working with Database Models

DBA transforms a logical data model into a physical database design and incorporates the knowledge of DBMS that creates an appropriate and efficient physical database design from a logical model.

(v) Performing Implementation and Operations for DBMS

The systems programming role creates an impact on DBMS implementations and operations. Sometimes, DBA plays a proxy role of system administration that is responsible for the installation and setup of DBMS. Enabling useful databases for applications and clients and designing optimal databases to administer and tune the databases are the activities performed by DBA. DBA is required to design the database so that it can be maintained with minimal disruptions.

NOTES

(vi) Enabling System Locking and Log Files

Locking the system; maintaining log errors, such as update log, binary update log, slow queries log; and monitoring startup messages and files are activities done by the DBA. Sometimes, username and timestamp are logged for queries. It supports and helps in repair works if the data file is changed during recovery. DBA changes the user file *tbl_file.MYD* as *tbl_file-datetime.bak*, if the backup file is required by the system.

(vii) Security and Authorization

Access of unauthorized data can also be prevented by DBA. In DBMS, users are granted permission to access the data which maintains certain views and relations. For example, for an online air ticket booking system, a traveller is allowed to find the details of his/her reservation and updates. However, if any user wants to see the details of others, then it is not possible because viewing important credentials and information is restricted by the database administration side. The DBA enforces this policy by granting travellers the permission to view in the read only mode, the traveller's information. The Oracle database administrator can be taken as an example. All the above functions of DBA can be performed with reference to Oracle database system whose functions are as follows:

- It is responsible to install and upgrade the application and server tools of Oracle.
- It allocates the system planning and storage management to the database systems.
- It creates primary storage space, known as tablespaces.
- It creates primary objects, such as tables, views and indexes, if application developers design an application.
- It modifies the database structure which is specified by application developers.
- It enrolls users to maintain the system security.
- It controls and monitors user access to databases.
- It plays an important role for backup and recovery database information.
- It also maintains archived data on the tape.
- It takes the backs up and restores the database.
- It monitors to optimize the performance of databases.
- It contacts Oracle Corporation to take technical support.

4.2.6 DBMS Engine

A database engine (or storage engine) is the underlying software component that a Database Management System (DBMS) uses to Create, Read, Update and Delete (CRUD) data from a database. Most database management systems include their own Application Programming Interface (API) that allows the user to interact with their underlying engine without going through the user interface of the DBMS. The term 'database engine' is frequently used interchangeably with 'database server' or 'database management system'. A 'database instance' refers to the processes and memory structures of the running database engine.

NOTES

4.2.7 Data Models

A characteristic of the database approach is that it provides a level of data abstraction by hiding superfluous details, while highlighting the details that are of interest to the application. A data model is a mechanism to provide this abstraction for different database applications.

A DBMS can choose from several approaches to manage data. Each approach constitutes a data model, which can be defined as an integrated collection of concepts or tools used to describe and manipulate data, relationships between data and semantics and constraints on data in an organization. Constraints imply a set of rules that imposes restriction on data in a database. The data model provides the necessary means to achieve the abstraction. Most DBMS provide mechanisms to structure data in the database being modelled, allow the set of operations to be defined on them and enforce certain constraints to maintain the integrity and security of data. Therefore, a data model is a mechanism for specifying the schema of a database.

Advantages of a Data Model

The advantages of a data model are as follows:

- It ensures that all data objects provided by the functional team are represented with accuracy and in complete form.
- It contains enough details to be used by the technical team who build the physical database.
- It can be used to communicate information within and across business organizations.

Entity–Relationship (E–R) Analysis

The E–R analysis is beneficial for a database designer in various ways. The following are the features of the E–R analysis:

- The constructs used in the E–R model can easily be transformed into relational tables.
- The E–R model is simple and easy to understand and is, therefore, used by a database designer to communicate a database design to users accessing the database.
- It is used as a design plan by a database developer to implement a data model in a DBMS.

Record-Based Logical Model

Data at the conceptual level and view level can be described with the help of record-based logical models. Such models are also used for specifying the overall logical structure of the database and providing a higher level description of implementation. The database in-record-based model is prepared in a fixed format with different types of records. Each record type refers to a limited number of fields or attributes and each field regards to a fixed length. There is no mechanism in it for the direct representation of code in the database but there are different languages related to models for expressing database queries and updates.

There are mainly three types of record-based logical models: relational, network and hierarchical models. You will learn the concepts of the record-based data model in the following paragraphs.

The record is maintained on the basis of existing data in a database. Data and relationship among the data is represented by a collection of tables in which every table has its own number of columns with unique names through the relational model. In the E-R model, entities and their relationships are corresponded as two-dimensional tables. The set theoretic relation that is a subset of the Cartesian product of the list of domains is the mathematical concept based on the relational model. A relation among a set of values is represented by a row in a table, where, the table is a collection of such a relationship.

Relations

Given a collection of sets H_1, H_2, \dots, H_N ; R is a relation on these N sets, if it is a set of ordered N -tuples $\langle h_1, h_2, \dots, h_n \rangle$ such that d_1 belongs to H_1 , h_2 belongs to H_2 , \dots , h_N belongs to H_N . Sets H_1, H_2, \dots, H_N ; are the domains of R and N is the degree of R . Any subset of the Cartesian product of one or more than one domain is their relation which can be represented through a table. Characteristics of the table are a lot similar to conventional sequential file with rows of the table corresponding to records of the file and columns corresponding to the fields of the records.

Student

Rollno Name House Place

108 BRIJESH YELLOW HOUSE AUSTIN

125 RADHIKA PINK HOUSE SHAWN

139 NEHA BLUE HOUSE TAKER

161 SUDHIR PURPLE HOUSE ORTON

Cartisian Products

A provided collection of sets E_1, E_2, \dots, E_N , the Cartesian product of these N sets indicated by $E_1 \times E_2 \times \dots \times E_N$ consists the set of every possible ordered N tuples $\langle e_1, e_2, \dots, \dots, e_N \rangle$ such that e_1 belongs to E_1 , \dots , $e_2 \in E_2, e_N$ belongs to E_N . ie $E_1 \times E_2 \times \dots \times E_n = \{(e_1, e_2, \dots, e_N) / e_1 \in E_1, e_2 \in E_2, \dots, e_N \in E_N\}$

NOTES

For instance, if $E1 = \{2,3,4\}$ $E2 = \{5,6\}$ then $E1 \times E2 = \{(2,5),(2,6),(3,5),(3,7),(4,5),(4,6)\}$.

Tuple

NOTES

Every row of the table corresponds to one N-tuple or tuple of the relation, .i.e., tuples are the members of a relation. An entity is represented by a tuple in the set of entity.

Cardinality

Cardinality of the relation (the number of rows in a table) is defined by the number of tuples in a relation.

Degree or Arity

Degree means the number of attributes (columns) within a relation (table). For example, each relation that is a subset of $M1 \times M2 \times \dots \times MN$ is said to have a degree N or arity N.

Components: A tuple $(s1, s2, \dots, sN)$ is said to have N components and the *i*th component is d_i .

Attribute: Attributes are defined as the columns of a table. It is a significant feature of entity as it helps in the identification of an entity. The actual content of the attribute or quality of the attribute is called attribute value.

Domain: It is a collection of values from which the values for a given attribute should be selected.

Relation scheme: It is the set of attribute names of a relation. If $B1, B2, \dots, BN$, are the attributes of a relation V then the relation scheme would be written as $V(B1, B2, \dots, BN)$; and $V(B1, B2, \dots, BN) = \{B1, B2, \dots, BN\}$. For instance, if Name, Date of birth (DOB) and Token No are the characteristics of a relation Employee, then the relation scheme for Employee will be $\{Token\ No, Name, DOB\}$ and the relation scheme would be written as $Employee(Token\ No, Name, DOB)$. The assortment of relation scheme used to signify information is termed as relational data base scheme. Relational database is the current value of the corresponding relations. 1st degree relation is called unary relation, 2nd degree is binary, 3rd is ternary and N is N-ary relations. (The relational model is also known as flat file.)

Properties of a relation

The properties of a relation are as follows:

- Relation is a set of tuples and it will not have any duplicate elements and therefore, there cannot be any duplicate rows. Two identical records can only exist if the file is conventional.
- Row and column order are insignificant.
- Information would be lost if the values are split, .i.e., the values are atomic.
- All values in the column are the values of the same attribute.

Various keys

Super key: A set of one or more attributes (columns), whose values when taken collectively can uniquely identify a tuple(a row) in a relation (table).

Primary key: A candidate key which is used as the principal means for identifying tuples in a relation.

Secondary key (Regular key): A key which identifies a set of tuples (rows / records) in a relation (table / file), or tuples having a certain property or common values for the different attributes (columns or fields) of that key. A secondary key does not identify a tuple uniquely.

Alternate key: A candidate key that is not the primary key is called an alternate key. A relation can have only one primary key and any number of candidate keys.

Concatenated key or composite key: If the primary key is formed by the combination of more than one attribute, it is termed as concatenated key.

Candidate key: A super key for which no proper subset is a super key is called candidate key.

Foreign key: An attribute or a set of attributes of a relation R1 is termed as a foreign key, if its values can be either null or values in the primary key (candidate key) of some other relation, say R2. Consider the STUDENT table with attributes RollNo., Name, Class, Place, Sex, DOB; and the CLASSES table with attributes classCode, and Description. RollNo is the primary key of STUDENT table and classcode is the primary key of CLASSES table. Class is a foreign key for the STUDENT table whose referenced key is classCode in CLASSES table.

Primary domain

A domain is a primary domain, if and only if there exists a single attribute primary key defined on that domain. For example, consider the table ITEM with attributes- ItemCode, ItemName, SUPPLIER with attributes - SupplierCode, SupplierName, PURCHASE with attributes ItemCode, SupplierCode, Qty, and CUSTOMER with attribute Name, Place, HouseName. If ItemCode is the primary key of ITEM table, SupplierCode is the primary key of SUPPLIER table, {ItemCode, SupplierCode} is the primary key of PURCHASE, and {Name, HouseName } is the primary key of CUSTOMER, then the domains of ItemCode, SupplierCode are primary domains. But the domains of Name or HouseName are not Primary domains.

Relational Model

In 1970, E. F. Codd formally introduced the relational model. Predicate logic and set theory form the basis of the relational model for database management. This model provides a simple, yet rigorously defined, concept of the manner in which data is perceived by users.

Strengths

Some of the strengths of relational models are given in Table 4.2.

NOTES

Table 4.2 Strengths of Relational Model

NOTES

Simplicity	End-users' requests are formulated in terms of the information content. These requests do not reflect any complexities due to system-oriented aspects. A relational data model is what one sees, and not necessarily, what will be implemented physically.
Non-procedural Request	Requests focus on 'What is to be done' rather than 'How it is done'.
Data Independence	Removes the details of storage structure and access strategy from the user interface. Structural flexibility is provided by relational databases; It is easier to maintain applications written for those databases. It also allows retrieval of combinations of data that may not have been anticipated as required or needed when the database was designed. To be able to make use of this characteristic, however, the design of the relations must be complete and accurate.
Mathematical Backbone	It is based on a formal theoretical model and is not only studied extensively but proven in practice. Almost every known aspect of it is actually proven in the form of mathematical theorems.

Components

The main principle of the relational model is the information principle—all information is represented by data values in relations. The three components—structural, manipulative and integrity—make up the relational model. These components are described as follows:

- The structural component is concerned with how data is represented. A set of relations represents the conceptual view of the database.
- The manipulative component is concerned with how data is operated upon. It comprises a set of high-level operations, which produces whole tables and acts upon them.
- The integrity component is concerned with determining which states are valid for a database. It is a cluster of rules for the maintenance of the integrity of the database.

The relational data model represents a database as a collection of relation values or relations where a relation resembles a two-dimensional table of values presented as rows and columns. A relation has a heading, which is a tuple of attribute names, and a body, which is a set of tuples having the same heading. The heading of a relation is also referred to as *relation schema* or *intension* and the body of the relation is referred to as *extension*. Thus, the intension of the EMP relation would be:

EMP (EmpCode, EmpName, Salary, Date_of_join, Deptno)

Intensions provide a convenient way of describing a database depicting the schema of the database. An extension refers to the rows of data values.

Relational terminologies

Domain: A domain is the set of defined atomic values for an attribute. It is a pool of values from which specific relations draw their actual values. A domain is specified in terms of data type (possibly system-defined or user-defined) and, optionally, in terms of size, range, etc.

Attribute: Attribute is the name of a role played by a domain in the relation. Each attribute A_i is defined over a domain D_i (the set of values that A_i can take on) and is the name of a feature of the real-world entity or relationship that the relation is representing. Formally, it is an ordered pair (N, D) , where N is the name of the

attribute and D is the domain that the named attribute represents, e.g., Emp Name, M GHOSH.

Relational schema: A relational schema is made up of a relation name and a list of attributes. A relation schema R is denoted by $R(A_1, A_2, \dots, A_n)$, where R is the name of the relation and A_1, A_2, \dots, A_n is a list of attributes. A relation schema is used to describe a relation.

Relational database schema: It is a set of relation schema, each with a distinct name. If R_1, R_2, \dots, R_n are a set of relation schemas, then we can write the relational database schema R as $R = \{ R_1, R_2, \dots, R_n \}$.

Relation: A relation (or relation state) r of the relation schema $R(A_1, \dots, A_n)$ is a set of n-tuples,

$$\text{i.e., } r = \{t_1, t_2, \dots, t_n\}.$$

Tuple: Each row in a relation is a set of related data values and is called a tuple. Formally, an **n-tuple**, is an ordered list of values $t = \langle v_1, \dots, v_n \rangle$ where each v_i is an element of D_i where D_i is the domain of A_i .

Degree (of a relation schema): The degree of the relation is the number of attributes (n).

Cardinality (of a relation state): The cardinality m is the number of tuples in a particular relation state.

Formally, a **relation** is defined as the subset of the subset of the Cartesian product of domains. In order to do so, first we define the Cartesian product of two sets and then the expanded Cartesian product. The Cartesian product of two sets A and B, denoted by $A \times B$ is

$$A \times B = \{(a,b): a \in A \text{ and } b \in B\}$$

The expanded Cartesian product of n sets A_1, A_2, \dots, A_n is defined by

$(A_1, A_2, \dots, A_n) = \{(a_1, a_2, \dots, a_n): a_j \in A_j, 1 \leq j \leq n\}$. The element (a_1, a_2, \dots, a_n) is called an n-tuple.

A relation $r(R)$ is a subset of the Cartesian product of the domains $D(A_i)$ that define R. Therefore,

$$r(R) \subseteq D(A_1) \times D(A_2) \times \dots \times D(A_n).$$

A **relation state** r of the relation schema $R(A_1, \dots, A_n)$ is a set of n-tuples, i.e., $r = \{t_1, t_2, \dots, t_m\}$.

Let us consider the following relation schema EMPLOYEE describing the employee information of a company. The relation EMPLOYEE is shown in Figure 4.4.

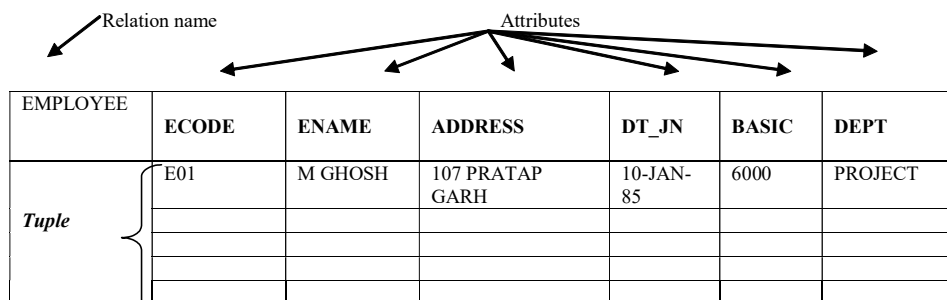


Fig. 4.4 Relation Schema Employee

NOTES

Characteristics of relations

NOTES

- A relation has a name that is distinct from all other relation names in the relation schema.
- Each attribute value of a tuple is atomic. Hence, composite and multi-valued attributes are not allowed in a relation.

According to this property, repeating groups or arrays should not form columns in a relational table. Such tables are said to be in the 'first normal form' (1NF). The foundation of the relational model is the atomic value property of relational tables and there it is important. The primary advantage of the one value property is that it makes the data manipulation logic simple.

- A distinct name is given to each attribute.
- In a relation, all the values of an attribute come from the same domain.
- There is no semantic significance in the order of attributes as long as correspondence between the attributes and their values in the relation is maintained.

This property is derived from the fact that the heading of the relation is a mathematical set (of attribute). According to this property, the ordering of the columns in the relational table is meaningless. Columns can be retrieved in various sequences and in any order. The advantage of this property is that it allows multiple users to share the same table without any concern for the manner in which it is organized. It also allows the physical structure of the database to alter without any impact on the relational tables.

- Each tuple is distinct; there are no duplicate tuples.

This property is based on the fact that the body of the relation is a mathematical set (of tuples). In mathematics, sets do not include duplicate elements. Therefore, theoretically, this property makes sure that two rows are never identical in a relational table; the values of at least one column, or set of columns, uniquely identify each row in the table. Such columns are referred to as primary keys.

- The order of tuples has no semantic significance.

This property is based on the fact that in mathematics, a set is not ordered. Since the body of the relation is represented following the set theory, this property is analogous to the one mentioned earlier. However, it applies to rows rather than columns. The primary advantage is that in a relational table, the rows are retrievable in varying sequences and orders. Addition of information to a relational table becomes simple and does not impact the existing queries.

- Derived attributes are not captured in a relation schema.

In an SQL schema, only two types of relation schema may be defined, that is VIEWS and BASE RELATION. These are called NAMED RELATIONS. Other tables, called UNNAMED RELATIONS, may be derived from these using relational operations such as join and projection.

- **Base relation:** This implies a named relation which corresponds to an entity in the conceptual schema whose tuples are physically stored in the database. A relational system must provide a means for creating the base

relations (specifically tables) in the first place. In SQL, this function is performed by the CREATE TABLE command. Base tables have independent existence.

- **View:** It is a virtual or derived relation: a named relation that does not necessarily exist in its own rights, but may be dynamically derived from one or more base relations. Its purposes may be cited as follows:
 - It provides a powerful and flexible security by hiding parts of the database from certain users.
 - It permits users to access data in a way that is customized to their needs so that the same data can be seen in different ways at the same time.
 - It can simplify complex operations on the base relations.

Disadvantages of relational model are as follows:

The relational model is the most dominant. However, it suffers from certain limitations data modal. Like the hierarchical and network models, the relational model has been developed to meet the requirements of business information processing. While applying the relational model to the application areas, such as Computer-Aided Design (CAD), simulation and image processing, many shortcomings have been noticed in this model. It is being suggested that a more sophisticated data model should be developed. The various shortcomings of this model may be discussed as follows:

- **Difficulty in modelling complex objects:** In certain circumstances, the strength of the relational model—its simple tabular data-model—becomes its weakness. The reason for this is that compressing some of the complex relationships, that exist in the real-world, into tables is a cumbersome exercise. Thus, the modelling of such complex, nested entities in a relational data model is not easy.
- **Lack of semantic knowledge:** ‘Semantic knowledge’ refers to knowledge about the meaning of data, i.e., how to interpret data, and the legitimate processes for which the data may be used. In the relational database model, this knowledge is scarce. Only the domain, entity and referential integrity rules possess semantic information. Moreover, many RDBMS do not fully support the domain concept. In such circumstances, application programmers are left with no other option but to compensate for the inability of the basic relational model to carry semantic knowledge, by building such knowledge into application programs.
- **Limited data types:** This limitation is also related to the two limitations just mentioned. An RDBMS can recognize only simple atomic data types, such as integers, characters, etc. It is one of the most critical disadvantages of RDBMS.

Network Model

The network data model can be defined as a database model used to represent objects and the relationships among these objects. In this model, a record can

NOTES

NOTES

have any number of parent records and it can also have multiple child records. Like the hierarchical model, the network model also supports the concept of data independence, which can be defined as the ability to change the representation of data at one level of a database system without the compulsion of changing the data representation at the next higher level. In the network data model, Data Manipulation Language (DML) is used for searching and retrieving records from the database. DML can also be used for connecting records from the set of instances, deleting and modifying records.

The network data model uses two types of data structures, records and set type, to define the data and relationship among data. Figure 4.5 represents a record type Employee that has three data items: Name, Sex and Birth Date.

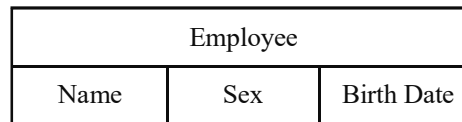


Fig. 4.5 Employee Record Type

Set type is a description of a 1:N relationship between two record types. Each set type definition has the following elements:

- Name for set type
- Owner record type
- Number record type

Figure 4.6 represents a set type R_Dept as an arrow. This representation is known as Bachman diagram. In the figure, Department is the owner record type and Employee is the child record type. This represents a 1:N relationship between the department of the company and the employees that are working in that department.

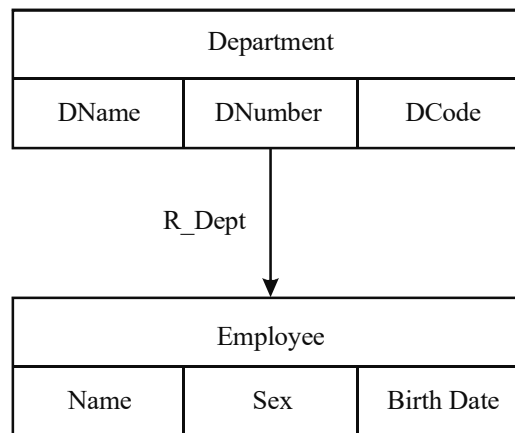


Fig. 4.6 Set Type R_Dept

In a database, there are set occurrences, also called set instances, corresponding to a set type. Each instance is used to relate one record from owner record type, i.e., Department to the set of records of member record types, i.e., Employee.

Owner serves as parent node and member serves as a child node. Each set occurrence consists of the following elements:

- One owner record from owner record type
- A member of related member records from the member record type

A record from the member record type cannot belong to more than one set occurrence of a particular set type. This represents a 1:N relationship. A set occurrence can be easily identified by the owner record or by any number of records. The following are the differences between the set instance of a database and the set in mathematics:

- The set instance in a database has one distinguished element called owner record, whereas in mathematics, there is no such type of distinction among set elements.
- In a database, all member records of a set instance are ordered. On the other hand, in mathematics, the elements of a set are not ordered.

The most commonly used implementation of a set type in a network model is the system-owned set. A system-owned set can be defined as a set that does not have any owner record type. In this set, the system can be regarded as an owner record type. It provides the following services to the network model:

- System-owned sets provide entry points into the database through the records of a specified member record type. Processing can be performed through the fields or data items of the member record type.
- System-owned sets can be used to order the records of a given record type by using set ordering specifications. By specifying the number of system-owned sets on the same record type, you can access your records in a different order.

Figure 4.7 shows a network model.

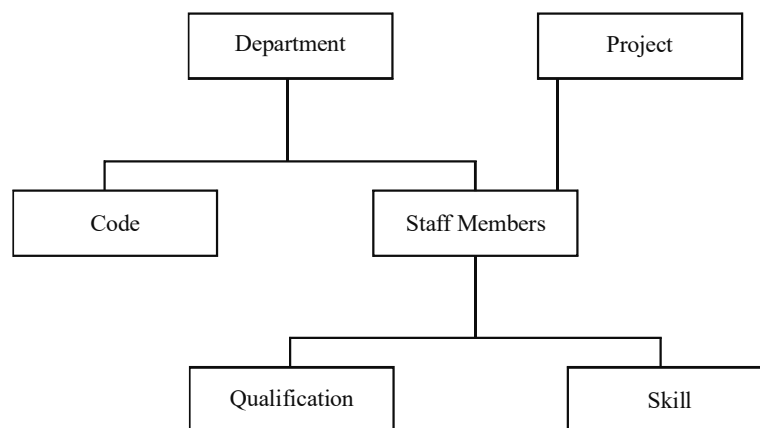


Fig. 4.7 Network Model

In Figure 4.7, Department, Project and Staff Members are the owner record types and Code, Qualification and Skill are the member record types.

NOTES

NOTES

The advantages of the network data model are as follows:

- It enables the representation of complex relationships and effect of operations, such as add and delete, on the relationships.
- It uses constructs, such as FIND, FIND OWNER and FIND NEXT, within a set that allows the users to navigate through the database.
- It can inherit the advantages of the hierarchical model.
- Many-to-many (M:N) relationships are easier to implement in a network model as compared to a hierarchical model.
- It ensures data integrity.

The disadvantages of the network data model are as follows:

- It provides a complex array of pointers, that thread through a set of records, that are not dealt with easily.
- It provides less scope for query optimization.

hierarchical model

Data models can be defined as a collection of various concepts used to describe the structure of a database. Implementing a data model includes specifying data types, relationships among data types and constraints on the data. In the hierarchical model, also called hierarchical schema, data is organized in the form of a tree structure. The hierarchical model supports the concept of data independence. Data independence is the ability to change the representation of data at one level of a database system without the compulsion of changing the data representation at the next higher level.

The hierarchical model uses two types of data structures, records and parent-child relationship to define the data and relationship among data. Records can be defined as a set of field values, which are used to provide information about an entity. An entity is a collection of objects in a database, which can be described by using a set of attributes. Records that have the same type can be easily grouped together to form a record type and assigned a name. The structure of a record type can be defined by using a collection of named fields or data items. Each data item or field has a certain data type such as character, float or integer. The Parent-Child Relationship (PCR) can be defined as a 1:N relationship between two different record types. The record type on the 1-side is called the parent record type and the record type on the N-side is called the child record type. Occurrence of the PCR type, also called instance, consists of one record of the parent record type and a number of records of the child record type. Figure 4.8 shows an example of 1:N relationship between a finance department and its employees.

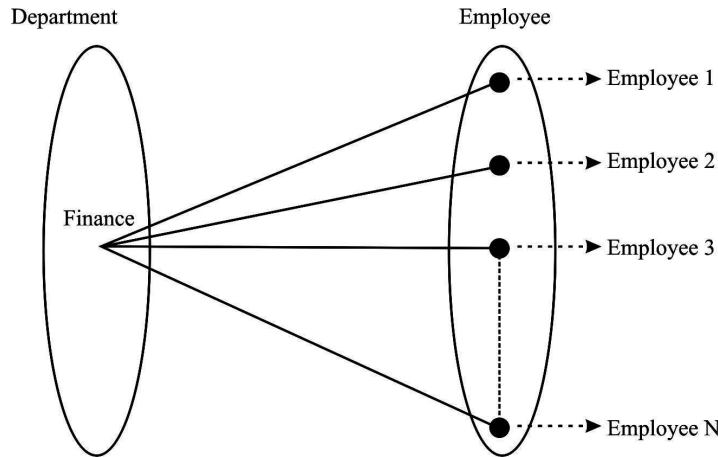


Fig. 4.8 1:N Relationship between a Finance Department and its Employees

Hierarchical schema consists of a number of record types and PCR types. In the hierarchical schema, record types are represented by rectangular boxes and PCR types are represented by the lines, which are used to connect a parent record type to a child record type. Figure 4.9 represents a hierarchical schema, which has three record types and two PCR types. Department, Employee and Project are the record types in Figure 4.9.

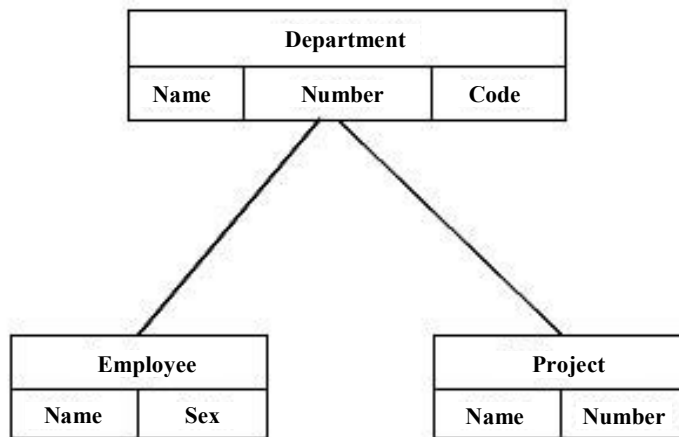


Fig. 4.9 Hierarchical Schema

Each record type can have a set of data items or fields. For example, the record type Department can have department name, department number and department code as the fields or data items. PCR type can be represented by listing pair in parentheses. For example, in Figure 4.9, there are two PCR types, which can be represented as (Department, Employee) and (Department, Project). In the figure, each occurrence of the (Department, Employee) PCR type relates one department record to the records of many employees, who work in that department. The occurrence of (Department, Project) PCR type relates a department record to the records of projects controlled by that department. Figure 4.10 represents the tree-like structure of the hierarchical schema shown in Figure 4.9.

NOTES

NOTES

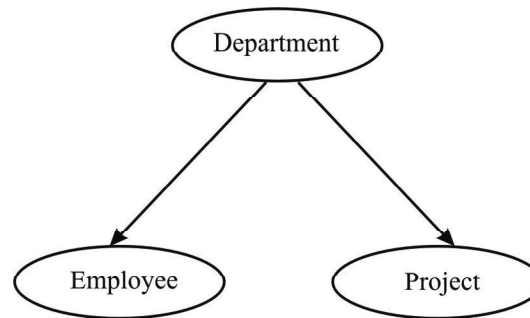


Fig. 4.10 Tree Representation of Hierarchical Schema

In a tree-like structure, a record type is represented by the node of the tree and PCR type is represented by the arc of the tree. The following are the properties of the hierarchical schema, which contains the numbers of record types and PCR types:

- One record type, called the root of the hierarchical schema, does not participate as a child record type in any PCR type.
- In the hierarchical model, each record can have only one parent record but can have many child records.
- Every record type except the root participates as a child record type in only one PCR type.
- A record type can participate as a parent record type in a number of PCR types.
- A record type which does not participate as a parent record type in any PCR type is called leaf node in hierarchical schema.
- If a record type participates as a parent node in more than one PCR type, then its child record types must be in a left to right ordered sequence.

The advantages of the hierarchical data model are as follows:

- It is simple to construct and operate on data in the hierarchical model.
- It involves hierarchically organized domains such as product info in manufacturing and employee information in organization.
- It uses constructs such as GET, GET UNIQUE and GET NEXT.

The disadvantages of the hierarchical data model are as follows:

- It requires the navigational and procedural processing of data.
- It provides less scope of query optimization.

Comparison of Network and Hierarchical Models

Table 4.3 shows the similarities and differences between network and hierarchical models.

Table 4.3 A Comparison of Network and Hierarchical Models.

	Network Model	Hierarchical Model
Data Independence	Yes	Yes
Structural Independence	No. Changes in the database structure require to be made in all related application programs.	No. Changes in the database structure require to be made in all related application programs.
Programming	Extensive programming required, as network model is implemented using linked list	Difficult to design, as you need to implement it using a tree.
Data definition	Network Data Definition Language (NDDL) is used to define data in network models.	Hierarchical Data Definition Language (HDDL) is used to define data in hierarchical models.
Data manipulation	Network Data Manipulation Language (NDML) is used to modify data in network models.	Hierarchical Data Manipulation Language (HDML) is used to define data in hierarchical models.
Constraint	A link depends on its start node and end node. If a start node or end node is deleted, the link is also deleted.	A record can only be occur if it is related to a parent record and not to a root record.

NOTES

4.3 BASIC COMMUNICATION SYSTEM

In data communication system, digital and analog communication together play a very important integrated role irrespective of many advantages of digital communication over analog. Figure 4.11 shows the integrated role of digital and analog communication to complete data communication system.

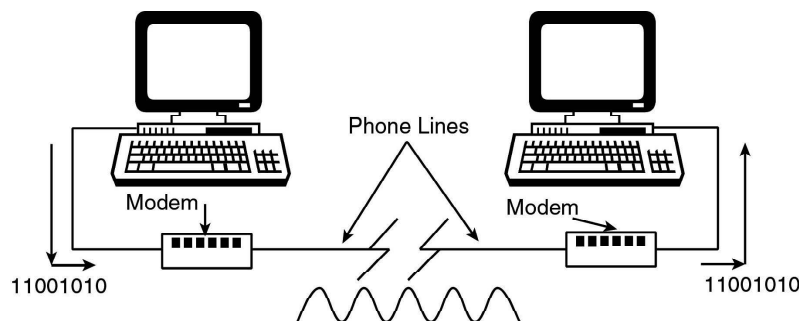


Fig. 4.11 Data Communication System

As these two signals are different in nature, how can these be connected together or communicated over the communication channel? This question can be understood by communication channels, which provide the link for data communications. Figure 4.11 shows that the link between modems is modulated analog signal created by the modem. Likewise, we may consider Figure 4.12 where a data communication system is presented in a wider sense. The communication from PC to modem consists of a binary signal whereas the communication between Central Telephone Office (CTO) and modem takes place in modulated analog signal. The communication between one CTO to another

NOTES

CTO is by digital signal using time division multiplexers, which are codecs. Thereafter, CTO feeds modulated analog signal to modem and the modem converts it into binary signal for the PC. Different types of signals emerge on the communication link and reach to CTO on their way across a big city. These can be multiplexed to share the same communication link for transmitting to destination.

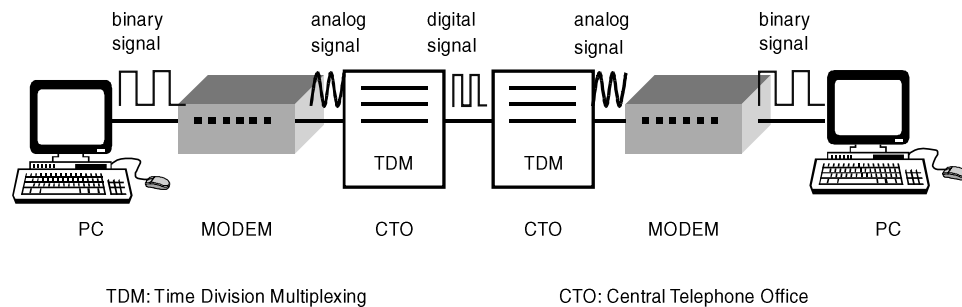


Fig. 4.12 Data Communication System in a Wider Sense

Circuits, Channels and Multichanneling

A circuit refers to a path between two or more points along which an electrical current flows. Since the discussion is about data communication, the underlying concept of circuit is to be modified accordingly. In data communication, you may consider a circuit as a specific path between two or more points along which signals can flow or be carried. The signals may be analog, binary or digital. It has been shown in the Figures 4.11 and 4.12, wherein the link between PC and modem, link between modem and CTO, and so on, constitutes a circuit.

The circuit may be a physical path comprising one or more wires. It may also be wireless. A network which may be wired or wireless is an arrangement of circuits consisting of a number of intermediate switches. A circuit may be classified based upon its uses. In case of dial-up connection or leased-line connection, a circuit is basically dedicated to a user or an owner. Again, in case of a dial-up connection, it is reserved for the user only for a limited period, the period for which the connection is invoked. On the other hand, in a leased line arrangement, a circuit is reserved in advance. Also, it can only be used by the circuit owner.

Further, a connection between two or more points may be established virtually unlike the connection made by dial-up and leased-line as physical in nature. Therefore, a circuit may be a **Virtual Circuit (VC)** based upon the type and nature of the connection. A virtual circuit may be defined as a logical path between two or more points. It is selected out of many possible physical paths available between those points. However, the connection is not guaranteed. Therefore, it seems like a fixed physical path. A permanent connection is ensured using a **Permanent Virtual Circuit (PVC)**. This provides a guaranteed connection between two or more points whenever needed without a need to reserve or commit to a specific physical path in advance. A **Switched Virtual Circuit (SVC)** is like a permanent virtual circuit, but enables users to dial in to the network of virtual circuits.

In this section, you will be mainly concentrating on the channel characteristics. With respect to data transmission, a channel may be defined as a path between transmitter and receiver. This path may be logical or physical in nature. It may also be hard wired or wireless. The path provides a passage for the information or data

from transmitter to receiver with certain amount of loss of information or data, which may be reproduced with other techniques. In some cases, the information may not be reproduced or the information may not reach the receiver at all. Such kinds of phenomenon may be very well-understood from the following channel characteristic issues.

4.3.1 Channel Characteristics

The following are the various characteristics of channel:

Channel Noise

Channel noise is the unwanted electrical or electromagnetic energy that carries no data or information but interferes with the information or data. The unwanted electrical or electromagnetic energy is nothing but a small amount of interference caused due to the disturbances produced by the electrical appliances in the neighbourhood. Thus, noise degrades the quality of information and data by affecting data and all types of communications including text, programs, audio, images, etc.

The noise may be classified as external or internal noise based upon the sources. External noise is generally caused because of the electrical appliances in the vicinity. External noise sources may range from electrical transformers to atmosphere and outer space. Generally, this noise does not hamper the performance seriously. However, if there are a number of electrical appliances in use or heavy current machines, external noise can affect communications. It also impacts communication during severe thunderstorms. A remarkably good external noise slows down the data transfer rate. In telephone communication, noise appears as a faint hissing.

The external noise is found to be directly proportional to the wavelength and inversely proportional to the frequency. It, therefore, has a remarkable impact in wireless systems as compared to the hardwired systems. The noise is generated because electricity or atmospheric disturbances is of the order of 300 KHz and therefore, tends to have more interference with the signal or information.

Noise generated in the channel or receivers is known as internal noise. Internal noise is less dependent on frequency but has a significant effect at higher frequency because external noise has less effect at higher frequencies. Minimizing the signal bandwidth may contain the internal noise but this will limit the maximum speed of the data that can be delivered. Internal noise is fairly low in case of digital signal processing, fibre optics technology.

Noise is measured in terms of the signal to noise ratio S/N or SNR with decibels (**dB**) as unit. Hence a low ratio implies that it is not worth paying attention to the medium in question. The signal to noise ratio is calculated as follows:

$$SNR = 20 \log_{10} \frac{Signal}{Noise} \text{ dB}$$

Channel Bandwidth

Channel bandwidth may be simply defined as the size of the range of frequencies that can be transmitted through a channel. In other words, we may define it as the volume of information per unit time that a computer, person, or transmission medium

NOTES

NOTES

can handle. It is measured in Hertz (Hz). Bandwidth is expressed as data speed in bits per second (bps) in digital systems while as the difference between highest frequency to lowest frequency in analog system. Bandwidth determines how fast data flows on a given transmission path. It is determined as the amount of data transmitted or received per unit time. As it has already been explained in noise that low bandwidth signal produces less internal noise compared to high bandwidth signal; therefore, this is preferred. However, in this case, we have to sacrifice data transmission speed. Therefore, a trade-off based on the performance requirements is required to be determined.

Bandwidth is dependent on the variety and physical characteristics of the transmission media, the amount of noise in the communication channel, the method of data encoding, etc.

Channel data transmission rate (bit rate)

The highest number of bits that transmits in unit time through the physical transmission media determines the channel data transmission rate. The unit of channel data transmission rate is bits per second (bps). In 1924, H. Nyquist gave the maximum rate of data of a noiseless communication channel. Further, C Shannon extended the work of H Nyquist and proposed a data rate for random noise.

Nyquist stated that if an arbitrary signal has been run through a low pass filter of bandwidth H , the filtered signal can be reconstructed by sampling the signal twice the frequency of the signal. Mathematically,

$$\text{Maximum data rate} = 2H \log_2 W / \text{Second}$$

Where, W represents the number of discrete levels in the signal.

The above is a case of a noiseless channel. If random (internal) noise is present the situation deteriorates rapidly. As we have already explained that SNR is given by a quantity $10 \log_{10} S/N$ dB.

Therefore, Shannon stated that the maximum data rate of a noisy channel whose bandwidth is H Hz and whose signal to noise ratio S/N is given by

$$\text{Maximum data rate} = H \log_2 (1+S/N)$$

Channel capacity

Channel capacity is the amount of information is passed through a communication link or transmission channel in unit time. It is measured in bits per second.

Transmission time

It is the time taken by a signal to pass over a communication link or transmission media. Transmission time is measured in seconds. It is calculated by dividing the maximum number of bits in a message by data rate. The data rate is measured in bits per second (bps). It is also given as the packet length divided by the channel capacity.

Propagation time (channel latency)

It is the time taken by a signal or information to propagate or pass from the source to the destination over a communication link or transmission media. Propagation

time is deduced by dividing the length of the communication channel or distance from the source to the destination by the speed of signal propagation. The propagation speed of electromagnetic signals is normally taken as the speed of light. The characteristics of the mediums, the speed of signal propagation and transmission distance are the major factors influencing channel latency.

Throughput

Throughput may be defined as the number of bits, characters, or blocks passing through a data communication system over a period of time.

$$\text{Throughput} = \frac{\text{Packet length in bits}}{\text{Transmission time} + \text{Propagation on time}}$$

Channel utilization

The fraction of the channels data rate that is used for the transmission of data is known as channel utilization. From the throughput it is observed that the propagation time and transmission time are two different parameters which are dependent upon the path length and packet length, respectively (number of bits in a message).

$$\text{Hence, Channel Utilization} = \frac{a}{1+a}$$

Where, a is given as the ratio of propagation time and transmission time and is known as bit length.

We may now consider an example to understand the above concepts. Suppose, a channel data transmission rate is 10 Mbps and time taken by one bit to transmit through channel 10^{-7} seconds. The signal propagation speed in the medium is 2×10^8 m/s.

The transmission rate is 1bit / 10^{-7} seconds that is equal to 10⁷ bit per second (bps).

Therefore, bit length will be equal to 2×10^8 m/s / 10^7 bps which is equal to 20 meters.

4.3.2 Modes of Transmission

From the viewpoint of the direction of transmission, communication modes can be classified into the following three types:

Simplex

Data is always transmitted only in one direction. TV broadcasting is an example. The information flows in one direction across the circuit, with no capability to support a response in the other direction. Simplex transmission generally involves dedicated circuits as shown in Figure 4.13. Simplex circuits are analogous to escalators, doorbells, fire alarms and security systems.

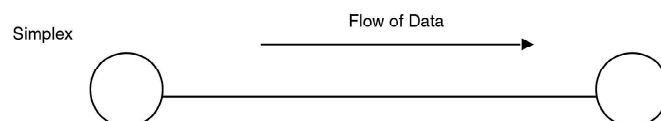


Fig. 4.13 Simplex

NOTES

NOTES

Half duplex

Data is transmitted in one direction at a time, for example, a walkie-talkie. This is generally used for relatively low-speed transmission, usually involving two-wire, analog circuits as shown in Figure 4.14. Due to switching of communication direction, data transmission in this mode requires more time and processes than under full duplex mode. Examples of half duplex application include line printers, polling of buffers, and modem communications (many modems can support full duplex also).

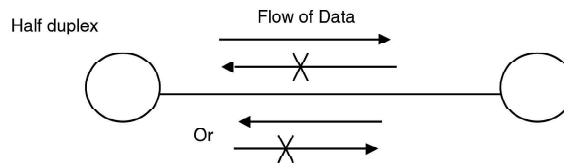


Fig. 4.14 Half Duplex

Full duplex

Data can be transmitted in both directions at the same time. In general, 4 wires, as shown in Figure 4.15, are required for full duplex transmission. Full duplex typically requires two simplex circuits, one operating in each direction. All wide-band and broadband circuits are full duplex in nature, as most of the multichannel circuits. More typical examples of full duplex applications include channel links between host processors, channel links between controllers/concentrators and hosts, and other applications involving the interconnection of substantial computing systems. Services, such as Frame Relay, SMDS and ATM based on full duplex transmission.

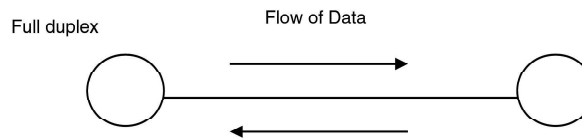


Fig. 4.15 Full Duplex

4.3.3 Data Transmission Systems/Technique

Data transmission can be divided into parallel and serial data transmission.

Parallel transmission

One or more bytes of data are sent over two or more wires. Each wire transmits one digit of binary code. Therefore, sending one byte (8 bits) of data requires 8 wires as shown in Figure 4.16. In this type of transmission, it is necessary to detect where each byte of data is separated from the next. Normally, this detection is made on elapsed time base. The interface of a printer with PC is a good example for this case.

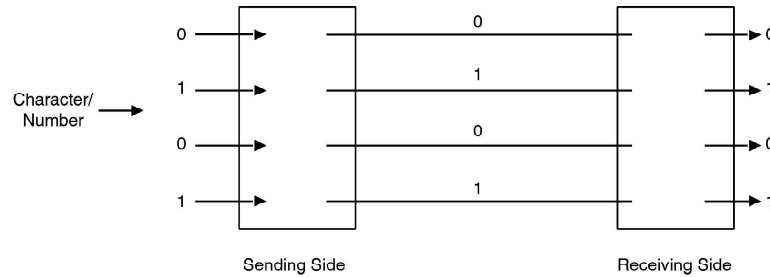


Fig. 4.16 Parallel Transmission System

Two key issues occur in parallel transfer. The wire itself is the first issue. Minimum of nine wires (eight for data bits, one for circuit ground) are required. Many a times extra wires are needed to control the flow of data across the interface. The other issue is with the nature of the bits or voltages itself. When there is change in the state of the bit/voltage from a one to zero, or vice versa, it happens at the rate of nanoseconds (one billionth of a second). A crucial part of the data transfer is the abruptness itself. The changes that occur slowly, i.e., between zero and one are not accepted as data. The electrical properties (capacitance and inductance) of a longer cable limit the suddenness with which a bit changes from zero to one, and corruption of data or loss becomes probable. Due to this, inherent speed in parallel transfer of data creates problems while transmitting over longer cables. Hence, its usage is limited to selected peripheral devices (such as, printers) used in close proximity to the computer or that operates at high speed.

Serial transmission

Data is sent over a single wire as shown in Figure 4.17. Therefore, sending one byte does not require 8 wires. These are sent one after the other. In this transmission, it is necessary to detect where each bit is separated from the next and also where each block is separated from the next. Normally, the former is detected based on elapsed time, the latter using one of a variety of so called synchronous system that will follow next. For example, the RS-232C interface and Ethernet can be cited in this category. Serial transmission is suitable for long distance data transmission because it is less costly and more resistant to noise. Therefore, almost all transmission lines for data communication systems are serial transmission lines.

Transmitting eight individual bits one after the other involves eight times more time than transmitting them all at the same time parallelly. This speed limit does not prove to be significant for several applications. Compared to the internal speed of the microprocessors, serial peripheral mechanisms are slower. They involve long, mechanical processes which restrict their speed: the speeds of their print-heads often limit the printers; the frequency limitations of the telephone lines affect the modems, and time-consuming rotational speed limits the disk drives. The speed that is built within the process of parallel data transfer is a waste on such peripheral mechanisms. The serial method thus, sacrifices a part of the speed while sufficiently servicing the peripheral devices. In such situations, the sacrifice in speed is unimportant compared to the added transmission range and reliability.

NOTES

NOTES

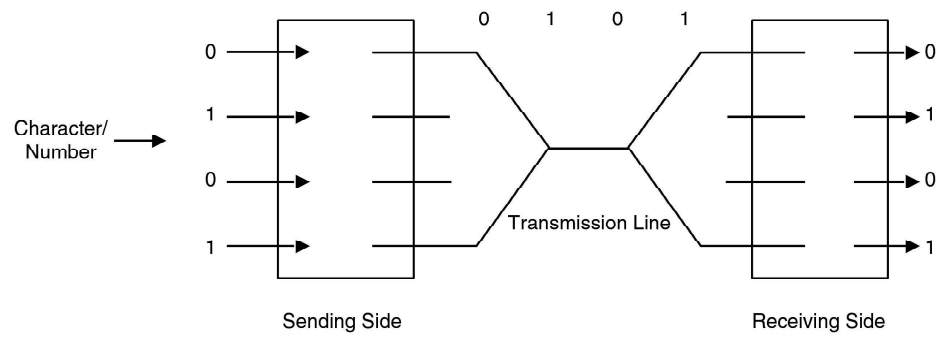


Fig. 4.17 Serial Transmission

4.3.4 Communication Parameters

Bit rate is the rate over network speed which is used to detect errors while transmitting data. The most popular method for detecting errors is inserting a parity bit alongside the data bits for a character. Receiving modems detect incorrect bit rate which is also called parity bit. It requests the sending modem to retransmit the character.

Table 4.4 String of Data with Start and Stop Bits

Stop Bit	Data Bits	Start Bit
0	1 0 0 0 0 1	0

Table 4.5 String of Data with Inserted Parity Bit

Stop Bit and Parity Bit	Data Bits	Start Bit
0	1 0 0 0 0 1	0

A modem is a bridge gap between digital and analog data transmission which allows the digital data to be transmitted/received over the telephone lines.

In telecommunication and network computing, bit rate is sometimes written as data rate and conveyed and processed per unit time. It is measured as bits per second (bit/s or bps).

Bit	SI Prefix
Kilo	Kbit/s or Kbps
Mega	Mbit/s or Mbps
Giga	Gbit/s or Gbps
Tera	Tbit/s or Tbps

The net bit rate, also called useful bit rate of a digital communication link, is the capacity of the physical layer protocol, such as Framing Bits, Time Division Multiplex (TDM) and redundant Forward Error Checking (FEC).

The operating system indicates the connection speed of the network access technology device. The speed of connection of bit rate is determined by FEC and refers to physical layer net bit rate.

For example, the connection speed of IEEE (The Institute of Electrical and Electronic Engineers) 802.11, a wireless network, is the bit rate of 6 Mbit/s and 54 Mbit/s while the gross bit rate is between 12 Mbit/s and 72 Mbit/s which includes error correcting codes. The connection speed of ISDN (Integrated Services Digital Network) BSI (Basic Rate Interface) that is 2 B-Channels+ 1D Channel) of $64+64+16=144$ Kbit/s refers to the user data rates, while the line rate is 160 kbit/s. The connection speed of Ethernet 100 BASE-TX (TRANSMIT) physical layer is 100 Mbit/s, which is also its gross bit rate.

Throughput and digital bandwidth consumption refer to the bit rate achieved in a computer network over a logical or physical communication link via network node. Throughput is affected by traffic load of network resources whereas data transfer rate is achieved by average net bit rate of all protocol overhead, data packet transmissions.

NOTES

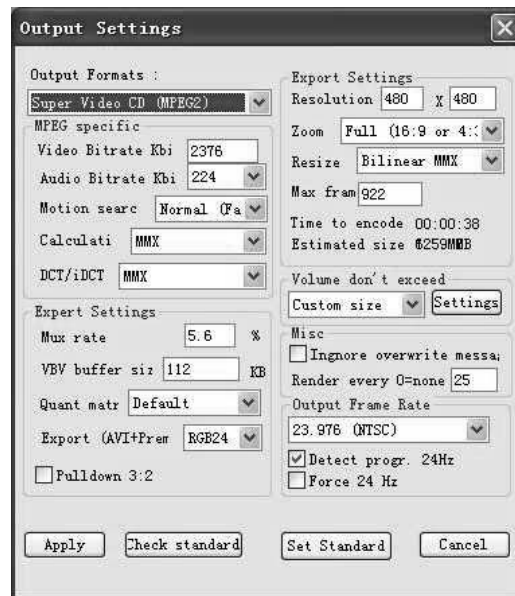


Fig. 4.18 Output Setting of Bit Rate

The Figure 4.18 shows the video bit rate of the output MPEG (Moving Picture Experts Group) file and Audio bit rate which are arranged for output setting. Motion search is set to achieve a balance between noise blurring and rendering speed around moving objects in the output file.

Multimedia Bit Rate

Bit rate is the number of bits used per unit to represent continuous medium such as audio or video following source coding (data compression) to the multimedia files. The size of the multimedia file is the product of bit rate (in bit/s) in bytes and the length of recording in seconds divided by eight. The bit rate is measured by input which avoids interrupts with reference to streaming multimedia.

Fundamentals of bit rate

- The sample uses different number of bits.
- The data is encoded by different number of bits.

- The material is sampled at different frequencies.
- The information is digitally compressed by different algorithms.

NOTES

Bit Rate Measured	Reference Standard
32 Kbit/s	MW AM Quality
96 Kbit/s	FM Quality
128-160 Kbit/s	Bass Quality
192 Kbit/s	Digital Audio Broadcasting (DAB)
224-320 Kbit/s	CD Quality
800 Bit/s	Recognizable Speech
8 Kbit/s	Telephone Quality
500 Kbit/s	Audio Format, such as FLAC, WavPack, Monkey's Audio
1411.2 Kbit/s	Compact Disc Digital Audio Video (MPEG2)
16 Kbit/s	Video Phone Quality
128-384 Kbit/s	Business – Oriented Videoconferencing System Quality
1.25 Mbit/s	VCD Quality
5 Mbit/s	DVD Quality
15 Mbit/s	HDTV Quality
36 Mbit/s	HD DVD Quality
54 Mbit/s	Blue-Ray Disc Quality

Properties of bit rate with reference to speed

The bit rate is fixed and uniform for a specific network. The gross bit rate is the number of bits transmitted per second by an ideal transmitter. It could be as high as 1 Mbit/s. The net bit rate means the number of useful bits carried per second. The latency of bit rate is the time interval between the instant of initialization, that is, transmission request and the actual start of transmission including all algorithms and parameters. It is used to calculate the speed of time to access the network without getting error frames.

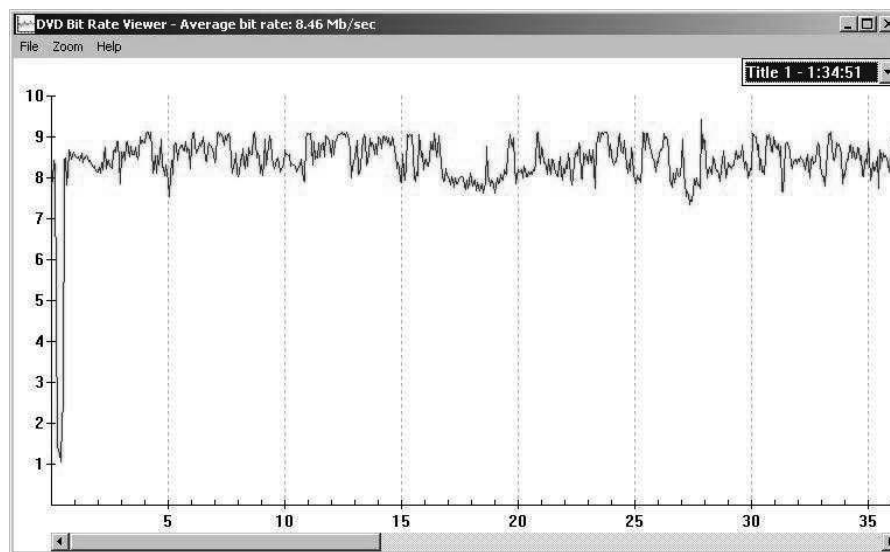


Fig. 4.19 DVD Bit Rate via Network Speed

Figure 4.19 shows the graph of DVD bit rate viewer as average bit rate according to network searching speed.

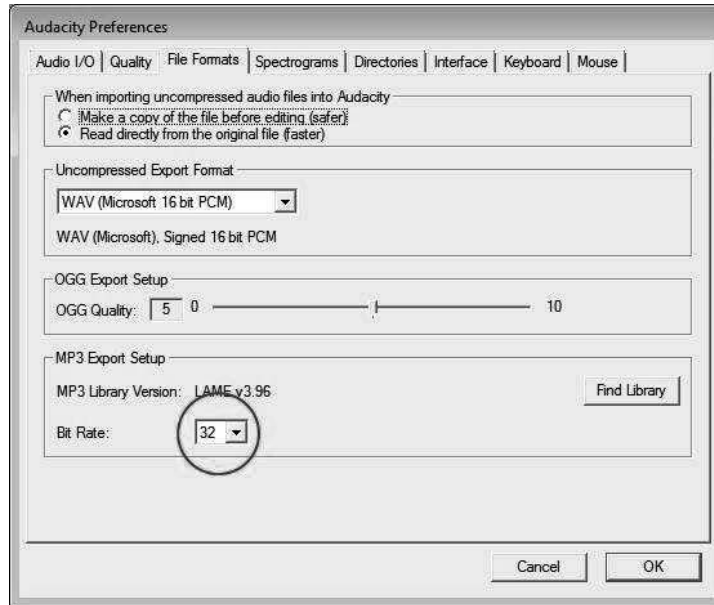


Fig. 4.20 Audacity Preferences

Figure 4.20 shows how the bit rate is set for MP3 (Music Player 3) Export set up as library version. Here, it is set as 32 Kbit/s. It is significant to export the messages as MP3s because they are smaller files as compared to WAV (Windows Media Audio) files.

4.3.5 Error Detection

In computer science, the error detection and correction or error controls are techniques that allow reliable delivery of digital data over unreliable communication channels. Many communication channels are subject to channel noise and thus errors may be introduced during transmission from the source to a receiver. Error detection techniques allow detecting such errors while error correction enables reconstruction of the original data. The following are definitions of the error detection and correction terms:

- *Error detection* is the detection of errors caused by noise or other impairments during transmission from the transmitter to the receiver.
- *Error correction* is the detection of errors and reconstruction of the original error free data.

Error correction is accomplished in following two different ways:

- **Automatic Repeat Request or ARQ:** It is also sometimes referred as backward error correction. This is an error control technique in which an error detection scheme is combined with requests for retransmission of erroneous data. Every block of data received is checked using the error detection code used and if the check fails retransmission of the data is requested. This may be done repeatedly until the data is verified.

NOTES

- **Forward Error Correction or FEC:** The sender encodes the data using an Error Correcting Code (ECC) prior to transmission. The additional information or redundancy added by the code is used by the receiver to recover the original data.

NOTES

ARQ and FEC can be combined to correct minor errors without retransmission while major errors are corrected via a request for retransmission which is called Hybrid Automatic Repeat Request (HARQ). The basic idea to perform error detection and correction is to add some redundancy, i.e., some extra data, to a message which receivers can use to check consistency of the delivered message and to recover data determined to be erroneous. Error detection and correction schemes can be either systematic or non-systematic. In a systematic scheme, the transmitter sends the original data and attaches a fixed number of *check bits* or *parity data* which are derived from the data bits by some deterministic algorithm. If only error detection is required a receiver can simply apply the same algorithm to the received data bits and compare its output with the received check bits. If the values do not match an error has occurred at some point during the transmission. In a system that uses a non-systematic code, the original message is transformed into an encoded message that has at least as many bits as the original message.

A parity bit is a bit that is added to a group of source bits to ensure that the number of set bits, i.e., bits with value 1 in the outcome is even or odd. This method is very simple and can be used to detect single or any other odd number, for example three, five, etc., of errors in the output. An even number of flipped bits will make the parity bit appear correct even though the data is erroneous.

A checksum of a message is a modular arithmetic sum of message code words of a fixed word length, for example byte values. The sum may be negated by means of a 1's complement prior to transmission to detect errors resulting in all zero messages. Checksum schemes include parity bits, check digits and longitudinal redundancy checks.

Error detection is done using a suitable hash function or checksum algorithm. A hash function adds a fixed length *tag* to a message which enables receivers to verify the delivered message by recomputing the tag and comparing it with the one provided. A Cyclic Redundancy Check (CRC) is a single burst error detecting cyclic code and non-secure hash function designed to detect accidental changes to digital data in computer networks. Cyclic codes have favorable properties and are well suited for detecting burst errors. CRCs are particularly easy to implement in hardware and are therefore commonly used in digital networks and storage devices, such as hard disk drives. Even parity is a special case of a cyclic redundancy check where the single bit CRC is generated by the divisor $x + 1$. Any error correcting code can be used for error detection.

The memory and network technologies are not entirely reliable and are prone to errors while communicating data. Data processing and transmission systems experience errors due to several reasons. Some of the reasons are:

- (i) Electrostatic interference can be caused from circuits or machines lying close by.

- (ii) Inductance and capacitance, loss in transmission due to leakages, impulses from static in the atmosphere, etc., lead to distortion.
- (iii) Resistance to current in a cable can be caused due to attenuation.

Briefly, the main constraints are due to the physical medium that produces noise, distortion and attenuation or dropouts. However, the digital signals do not suffer from noise or distortion but they are susceptible to dropouts. The simplest way of detecting errors is to add up the sum of a certain number of words and embed the sum with the data. If the received data does not add up to the so called checksum, the data has an error. However, this approach is good in detecting errors but this cannot reconstruct the original signal. Another way to minimize the effect of data loss is to use the interleaving process in which bits of each word are spread out between the bits of previous and subsequent words to reduce the chances that all the bits of any word are lost. This makes reconstruction of the data a bit easier. Thus, the error detection and correction techniques involve the approaches to improve the ability to determine what the original data was after a loss or attenuation by using mathematical processes. A set of techniques that involves coding are used to detect errors in stored or transmitted data and rectify them. The use of several parity bits, helps to detect any error, find if any bits are inverted, and in case they are, they should be re-inverted so that the original data is restored. Adding extra bits add to chances of detecting and correcting multiple errors. There are different methods depending upon Single Error Correction, Double Error Detection (SECDED).

NOTES

Check Your Progress

1. What is a database?
2. Define the term metadata.
3. What is the function of DDL compiler?
4. Write the advantages of a data model.
5. What do you mean by channel noise?
6. Define the term parity bit.
7. What is error detection?

4.4 COMPUTER VIRUS

It is believed that the program 'Elk Cloner' was the first computer virus and was created by Rich Skrenta. He developed this virus in 1982. This program was attached with Apple DOS 3.3 OS and was spread by floppy disks. In fact, other accounts related to virus also include a virus that was inputted to the system as a game by a high school student. If the game was used after the 50th time, the virus in the game replicated itself rapidly. The virus then displayed a message in a poetic format on the blank screen (see Figure 4.21). The term virus refers to an unwanted computer code.

NOTES

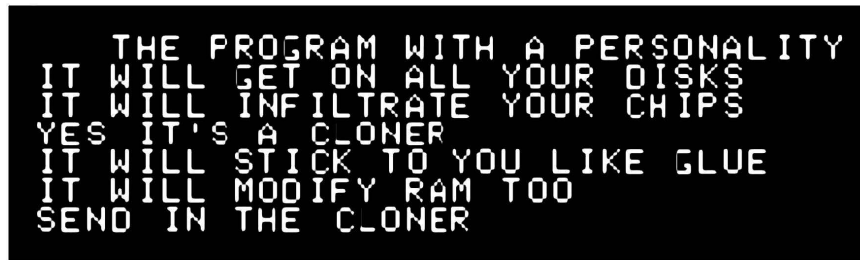


Fig. 4.21 Message Shown on Disks Infected by Elk Cloner

4.4.1 New Technology and Viruses

A computer virus refers to an executable code that is recursively attached to various other computer programs to replicate itself without the awareness of a system user. These computer viruses damage the computer system. A virus attempts to replicate itself in the system unit. It can spread via means of files, disks and recreate itself automatically without getting permission of a new host. It is estimated that almost 30,000 computer viruses exist at present. Each month, approximately 300 new viruses are created by hackers and malicious programmers. In the Internet era, 87 per cent viruses are spread across networks to which many PCs are connected. Different types of computer virus include boot sector virus, multi-partite virus, companion virus, macro and link virus. Every virus first occupies the main memory in CPU, which in turn affects the processing time as taken by a CPU. Viruses are frequently transmitted through e-mail attachments, peer-to-peer downloads, phishing (a fraudulent process to obtain a user's credentials) and instant messages. Among these, e-mail attachments carry and spread virus the fastest in an address book or a random combination of such address books. If these viruses are not controlled quickly, they can negate the functioning of servers and hence disrupt the e-mail services to all systems. Table 4.6 shows the types of computer viruses and the years of their emergences:

Table 4.6 List of Some Viruses

Computer Viruses	Year
Brain	1986
Michelangelo	1991
Mellisa	1999
I Love You	2000
Code Red	2001
Nimba	2001
Klez	2001
Slammer	2003
My Doom	2004
Storm	2007

Table 4.7 shows the recently found viruses and their effects.

Table 4.7 Viruses and their Effects

Recently Found Viruses	Effects
Oprah Winfrey Virus	The capacity of memory sector specified in hard drive reduces the size to 80MB and then expands to 200MB.
AT & T Virus	This virus alerts you on great services after every three minutes.
MCI Virus	This virus reminds you about expensive payments to be made for the AT&T virus.
Paul Revere Virus	It attacks the hard disk as well as LAN the setup.
Politically Correct Virus	It allures you to refer to a facility of extra memory that is technically referred to as 'electronic microorganism'.
Right to Life Virus	It facilitates alternatives before deleting a file.
Ross Perot Virus	It activates all the resources if the system unit starts processing files and programs.
Mario Cuomo Virus	This virus pretends to run the corrupt files and programs.
Ted Turner Virus	This virus changes the resolution setting of a monochrome monitor.
Government Economist Virus	This virus refers to itself as a diagnostic software. It also spreads the message that the system is 'OK'.
Federal Bureaucrat Virus	It segments the hard disks into little units.
Terry Randall Virus	It prints the message as ' oh no you do not ' if abort operation is done in 'Abort' 'Retry' 'Fail' prompt.
Airline Virus	This virus shows that data is in Singapore no matter where the domain setup is associated to.
PBS Virus	This virus prompts money message after every certain minutes when system is processed with files, programs and even applications.
Chicago Cubs Virus	This virus makes frequent mistakes to PC which is reviewed at the last processing.

NOTES

Following are the symptoms that indicate that a PC has been affected by viruses:

- System unit runs slower than the usual speed.
- Increase in time taken to boot up the system.
- The computer screen starts to flicker.
- The computer speakers beep at regular intervals.
- Sometimes files and directories disappear.
- System resorts to Denial of Service (DoS).

The awkward behavior of a computer system cannot be attributed to a computer virus because the OS and the programs work differently and not according to the system. The symptom of a computer virus attack is mostly in form of unusual screen displays or messages. A virus slows down the operations of a computer system. The computer drive is accessed unexpectedly if the virus infected programs reside there. The virus infected files change the length and size of the *.com or *.exe files. Details pertaining to the content of the file and date/

time change in the directory listing. Some viruses hide the details of a file if they work in RAM memory. The location of system resources also change in virus infected files.

NOTES

4.4.2 Classification of Viruses

Viruses can be classified into two types, on the basis of their location. They are as follows:

- **Non-resident virus:** The non-resident virus looks for other sources to infect and transfers control to the host program.
- **Resident virus:** The resident virus looks for a new host if the infected files are accessed by other programs.

Characteristics and Types of Viruses

The characteristics of a computer virus are as follows:

- A virus deletes registry attaches, system files and log files. It also destroys the OS of system units.
- Many viruses slow down the system performance of Internet Browsers.
- Viruses lead to a sudden decline in the speed and performance of the computer and may also lead to the loss of files or data.
- They lead to display of unknown and uncommon messages or even at times may lead to playing of a time.
- They are responsible for the loss of hard disk partition.
- They usually gain access to the system via e-mails.

Types of viruses

When a virus infects a computer system, it plays havoc on the system, generating hoax messages for every function or operation. Viruses cause unnecessary network traffic and can even cause potential damage by instructing the users to delete system files.

On the basis of their function, the primary types of viruses are as follows:

(i) Trojan horse virus

This virus was discovered on 15 October 1998 on a website called 'pro-hacker'. This virus basically attacks the secure and trusted operating system and makes users rethink if it is a beneficial and secure application. A virus gets triggered to motion when a user, in an attempt to make the system legitimate, wants to get read/write permission for a file. If the file is shared by a second user, a Trojan horse program is created which motivates data destruction. This infection is usually spread through email attachments and the chatting process, via Skype and Yahoo Messenger. For instance, a user sends an attachment, 'saxophone.wav' file to another user which is infected with a trojan horse. This file collects sound samples that are virus files leading to a spread of the virus across the network. It is different from other destructive viruses because it does not replicate. It also steals the

passwords of the authorized user and sends email disguising as the authorized user. It does not replicate the infection due to its leakage capacity.

The Trojan horse virus requires Windows to work because remote access can infect the other system unit. Once a PC is infected by this virus, the system unit tries to work with UDP port 54321, Back Orifice 2000 and UDP port 57341. Netraider 0.0 is the most commonly-used default port as used by some of the known Trojan horse viruses. In Australia, about 72 per cent ISPs were infected with Back Orifice.

The major malfunctions as shown by a Trojan horse virus infected computer are starting queries and pings to servers, rebooting the system automatically, displaying system information, logging keystrokes, editing registry file to control the server, etc. The latest dangerous trojan horse is *TR/Dldr.FraudLoad.51200* which was discovered on 16 September 2009. It represents the download subtype which has a file size of 51.200 bytes.

(ii) E-mail viruses

E-mail viruses spread through the transmission of emails, which includes both sending and receiving. For example, the **Mellisa virus**, embedded in an MS Word document was spread via e-mail. Anyone who would download and open the document ran the risk of getting infected with this virus. This virus contained a friendly note, with a person's name and other details so that the users would understand it to be authentic and asking for the recipient's personal details. After entering the required information, the virus created 50 new mails from the recipient's device. It was the fastest spreading virus that forced many big organizations to close their e-mail systems. The **I Love You** virus, that appeared on May 2000, also contained a part of the virus program code as an e-mail attachment. After clicking on the the attachment, the virus launched the code and then sent a copy of the email to other's address which in turn started corrupting files on other's systems. This email virus was termed to be more dangerous than the trojan horse virus.

(iii) Macro viruses

The macro virus (Melissa) was discovered on 26 March 1999 and it originated in Newsgroup Posting. A macro virus is written in a macro language. Macro viruses are spread via means of various applications which use and run macros. For example, a word processor allows macro programs to be embedded in documents so that the program can run automatically when the word document is opened. Due to this, the macro virus also gets initialized along with the embedded macro program and thus damages the system files. The macro virus is specific to Word 6.0, WordBasic and Excel. A **CAP** macro virus infects macros in the Word application generally attached with Word 6.0 in Windows OS, Word 6.0.1 in Windows Macintosh and MS Word in Windows 95. A **Concept** virus only spreads after a document is opened that contains the virus. Modern antivirus software can detect macro viruses as well as the other types. Figure 4.22 illustrates the macro virus and 4.23 examines the growth of macro virus from July 1995 to May 1997.

NOTES

NOTES

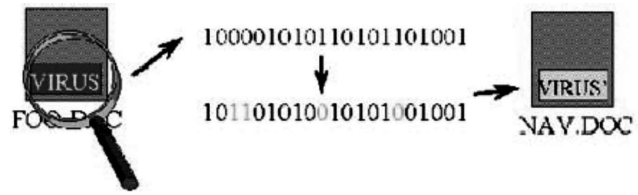


Fig. 4.22 Macro Virus

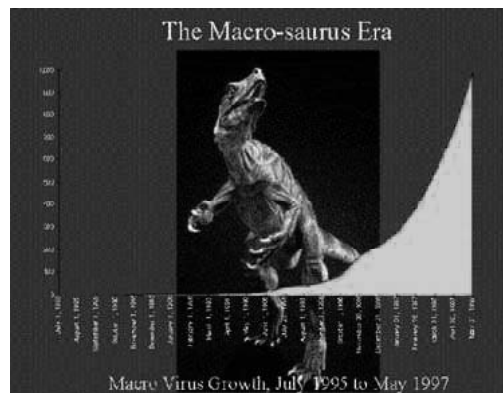


Fig. 4.23 Growth of Macro Virus from July 1995 to May 1997

(iv) Worm Virus

Worms are also computer viruses that lead to malfunctions across the network. If a user downloads the worm infected file, it spreads in his system and to other computers that are connected with that system unit. As a result, it heavily infects those system units that are on the same network. This virus is especially dangerous for businesses that are set up on a large network.

Working and Transmission of a Virus Program

This section further explains the working and spread of a virus.

(i) Working of a virus program

In order to understand how viruses work and spread on PC or on the Net, you must also understand the life cycle of computer viruses. The life cycle of a computer virus can be defined as follows:

- **Coding:** The virus program is coded and developed.
- **Releasing:** Once the code is ready, it is spread to the system and network.
- **Spreading:** It is forwarded through a simple e-mail.
- **Quarantining:** The virus gets quarantined. This phase often happens when an antivirus validates the signature of the virus and develops an antivirus update.

(ii) Transmission of a virus

Viruses spread from machine-to-machine and across networks in a number of ways. Viruses always try to trigger and execute malicious programs that are intended to be spread through the computer system. For instance, a macro virus is booted along with the infected disk which then spreads the viruses to the boot sector. It

then starts exchanging infected files on the network drive, across the Internet by downloading files and attachments from the Web. Following are the ways in which viruses can be transmitted:

- If the system unit is booted using infected files.
- If programs are executed using infected programs.
- Virus infiltration may occur via means of common routes, floppy disks and e-mail attachments.
- If pirated software and shareware are used in the system files.

NOTES

Detecting and Cleaning A Virus

Following are the two types of methods used for the detection of a virus:

(i) Static detection method

In this method, a virus is detected by searching for the pattern of viruses; for example, running a code by examining the records and files. Table 4.8 shows the list of static detection methods and their respective functions:

Table 4.8 List of Static Detection Methods

Method	Function
String Scanning method	Strings represent the sequence of bytes that are searched for in the programs.
Wildcard method	It is a detection method that allows checking the byte range; for example, wild card '?' is skipped, whereas '%' wildcard is matched with the next byte by the scanner.
Mismatch method	It detects the virus by checking the arbitrary value in the specified string.
Generic Detection method	This detection method fixes a common string. This string then detects variants of the viruses' family.
Bookmark method	It calculates a specific distance through the detection string from the starting point of the virus body.
Smart scanning method	This method skips over the junk instructions. For example, NOPs (NULL bytes). It basically selects an area of the virus body which has no subroutines or references to data files.
Skeleton Detection	The scanner checks the virus infected statements line-by-line.
Heuristics Analysis	This detection method determines the threat/risk in a system by following the weighing method based on decision rule; for example, Multi-Criteria Analysis (MCA).
Virus Specific detection	This method is based on standard algorithms due to which the scanner sometimes does not detect viruses. It introduces virus-specific detection methods, such as filtering, decryptor detection, etc.

(ii) Dynamic detection method

This method usually decides the behavior and the method of the infected code. It attempts the boot sector to modify the interrupted vectors and system files. For

example, virus activity frequently calls for I/O operations if the system unit is infected. This method is useful to search for viruses because most of the viruses exhibit a dynamic signature as follows:

NOTES

- The signature leaves an executable with read and writes permission.
- The signature reads file header that contains the starting address of the executable.
- At times, the signature is left as a portion of the file header while modifying the code.
- The signature seeks the end of the file.
- The signature appends to the file.

Cleaning a virus infected file

Viruses are malicious and smart. Many a times, users are not able to realize that their system unit has got infected with viruses. The basic property of viruses is that they hide themselves among regular system files or as attachments and camouflage themselves as standard files. The following steps are usually taken if the system gets infected with viruses:

- The Internet facility and LAN utilities must be disconnected for the time being.
- The OS must be installed within the system unit if the system has not been booted properly. If the system does not recognize the hard drive it means that it is infected with virus. It is better to boot from Windows to rescue the disk. The partition table of scandisk must be recovered using scandisk for a standard Windows program.
- A back-up of all important and critical data must be taken at regular intervals by using external devices, such as CD, USB, floppy disks or even flash memory etc.
- Anti-virus software must be installed in the system. A good quality antivirus disinfects infected objects, quarantines infected objects and is able to delete Trojans and worms.
- The latest updates must be taken to remove the antivirus database. The infected computer is not included to download these updates.
- Scans disinfect the mails of a client's database and ensure the malicious programs are not reactivated if messages are sent from one address to another across the network.
- Firewall security features must be installed in the system that prevent attacks from malicious and foreign programs.
- Delete and clean the corrupted applications and files. Try to reinstall the required applications in your system but be sure that the corresponding software is not pirated.

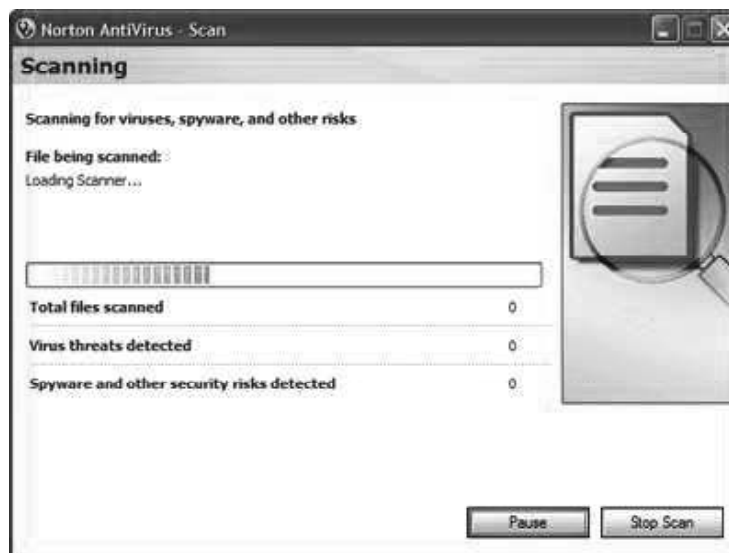
Antivirus Programs

The antivirus software available in the market to deal with virus-related issues are as follows:

- Symantec Antivirus (used to check the security of foreign programs and applications),
- Windows Vista Antivirus Spyware Antivirus Norton Antivirus (used to catch worms, rootkits, spywares, viruses etc.)
- Avast Antivirus
- Kaspersky Antivirus (used for HTTP traffic-checking and for providing a security wizard)

These antiviruses are useful for those types of viruses that are downloaded from the net or from email attachments. The most popular antivirus programs are Data Fellows F-Prot, EliaShim ViruSafe, ESaSS ThunderBYTE, IBM Antivirus, McAfee Scan, Microsoft Anti Virus, Symantec Norton Antivirus and S&S Dr Solomon's AVTK. The hard disks and drives must be scanned on a daily basis. Every week, hackers and malicious programmers release their virus programs across the net so it is better to keep one updated with the latest antivirus programs. The updated user manual and help files must be provided to the users during the installation of expensive applications and the operating system. In fact, automatic updates to the list of antivirus and multi-thread detection are the standard features of an antivirus program.

NOTES



Antivirus software is based on virus detection abilities, various features and breadth of protection. Following are the various types of protection products that are based on Windows-based system units:

(i) ZoneAlarm Extreme Security

This antivirus approach security is mainly based on the functions of intrusion prevention and malware. It also provides full drive encryption security as well as online backup support for protecting the data if a the system unit is lost or even

NOTES

stolen. The hackers can steal the password and corrupt the data in stolen computer. The main features of this antivirus are as follows:

- It analyses browser downloads in three unique ways before they can infect your PC and warns if they are malicious.
- It monitors changes within your computer to spot and stop new attacks that bypass traditional anti-virus protection
- It protects your identity with daily credit report monitoring and provides victim recovery services.
- It blocks fraudulent websites including 'phishing' sites that trick you into entering personal data.
- It protects your passwords and other key strokes from being stolen by identity thieves.
- It keeps hackers out by making your PC invisible online and blocking intrusions.
- It cleans, organizes and streamlines your computer for improved performance.
- The unified engine makes detection and removal of viruses, spyware, Trojan horses, worms, bots, and other malicious software faster and easier.
- It erases your surfing tracks and allows you to surf the Internet in complete privacy.
- It keeps your identity and data private in the event of PC loss or theft.

(ii) F-Secure Internet Security Suite

This antivirus controls the Internet access for various user types. It includes the time limit services that are imposed on Web surfing. It increases the speed by 60 per cent and consumes 70 per cent less memory space for antivirus file processing. It also provides an enhanced protection against malware, spam, e-mail, cyber criminals, viruses, etc. Following are the key features of this antivirus software:

- It is easy to install.
- It uses real-time blocking of the viruses, malware and spywares.
- It provides instant protection against new threats, for example 'DeepGuard'.
- It works with firewall to protect against hackers. It also protects from spamming and email phishing and filtering.

(iii) Panda Antivirus Pro

This antivirus software combines high technology to remove the spyware, phishing and other foreign particles from the system with the help of firewall protection, antivirus and antiphishing. It also provides free technology to support the email protection.

(iv) Trend Micro Internet Security Suite

This antivirus comes with a complete kit that provides network setup, micro servers, reporting file activities, etc. This micro server is used to analyse the data and

create the signature for checking the authentication and authorization. It is also helpful to push back the updates to the users at regular interval.

(v) BitDefender Antivirus Plus

This antivirus protects the system from instant messaging threats, spyware and viruses that offers firewall rules to stop the undesirable data trafficking and privacy gateway to identify the preferences when surfing the information.

NOTES

(vi) Eset Nod32

The Nod32 provides various features for fast scanning speeds for virus protection. It is used with firewall protection.

(vii) McAfee VirusScan Plus

This antivirus seeks the firewall and antivirus combination to make the job easy for removing spyware. SiteAdvisor service comes with this kit to protect from malicious web sites.

(viii) F-Prot Antivirus

This antivirus is used for Windows that includes command-line scanner to create a rescue CD. This CD is used to scan the system in safe mode. Stubborn resuscitators are effectively used to remove the spyware and malware but firewall is not included.

4.4.3 Data Attacks

An organization's information or computer systems may fail in many ways. Some of these are happen on purpose (maliciously) and others occur by accident. No matter why the event occurs, damage is done to the organization. There are four primary categories of attacks:

- Access
- Modification
- Denial of Service
- Repudiation

Access Attacks

An access attack is an attempt to gain information that the attacker is unauthorized to see. This attack can occur wherever the information resides or may exist during transmission. This type of attack is an attack against the confidentiality of the information.

- (a) **Snooping:** Snooping is looking through information files in the hope of finding something valuable. If the files are on paper, an attacker may do this by opening a filing cabinet or file drawer for searching. If the files are on a computer system, an attacker may attempt to open one file after another until information is found.
- (b) **Eavesdropping:** When someone listens to a conversation that they are not a part of, that is eavesdropping. To gain unauthorized access to information, an attacker must position himself at a location where information of interest to the attacker is likely to pass by. This is most often done electronically.

NOTES

- (c) **Interception:** Unlike eavesdropping, interception is an active attack against the information. When an attacker intercepts information, he is inserting himself in the path of the information and capturing it before it reaches its destination. After examining the information, the attacker may or may not allow the information to continue to its destination.

Modification Attacks

A modification attack is an attempt to modify information that an attacker is not authorized to modify. This attack can occur wherever the information resides. It may also be attempted against information in transit. This kind of attack is against the integrity of the information.

- (a) **Changes:** One type of modification attack is to change existing information, such as an attacker changing an existing employee's salary. The information already existed in the organization but it is now incorrect. Change attacks can be targeted at sensitive information or public information.
- (b) **Insertion:** Another type of modification attack is the insertion of information. When an insertion attack is made, information that did not previously exist is added. This attack may be mounted against historical information or information that is yet to be acted upon. For example, an attacker might choose to add a transaction in a banking system that moves funds from a customer's account to his own.
- (c) **Deletion:** A deletion attack is the removal of existing information. This could be the removal of information in a historical record or in a record that is yet to be acted upon. For example, an attacker could remove the record of a transaction from a bank statement (thus causing the funds that would have been taken from the account to remain).

Denial-of-Service Attacks

Denial-of-Service (DoS) attacks are attacks that deny the use of resources to legitimate users of the system, information, or capabilities (see Figure 4.24). DoS attacks generally do not allow the attacker to access or modify information on the computer system or in the physical world. DoS attacks are nothing more than vandalism.

- (a) **Denial of Access to Information:** A DoS attack against information causes that information to be unavailable. This may be caused by the destruction of the information or by the changing of the information into an unusable form. This situation can also be caused if the information still exists but has been removed to an inaccessible location.
- (b) **Denial of Access to Applications:** Another type of DoS attack is to target the application that manipulates or displays information. This is normally an attack against a computer system running the application. If the application is not available, the organization cannot perform the tasks that are done by that application.
- (c) **Denial of Access to Systems:** A common type of DoS attack is to bring down computer systems. In this type of attack, the system along with all

applications that run on the system and all the information that is stored on the system become unavailable.

- (d) **Denial of Access to Communications:** DoS attacks against communications have been prevalent for many years. This type of attack can range from cutting a wire, to jamming radio communications, to flooding networks with excessive traffic. Here the target is the communications media itself. Normally, systems and information are left untouched but the lack of communications prevents access to the systems and information.

NOTES

Repudiation Attacks

Repudiation is an attack against the accountability of the information. In other words, repudiation is an attempt to give false information or to deny that a real event or transaction should have occurred.

- (a) **Masquerading:** Masquerading is an attempt to act like or impersonate someone else or some other system. This attack can occur in personal communication, in transactions, or in system-to-system communications.
- (b) **Denying an Event:** Denying an event is simply disapproving that the action was taken as it was logged. For example, a person makes a purchase at a store with a credit card. When the bill arrives, the person tells the credit card company that he never made the purchase.

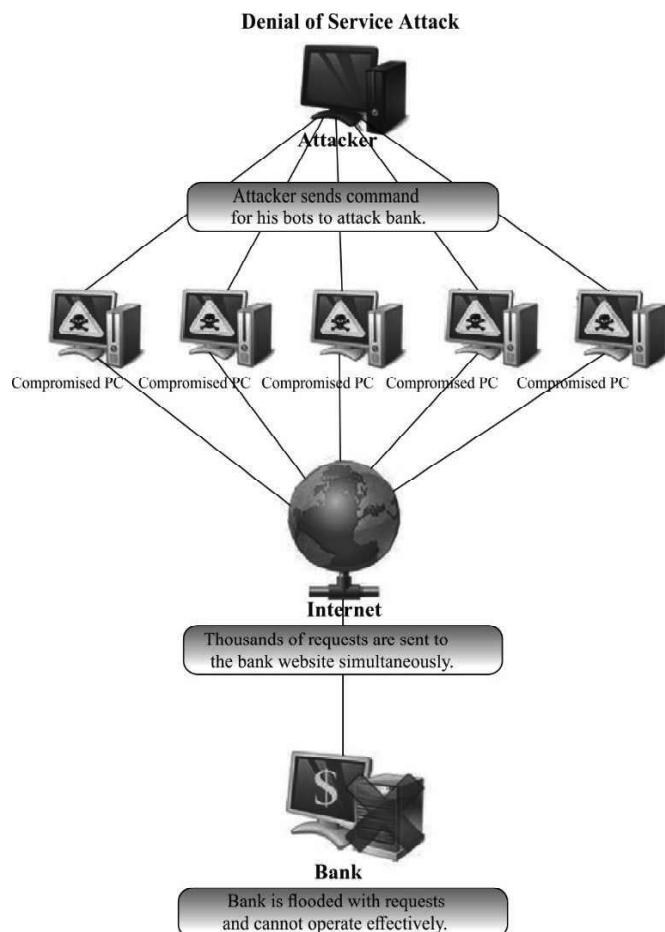


Fig. 4.24 A Typical DoS Attack

4.5 INTERNET

NOTES

Internet is a ‘**network of networks**’ that links computer systems around the world. It allows communication across networks, i.e., communication can take place between one network and another. This allows people to have access to information from different Web sites, locations and machines. It virtually makes accessible a sea of information and a nearly worldwide audience at the click of a button.

It was the Sputnik’s launch and the subsequent cold war, space race and the development of ARPANET (Advanced Research Projects Agency Network) in 1950 that led to the establishment of the Internet. However, it actually gained momentum in the 1980s when ARPANET was used by the National Science Foundation to connect the five supercomputers at its regional centres. Thus emerged a high-speed Internet service that enabled access to many universities, networks, bulletin board systems, commercial online services and institutions. The decade closed with the coming into being of the World Wide Web, which proclaimed the emergence of an independent platform of communication that was further augmented by a relatively easy-to-use and pleasant graphical interface.

Important Features

Some of the facilities that are available on the Internet are:

- **World Wide Web:** The Internet application that is currently drawing maximum attention is the World Wide Web. It has dramatically influenced the online world and continues to grow in popularity.
- **Direct Communication:** Through e-mail (electronic mail), messages can be sent to or received from any part of the world within a few minutes.
- **Round-the-Clock Availability:** Information on the Internet is available to users 24 hours a day.
- **Central Repository of Data:** The Internet is like a huge central warehouse of data that people from all over the world can access.
- **Search Engines:** These are like directories which help get any kind of information from the world over within a few seconds.
- **Advertisement:** A company can advertise its products/services through the Internet.
- **E-Commerce:** The Internet is increasingly being used for conducting monetary transactions. Through the Internet, you can shop and pay through your credit card or ask your bank to transfer your money to a different account, without even leaving your desk.
- **Distance Learning:** Several online distance learning courses are now being offered by Indian and foreign universities on the Internet.
- **BBS and New Services:** The Internet is perhaps the cheapest medium for online help. BBS services are available on the Net through which you can ask questions and get immediate troubleshooting assistance.

- **Wide Area Networks:** Using the Internet, organizations can collect and compile information from offices spread over large geographical areas.
- **Shareware Software:** The Internet is also a great medium for downloading free software. You can get a truckload of free games, utilities and trial versions of software through the Net.

Key Concept

The Internet has made things simpler. It can serve the following purposes:

- *Direct Communication:* You can send messages to family and friends, business associates and acquaintances using the electronic mail facility. Using electronic mail, you can send and receive messages within a few seconds anywhere in the world. Using Internet Relay Chat (IRC), you can communicate online with people over the Internet. You can log into a chat room and converse with others by typing messages that are instantly delivered. With the improvement of network technologies and increase in broadband, not only can you use text messages but also graphics, audio and video for communication with other people.
- *Online Shopping:* Logically, the Internet has removed all barriers of distance and nationality. You can shop for products and services across the world by logging on to a Web portal. You can also pay your bills online using credit and debit cards. You can also transfer money between different accounts with the click of a mouse.
- *Distance Education:* The Internet provides a perfect medium for knowledge sharing and information dissemination. Courses are available on the Internet. You can register and pay online, and complete a course on different interest areas. You can also pursue specialized higher studies now in the comfort of your own office or home.
- *Knowledge Base:* The Internet provides a rich information base that people from across the globe can access. In fact, it is one of the richest information bases that can be accessed at the click of a mouse. Using search engines, you can search for detailed information on any topic of your interest.
- *Banking:* Banks are using information technology to provide online banking facilities to their customers. Using the Internet, you can now view your account details, get drafts made, request for chequebooks and transfer money from one account to another. The use of ATMs has shifted the mundane back-office work to the customer himself. Instead of hiring an army of bank clerks, banks can now use ATMs to considerably reduce time and operational costs.
- *Travel:* Using the Internet, travel agencies can publish their services on the Web along with the latest discounts, packages and availability details, so that customers can compare rates, make online bookings and avail discounts without having to run around multiple offices.
- *Bill Payments:* The government sector has also realized the benefits of IT. Now you can make online payments for public utilities such as water, electricity and phones, using credit cards as the payment medium.

NOTES

Types of Connections and Internet Protocols

You can use two options for connecting to the Internet.

NOTES

Direct Connection

Through a direct connection, a machine is directly connected to the Internet backbone and acts as a gateway. Though a direct connection provides complete access to all Internet services, it is very expensive to implement and maintain. Direct connections are suitable only for very large organizations or companies.

Through Internet Service Provider (ISP)

You can also connect to the Internet through the gateways provided by ISPs. The range of Internet services varies depending on the ISPs. Therefore, you should use the Internet services of the ISP that is best suitable for your requirements. You can connect to your ISP using two methods:

Remote Dial-Up Connection

A dial-up connection (refer Figure 4.25) enables you to connect to your ISP using a modem. A modem converts the computer bits or digital signals to modulated or analogue signals that the phone lines can transmit and vice versa. Dial-up connections use either SLIP (Serial Line Internet Protocol) or PPP (Point-to-Point Protocol) for transferring information over the Internet.

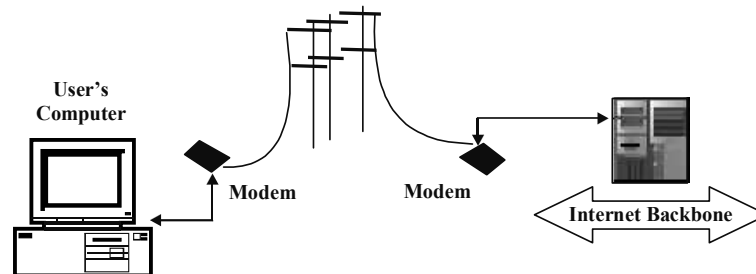


Fig. 4.25 Remote Dial-Up Connection

For dial-up connections, regular telephone lines are used. Therefore, the quality of the connection may not always be very good.

Permanent Dedicated Connection

You can also have a dedicated Internet connection that typically connects you to ISP through a dedicated phone line. A dedicated Internet connection is a permanent telephone connection between two points. Computer networks that are physically separated are often connected using leased or dedicated lines. These lines are preferred because they are always open for communication traffic unlike the regular telephone lines that require a dialling sequence to be activated. Often, this line is an ISDN (Integrated Services Digital Network) line that allows transmission of data, voice, video and graphics at very high speeds. ISDN applications have revolutionized the way businesses communicate. ISDN lines support upward scalability, which means that you can transparently add more lines to get faster speeds, going up to 1.28 Mbps (million bits per second).

T1 and T3 are the two other types of commonly used dedicated line types for Internet connections. Dedicated lines are becoming popular because of their faster data transfer rates. Dedicated lines are cost-effective for businesses that use Internet services extensively.

A Web browser, commonly known as a browser, is a computer application that creates requests for HTML pages or web pages and displays the processed HTML page. Web browsers use HTTP (HyperText Transfer protocol) to request for information from web servers. The two most commonly used web browsers are:

- Mozilla Firefox
- Microsoft Internet Explorer

Other examples of Web browsers include Mosaic, Cello, Lynx and Netscape Navigator.

Having knowledge of the basic features of browsers can be helpful in using them easily.

Toolbar of the Internet Explorer



The toolbar consists of various icons that can be used to execute functions. In fact, most of the options available through the menu bar are also available through the icons on the toolbar. Some commonly used icons are:

Back: The Back button allows you to navigate to the Web page you viewed last.

Forward: The Forward button on the toolbar navigates to the next Web page that was accessed previously. To view a list of the last few Web pages visited, you can click the down arrow button beside the Back and Forward buttons.

Stop: The Stop button can be used to terminate the current Web page request. This is usually used when you type the wrong URL by mistake and want to stop the request for the Web page or if the Web page takes too long to download.

Refresh: The Refresh or the Reload button is used to load the current Web page again. In other words, it refreshes the contents of the current page by fetching a new copy of it.

Search: The Search button allows you to find information on the Web. You can find information by clicking on the Search button on the toolbar. This activates the search text box as shown. You can then type in a word or phrase and click the 'search' button to start the search.

Favorites: The Favorites button is used to record the addresses of frequently visited Web sites. Once a website or a webpage is added to the favorites list, it can be revisited by simply clicking on the link in the list. This saves the effort of typing the URL each time the user wishes to visit the same site.

History: The History button is used to view the list of all the web pages visited in the last few days, hours or minutes. To revisit any one of them, simply click on the address.

NOTES

Print: The Print button is used to print the contents of the current web page.

NOTES

Net Surfing

Net surfing, Internet browsing or exploring a network on the World Wide Web is associated with visiting different websites on the Internet. It is typically finding places of interest at the click of a mouse. It is analogous to surfing TV channels with a remote control.

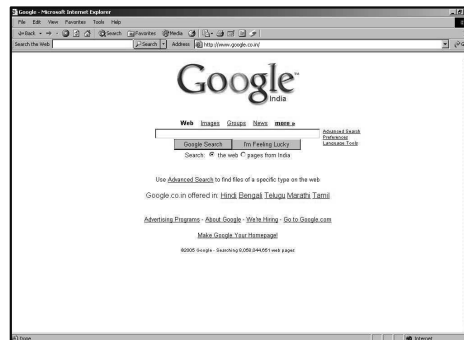
Searching

Searching is one of the most common uses of the Internet. You can search for any topic or information on the Internet. This is possible by using websites that provide a search engine.

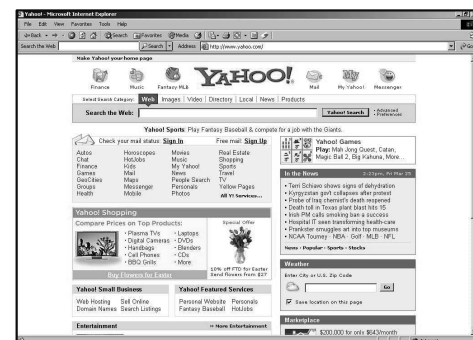
Search Engines

A search engine is a software system that enables users to search for information on the Web using keywords. It is designed to help Internet users locate Internet resources such as Web pages, documents, programs and images using a keyword search mechanism. Search engines typically use databases that contain references to a host of resources. The users interact with a search engine using an interface. There are many search engines available with different appearances and search mechanisms. Some commonly used search engines are: Google, Yahoo, MSN, Altavista, AOL, Ask Jeeves, Lycos, Excite and HotBot.

Google (<http://www.google.com>)



Yahoo (<http://www.yahoo.com>)



Information is collected by programs; called 'spiders' or 'robots'; following which, the search engine indexes them. 'Directories' also provide similar services by maintaining a systematic list of websites. Search engines function as card catalogues for the Net. They try to index and locate the requested information by looking for the keywords specified by the user. They usually maintain indices of Web resources, from which the necessary information is retrieved by querying for the keywords that the user enters.

4.5.1 Uniform Resource Locator (URL)

Each Web page has a unique address called the Uniform Resource Locator (URL) that identifies its location on the Internet. The URL consists of four parts: protocol,

name of the Web server, the directory on that server and the file within that directory. The syntax of URL is

```
protocol://domain name: port number/<directory path>/
<object name>#spot
```

Example:

```
http://www.yahoo.com/education/FIIT/home.htm#top
```

The various parts of a URL are described in Table 4.9.

Table 4.9 Parts of a URL

Part	Example	Description
Protocol	http://	represents the name of the protocol like ftp, http, etc.
Domain Name	www.yahoo.com	represents the name of the web server where the desired web page or other resource resides
Directory Path	/education/FIIT	represents the location of the web page in the web server's file system
Object Name	/home.htm	specifies the name of the file for the desired web page or the name of the other resource that is required
#Spot	#top	specifies the particular location of the text or the graphic on the web page

NOTES

Types of URL

There are two types of URL, namely, *absolute URL* and *relative URL*.

- **Absolute URL:** An absolute URL defines the exact location of the Web page or any other resource on the Internet. For example, *http://www.yahoo.com/education/FIIT/home.htm*. Here, the Web page *home.htm* is stored under the folders */education/FIIT* on the Web server *www.yahoo.com*.
- **Relative URL:** In a relative URL, each part of a URL does not need to be specified. You can abbreviate a URL by making it 'relative' to the current location. For example, suppose there exists a file *mypicture.gif* residing in the *image* folder under the *FIIT* folder. The relative URL for the file will be */images/mypicture.gif*.

4.5.2 Domain Name System (DNS)

Every machine on the Internet is assigned a unique IP address, which is a 32-bit or 128-bit numeric address. An IP address is the contact detail that the Internet Protocol (IP), part of TCP/IP, uses to identify host computers and route data to them. It is written as a string of four numbers separated by dots ('.'). Each number can range from 0 to 255. For example, 1.160.10.240 could be an IP address. However, it is difficult to remember IP addresses, so domain names are used.

NOTES

A **domain name** is a unique name that identifies a particular Website and represents the name of the server where the Web pages reside. They are the human-readable addresses and hence are easier to remember. Some examples of domain names are *www.yahoo.com* and *www.microsoft.com*. Domain names have two or more parts separated by dots. The letters occurring after the second 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. Three-letter abbreviations are used for such organizations. Some of the abbreviations reserved for various organizations are listed in Table 4.10.

Table 4.10 Reserved Abbreviation for Organizations Example

Domain Type	Type of Organization	Example
.com	Commercial and profit organization	yahoo.com
.edu	Educational institutions	stanford.edu
.gov	Government agencies	nasa.gov
.org	Non-profit organizations	who.org
.mil	United States military	army.mil
.net	Networking organizations	nsf.net

In addition to the three-letter abbreviations, two-letter abbreviations are also used. The two-letter abbreviations are reserved for countries. Some of the abbreviations reserved for countries are listed in Table 4.11.

Table 4.11 Reserved Abbreviations for Countries

Domain Type	Country Name
.in	India
.au	Australia
.ca	Canada
.uk	United Kingdom
.nz	New Zealand
.dk	Denmark
.jp	Japan
.fr	France

The relationship between the IP addresses and the domain names is maintained by an Internet service called the Domain Name System (DNS). This system maps a domain name to an IP address using a set of database servers, allowing you to look up a site by domain name without knowing the IP address of the site. This translation is done by a server called the DNS Server.

4.5.3 Web Browsers

A Web browser is a software application that is used to view and browse information on the Web. With the help of a Web browser, one can easily visit

various Websites and gather information, conduct market research, and stay in touch with the current happenings.

A browser handles requests for HTML files, interprets links, and deals with embedded images, audio, and video elements. A browser supports multimedia data in three ways, which are as follows:

- **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.

Internet Explorer

Internet Explorer is the most widely used Web browser developed by Microsoft. Let us take an example to illustrate how the Web can be accessed using the Internet Explorer.

Open and Navigate a Website

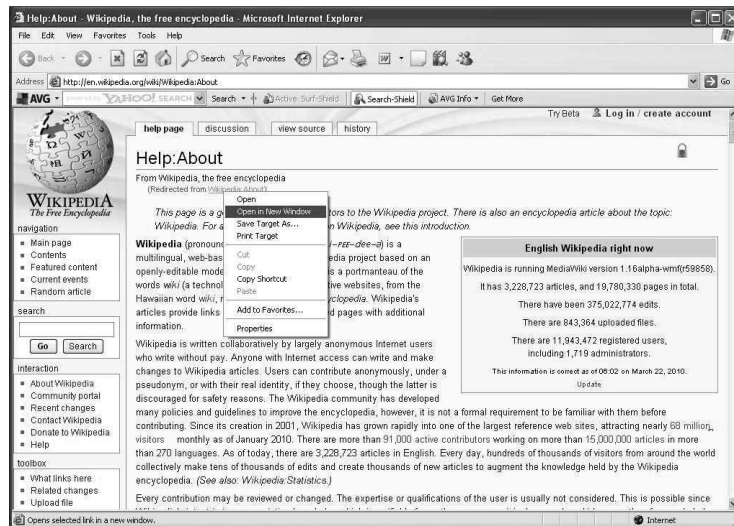
Open Internet Explorer by double-clicking the Internet Explorer shortcut icon placed on the desktop. After opening the Internet Explorer, enter the address of the Website the (URL) in the address bar and press **Enter** to open the Website. The home page of the Website will be displayed. You can move around the home page using various methods which are as follows:

- Use the vertical scroll bar to move around a document.
- Press the **Up** and **Down** arrow keys to move up or down one line.
- Press the **Page Up** key to move up one window length and **Page Down** to move down one window length.
- Press **Ctrl+Home** and **Ctrl+End** to go to the beginning of the document and the end of the document, respectively.

The home page displays various hyperlinks which link to other Web pages or to other parts of the same Web page. To open a hyperlink, just click any of the hyperlinks to open the linked document. Alternatively, right-click on the hyperlink and then click **Open** option from the shortcut menu to open the Web page in the same window or click **Open in New Window** option to open the Web page in a new window.

NOTES

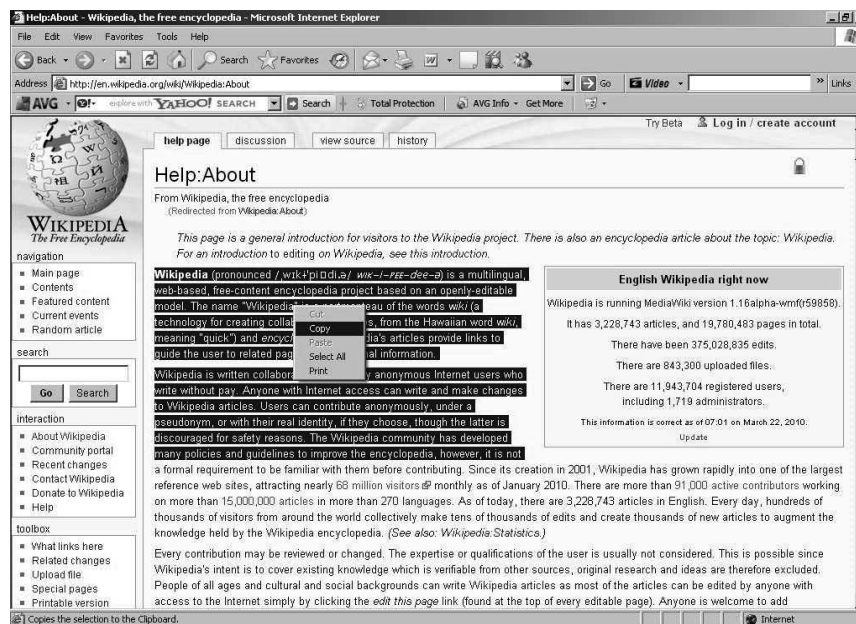
NOTES



After browsing through several pages in a window, you can move backward and forward by clicking the **Back** and **Forward** button on the Standard toolbar.

Save Text, Image and Web Page

Internet Explorer helps not only to find the required information from the Internet but it also allows to save the retrieved information on the local system for future reference. To save a part of the text from a Web page, select the text that has to be saved. Right-click on the selected text and then click **Copy** option from the shortcut menu that appears. Now you can paste the selected text in any of the word processor such as Notepad, Wordpad, MS Word, etc., and save it for future reference.



Similarly, images can also be saved from a Web page. To save the image, right-click the image and then click **Save Picture As** option from the shortcut menu. This displays the Save Picture dialog box. Select the location for the image from the **Save in** drop-down list box and enter a name for the image in the **File**

name box. Click **Save** to save the image at the desired location.

If you want to save the entire Web page, click the **File** menu and then click **Save As** option to display the Save Web Page dialog box. Select the location for the Web page from the **Save in** drop-down list box and enter a name for the Web page in the **File name** box. Click **Save** to save the Web page.

Add to Favorites

Internet Explorer provides a feature called **Favorites**. This feature maintains a list called the Favorites list, of all the Web pages that you visited and want to visit again. To add a Web page to the Favorites list, follow the given steps:

1. Open the Web page that you want to add to your favorites. Click **Add to Favorites** option on the **Favorites** menu. The Add Favorite dialog box appears.
2. Enter a different name for the Web page, if required, in the **Name** box.
3. Select the **Make available offline** check box to view the page even when you are not connected to the Internet and click **OK**.



After the Website is added to the Favorites list, you can view it by clicking on the desired link on the **Favorites** menu.

4.5.4 Internet Service Provider (ISP)

Internet Service Provider (ISP) is a company that access Internet services. This service provider provides a software package in which you get registration with the providing services. Once you registered with username, password and dialing phone number, you can access ISP by paying the monthly fee. This software package is equipped with modem that is connected with Internet services. Good ISPs have their own leased line provided by telecommunication providers. Some of the largest and popular ISPs are AT&T WorldNet, MCI, IBM Global Network, UUNet, PSINet, NetCom, etc. It is sometimes known as the Internet access provider. There are 183 ISPs in India. The Table 4.12 shows the list of ISPs having all India license:

NOTES

Table 4.12 ISPs in India

NOTES

BSNL	RPGInfotech	Gateway systems	RailTel Corporation	i2i Enterprise
CMC	Sifi	ERNET India	GTL	Tata Power Broadband
Essel	VSNL	Jumpp India	Bharti Infotel	RailTel Corporation
Astro India Network	Primus Telecommunication India	Siti Cable Network	World Phone Internet Services	Escorts Communication
Reliance	L&T Finance	In2Cable (India) Reliance	Spectra Net Reach	Estel Communication

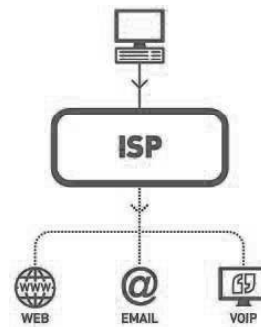


Fig. 4.26 Services of ISP

In the Figure 4.26, ISP provides Web, E-mail and VoIP, etc. as main services. ISP includes domain name registration and hosting, Internet transmit, dial-up or DSL access, lease line and collocation. You can take your domain name, secured Website and high-availability Web servers with this facility. Suppose a firewall is implemented with two separate Ethernet interfaces. Figure 4.27 shows how two ISPs are connected with the Internet.

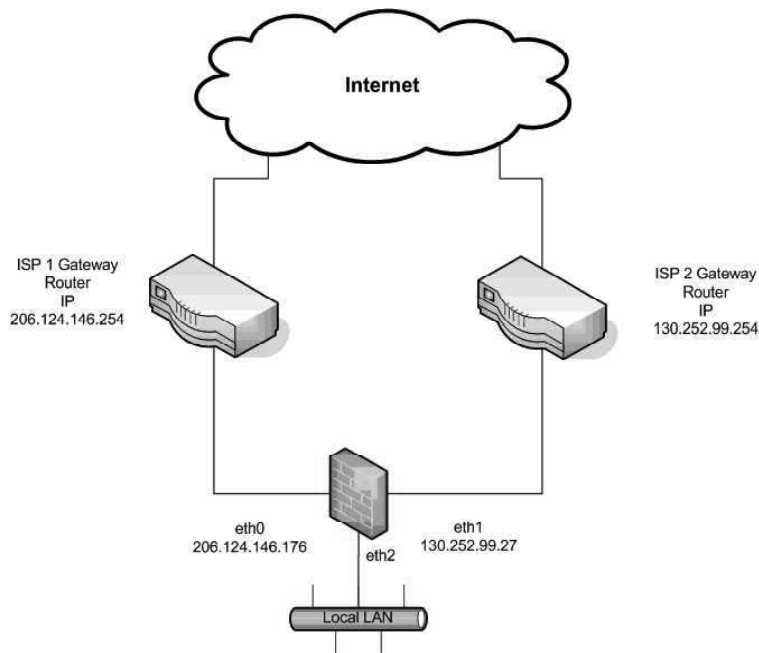


Fig. 4.27 Two ISPs Connected with the Internet

In the Figure 4.27, the following explanation can be analysed as follows:

- The Ethernet eth0 connects to ISP1. The IP address of eth0 is 206.124.146.176 and ISP's gateway router has address as 206.124.146.254.
- The Ethernet eth1 connects to ISP2. The IP address of eth1 is 130.252.99.27 and ISP's gateway router has address as 130.252.99.254.
- The Ethernet eth2 connects to local LAN.

NOTES

Function of ISP

Commercial ISPs easily access and communicate with individual or various organizations across net. They are facilities-based carriers, for example, telephone and cable companies. The interconnected routers are assembled with ISP known as Autonomous System (AS). ISP operates AS to information providers via Google and Yahoo search engines. They exchange traffic networking from other network. This process is called peering. The networks are connected to Internet eXchange (IX).

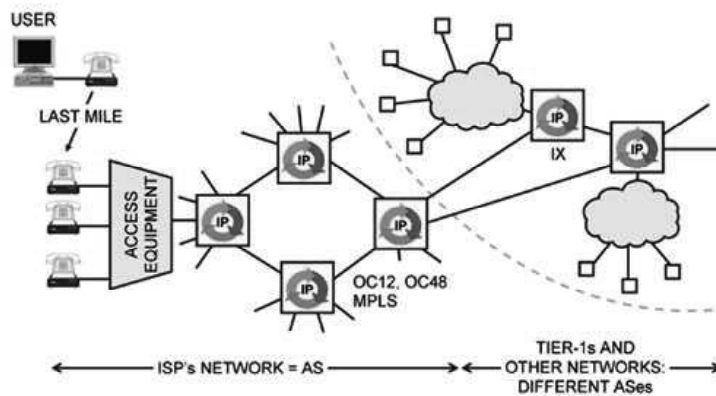


Fig. 4.28 ISP Network

In the Figure 4.24, ISP interconnects with IX providing Tier-1 and other networks. The Tier-1 network provides the largest service with reference to ISP. Peering is settlement free therefore no money transaction is done between ISP and commercial business houses.

4.5.5 Internet Services

The Internet is known as 'the Network of Networks'. It is like a phone system that connects almost anywhere around the world. It exchanges information and acts as global link between small regional networks. Internet services offer a gateway to a myriad of online databases, library catalogues and collections, and software and document archives, in addition to frequently used store-and-forward services, such as UserNet News and e-mail. The widely used Internet services are as follows:

E-mail

E-mail is the prime Internet service that facilitates services to people or users across the world. Full Internet connectivity is not required for this. For example,

an electronic address provides these services to FTP sites through which mail can be exchanged. Other Internet services, such as IP address resolver, Archie lookup, WHOIS service is done via e-mail.

NOTES

The header and body of the message make an e-mail message. The header contains the information where the message is to be sent and the complete path for reaching the destination, date and return path. The body of the message is the actual message that has to be sent. The syntax of an e-mail address is user@subdomain.subdomain.domain, e.g., **abc@gmail.com**. A service provider must be connected with leased line, dial-up or connection with any network for sending e-mail.

File Transfer Protocol (FTP)

FTP is also prime Internet service that acts as protocol and transfers files over TCP/IP network (Internet, UNIT, etc.). Once HTML page is developed on a local machine for a website, it is first uploaded to the Web server through FTP. Local machine is the machine on which you are initially logged into. It includes functions to log on to the network, gives a list of directories and copies files. FTP transfer is possible by entering URL preceded with ftp:// within address bar of a web browser. The FTP operations can be performed by issuing FTP commands at the command prompt or by using FTP utility running under a graphical user interface on Windows OS. FTP tasks can be performed through a browser. For example, type an IE address bar URL as ftp:// to get ftp services. For example, ftp://YourLoginName@IPaddress

The required steps used in connecting with FTP operations are as follows:

The local machine is connected with remote machine by typing 'ftp machinename'. The machinename is the full name, written as aaa.cs.state.edu, of the remote machine to which the local machine is to be connected. Basically, the machine name is the remote machine's full name. If the machine name is not available other option is taken as to type the 'ftp machinenumber' that demands the Net address of the remote machine, e.g., 129.15.0.11. The FTP responds to the users to enter their loginname and password. The anonymous ftp is used widely these days. Many computer systems provide this facility so that you can access the information of specific machine without creating an account on that particular machine. These types of services are provided by anonymous FTP. You need not to be registered user of the system. The anonymous FTP server contains relevant software, documents and files used to configure networks, graphics, images, songs, lyrics and other useful information. An electronic e-mail can be archived through the anonymous FTP. The ready information is stored in machines for any user across the Net who wants to get the required information.

Telnet

Telnet is used to connect remote network computers. It is the Internet service that executes commands on remote host as if you are going to log in locally. For this,

the machine name and valid user name are required to be connected. The commands that are issued on telnet are as follows:

Telnet hostname: A connection to the host name is opened by this command. For example, issuing the command as ‘telnet abc.maths.edu’ with that machine which keeps the required information of abc.maths.edu site can connect you.

Telnet address: It gives the IP address of the connected host.

Archie

If some programs are installed in a system unit and you want to know the availability of the program on the Internet, you can get to know the machine along with such programs via Archie. Basically, Archie is a program that searches files anywhere on the Net by filename. This facility is maintained by a database with the Internet sites accessible via anonymous ftp. The following table shows the various types of Archie servers:

Table 4.13 Various Types of Archie Servers

Name	IP Address
archie.rutgers.edu	128.6.18.15
archie.cs.mcgill.ca	132.206.51.250
archie.funet.fi	128.214.6.102
archie.rediris.es	130.206.1.2
archie.sura.net	192.239.16.13
archie.doc.ic.ac.uk	146.169.16.11

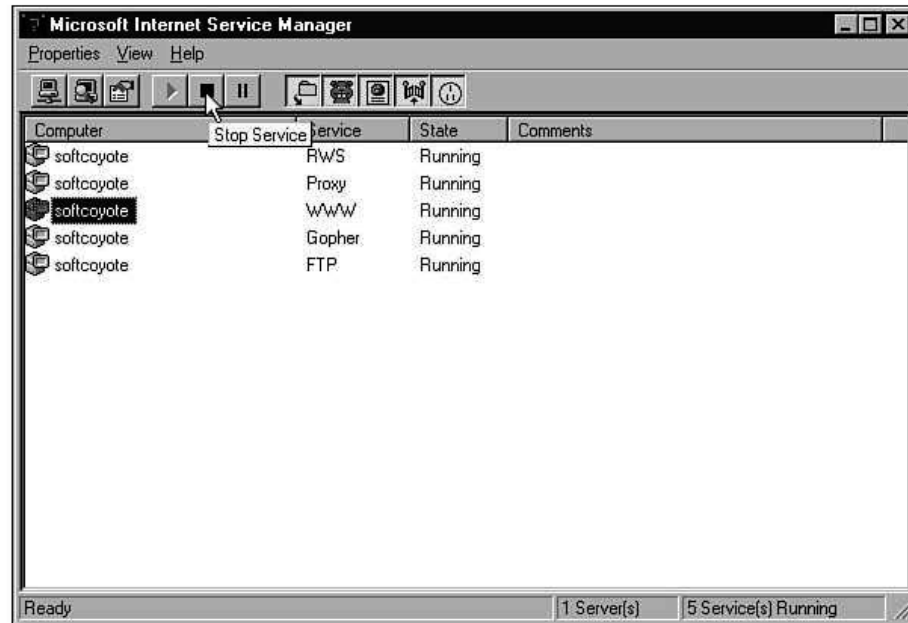
The Archie server can be accessed via Telnet, for example, ‘Telnet archie.rutgers.edu’. For getting Archie server login to ‘Archie’. It requires no password. You can look for files by its full name. For this, either ‘set type exact’ syntax is used or you can use ‘set type sub’ syntax. The ‘set type sub’ syntax is used if the required name of the file is known. The ‘find file-name’ syntax is also used to find the required file name.

Gopher

The Gopher protocol supports client–server software that searches files on the Internet. A Gopher client is required for validating and testing of Gopher publishing service. For example, WS Gopher 4.14 is available on the Internet as shareware. The server based text files are hierarchically organized and viewed by end-users. These end-users access the server by using Gopher applications of remote computers. Gopher browsers initially display the text-based files. Most of the files along with database are available on Gopher that converts HTTP compatible formats and makes them available on the net.

NOTES

NOTES



In the preceding screen, the Internet service manager displays the services that are installed on the server to which the Internet service manager is attached.

Finger

Finger service gives information about users, for example, username, person's first name and last name, information about recently logged in and also where they logged in. But the users must enter the required information where they get registration for particular e-mail services. Finger is also used to get a list of users who are currently logged into the host. In fact, the Finger program accepts input as an e-mail address that returns information of user. In some systems, Finger gives the information about the currently logged on users.

WWW

WWW provides hypertext access to documents located anywhere on the Internet. It is a very successful distributed information system. It is basically client-server data transfer protocol that communicates via application level protocol. Its structural components are clients-browsers, servers and caches. The Internet and semantic components include hypertext transfer protocol (HTTP), hypertext markup language (HTML) extensible markup language (XML) and uniform resource identifiers (URIs). The clients who get various sites requested to the server via HTTP determine the structure of WWW. Then web pages constructs HTML consists of graphics and sound embedded files. For running the complete system, TCP/IP, DNS networking protocols are required.

The reason behind the evolution of Java programming language is to develop distributed application. Distributed application means many CPUs are interconnected through different network topology so that each CPU can communicate with one another. Java introduced the remote method invocation technique to implement distributed application. The **java.net** package provides classes and methods to develop networking-applications through different network protocols.

A group of computers connected by cable to share information is popularly known as network. A network is a set of computers and peripherals that are physically connected. Networking enables sharing of resources and communication. Java applets can be downloaded from a website. This is one of the main attractions of Java. Networking in Java is possible through the use of **java.net** package. The classes within this package encapsulate the socket model developed by Berkeley software division. The network requires some components, such as:

- (a) server
- (b) client
- (c) peer
- (d) protocol
- (e) physical media
- (f) physical devices

Servers provide services to the client. If a server provides application services, then it is treated as an application server.

The **client** accesses services from the server. Peer is a computer that works as a server as well as a client.

Clients

A computer, which requests for some service from another computer, is called a **client**. The one that processes the request is called a **server**. A **server** waits till one of its **clients** makes a request. It can accept multiple connections at a time to the same port number. Multithreading is used to serve multiple users at the same time.

Write the client side application

The following example shows this:

Example 4.1

```
import java.net.*;
import java.io.*;
public class DataClient1
{
    public static DatagramSocket ds;
    public static int clientport=9999,serverport=10000;
    public static void main(String args[])throws
    Exception
    {
        byte buffer[]=new byte[2300];
        ds=new DatagramSocket(clientport);
        System.out.println("Client is waiting for server
        to send data");
        DatagramPacket p=new
        DatagramPacket(buffer,buffer.length);
```

NOTES

NOTES

```
while(true)
{
    ds.receive(p);
    String str=new
String(p.getData(),0,p.getLength());
    System.out.println(str);
}
}
```

Steps for compilation and execution of the client side application

The various steps are as follows:

- (a) Compile client side application: `javac DataClient1.java`
- (b) Execute client side application: `java DataClient`

```
D:\net>javac DataClient1.java
D:\net>java DataClient1
Client is waiting for server to send data
```

Servers

Write the server side application

The following example shows this:

Example 4.2

```
import java.net.*;
import java.io.*;
public class DataServer
{
    public static DatagramSocket ds;
    public static int clientport=9999,serverport=10000;
    public static void main(String args[])throws
Exception
    {
        byte buffer[]=new byte[2300];
        ds=new DatagramSocket(serverport);
        BufferedReader br=new BufferedReader(new
InputStreamReader(System.in));
        System.out.println("Enter Text");
        InetAddress ia=InetAddress.getByName
("localhost");
        while(true)
        {
            String s=br.readLine();
            if((s==null)||s.equalsIgnoreCase("end"))
```

```

        {
            break;
        }
        buffer=s.getBytes();
        d s . s e n d ( n e w
DatagramPacket(buffer,buffer.length,ia,
                clientport));
    }
}
}

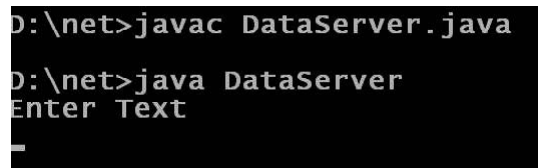
```

NOTES

Steps for writing, compilation and execution of the server side application

The various steps are:

- (a) Write the server side application.
- (b) Compile it, javac DataServer.java.
- (c) Execute it, java DataServer.



```

D:\net>javac DataServer.java
D:\net>java DataServer
Enter Text
_

```

Communication between Server and Client application through multithreading

This can be seen as follows:

Write the server side application

The following example shows this:

Example 4.3

```

import java.awt.*;
import java.awt.event.*;
import java.io.*;
import java.net.*;

public class AppServer extends Frame implements
ActionListener,Runnable
{
    Button b1;
    TextField tf;
    TextArea ta;
    ServerSocket ss;
    Socket s;
    PrintWriter pw;
    BufferedReader br;

```

NOTES

```
Thread th;
public AppServer()
{
    Frame f=new Frame("Server Side Chatting");
    f.setLayout(new FlowLayout());
    f.setBackground(Color.orange);
    b1=new Button("Send");
    b1.setBackground(Color.pink);
    b1.addActionListener(this);
    tf=new TextField(15);
    ta=new TextArea(12,20);
    ta.setBackground(Color.cyan);
    f.addWindowListener(new W1());
    f.add(tf);
    f.add(b1);
    f.add(ta);
    try{
        ss=new ServerSocket(12000);
        s=ss.accept();
        br=new BufferedReader(new
            InputStreamReader(s.getInputStream()));
        pw=new PrintWriter(s.getOutputStream(),true);
    }catch(Exception e)
    {
    }
    th=new Thread(this);
    th.setDaemon(true);
    th.start();
    setFont(new Font("Arial",Font.BOLD,20));
    f.setSize(200,200);
    f.setLocation(300,300);
    f.setVisible(true);
    f.validate();
}
private class W1 extends WindowAdapter
{
    public void windowClosing(WindowEvent we)
    {
        System.exit(0);
    }
}
public void actionPerformed(ActionEvent ae)
```

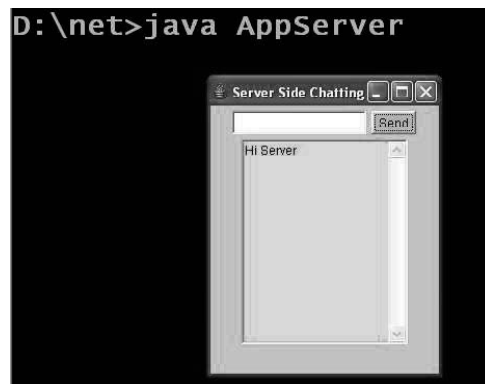
```
{
    pw.println(tf.getText());
    tf.setText("");
}
public void run()
{
    while(true)
    {
        try{
            ta.append(br.readLine() + "\n");
        }catch(Exception e)
        {
        }
    }
}
public static void main(String args[])
{
    AppServer a=new AppServer();
}
}
```

NOTES

Steps for compilation and execution of the server side application

The steps are:

- (a) Compile server side application, javac AppServer.java.
- (b) Execute server side application, java AppServer.



Write the client side application

The following example shows this:

Example 4.4

```
import java.awt.*;
import java.awt.event.*;
import java.io.*;
import java.net.*;
```

NOTES

```
public class AppClient extends Frame implements
ActionListener,Runnable
{
    Button b;
    TextField tf;
    TextArea ta;
    Socket s;
    PrintWriter pw;
    BufferedReader br;
    Thread th;
    public AppClient()
    {
        Frame f=new Frame("Client Side Chatting");
        f.setLayout(new FlowLayout());
        f.setBackground(Color.orange);
        b=new Button("Send") ;
        b.addActionListener(this);
        f.addWindowListener(new W1());
        tf=new TextField(15);
        ta=new TextArea(12,20);
        ta.setBackground(Color.cyan);
        f.add(tf);
        f.add(b);
        f.add(ta);
        try{
            s=new Socket(InetAddress.getLocalHost
(),12000);
            br=new BufferedReader(new
                InputStreamReader(s.getInputStream()));
            pw=new PrintWriter(s.getOutputStream(),
true);
                }catch(Exception e)
            {
            }
            th=new Thread(this);
            th.setDaemon(true);
            th.start();
            setFont(new Font("Arial",Font.BOLD,20));
            f.setSize(200,200);
            f.setVisible(true);
            f.setLocation(100,300);
            f.validate();
        }
    }
    private class W1 extends WindowAdapter
```

```
{
    public void windowClosing(WindowEvent we)
    {
        System.exit(0);
    }
}
public void actionPerformed(ActionEvent ae)
{
    pw.println(tf.getText());
    tf.setText("");
}
public void run()
{
    while(true)
    {
        try{
            ta.append(br.readLine()+"\n");
        }catch(Exception e) {}
    }
}
public static void main(String args[])
{
    AppClient al=new AppClient();
}
}
```

NOTES**Steps for compilation and execution of the client side application**

The steps are:

- (a) Compile the client side application, `javac AppClient.java`.
- (b) Execute the client side application, `java AppClient`.

4.5.6 Internet Requirements

Surfing the Internet is quite similar to scuba diving, with regard to the sophisticated equipments deployed to access Internet. Just as we need certain equipments to dive in the deep ocean for scuba diving, we require adequate paraphernalia to successfully plunge into the huge ocean of interconnected computers and networks.

Now-a-days, Internet access necessitates a broadband connection, which is a high data rate Internet access. The dial-up access deploys a 56K dial-up modem, which uses a dedicated telephone line and is limited to the bit rate of less than 56 Kbps. In contrast to this is the broadband technology, which provides more than double the dial-up bit rate and that too without intervening with the telephone use. In other words, it means that Internet access and voice call can be carried out simultaneously. The broadband connections are characterized by various minimum bandwidths ranging from 64 Kbps up to 2.0 Mbps. Some standards

NOTES

define the broadband connection as having download data transfer rates equal to or faster than 256 Kbps, whereas others define it as having data transmission speed exceeding 768 Kbps in either downstream or upstream direction. In general, any connection of 256 Kbps or greater comes under Broadband Internet.

Certain equipments which are required to access the Internet are as follows. Amongst these, some of them are mandatory and some are optional.

- **Computer:** A computer which is used to browse the Internet may either be a personal computer with Pentium processor or a Macintosh. It should have enough power and memory concomitant with multimedia features. Though 128 MB RAM is sufficient to have access to Internet but 512 MB RAM or more is recommended. Now-a-days, devices like smart phones, mobile phones, Pocket PCs, etc., are also used to browse the Internet.
- **Modem:** It stands for Modulator/Demodulator. This may either be internally built in or externally connected. The modem is a device that converts data in binary code used by the computer, to an analog signal that can be transmitted over the telephone network and vice versa.
- **Internet account with a service provider:** An account with a service provider is essential to create a link between the user's computer and the Internet. A service provider, which is popularly referred to as ISP (Internet Service Provider), signifies phone or cable companies that provide last mile connectivity. It may also refer to a cable line from the subscriber's home to his office and also to an exchange for long distance connectivity based on monthly or annual charges.
- **Widely used current standard broadband technologies:** These technologies are DSLs, that is, Digital Subscriber Line (DSL) and cable modems. However, recent technologies like VDSL and optical fiber connections are also gradually becoming popular in providing Internet access in a much more cost-effective way than copper wire technology. Wi-Fi networks are also used to provide the Internet connections. However, these are not served in the areas by cable or ADSL. The WiMAX has been gaining popularity with regard to mobile and stationary broadband access.
- **Internet browser software:** It is the software tool which enables a user to browse the Internet with the help of Web addresses or URLs. A few of the widely used browsers are Internet Explorer version 7 or 8 (IE), Netscape, Mozilla Firefox, Chrome, AOL, Opera, etc.
- **Anti-virus software:** These are used to protect the user from the onslaught of the nasty programs that obtain access to the user's terminal when he/she is surfing the network or downloading contents from there. Some examples of anti-virus software are Symantec, Norton, McAfee, etc.
- **E-mail software:** The e-mail software may be chosen from the Outlook or Outlook Express. Google, Yahoo and Hotmail offer free Web-mail for the same.

- **Plug-In software:** It is considered an add-on to the user terminal. It enables the user to avail services like music, video, multimedia, etc., on the Internet. The most popular plug-in-software include Real Audio music player, Macromedia Flash Player, Windows Media Player, Apple Quick Time, Java Virtual Machine, etc.
- **Stereo speakers, microphone and webcam:** These equipments enable the user to play sounds, videos, to conduct Internet telephoning and to send images to other users connected to the Internet.

NOTES

4.5.7 Intranet

Intranet is a private and secure internal information system within an organization. It uses Internet technology to collect and disseminate information of interest to all staff (see Figure 4.29).

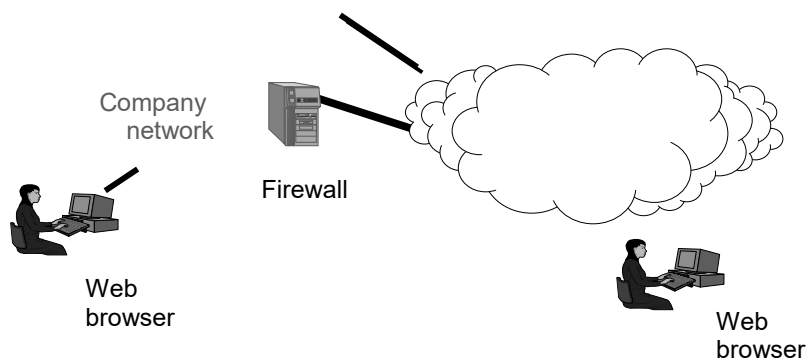


Fig. 4.29 Intranet

The following are the characteristics of the intranet:

- Universal access to staff
- Improved internal communication
- Improved access to information
- Improved productivity
- Promote user participation
- Improved management

Check Your Progress

8. Who created the first computer virus and what name did he give to it?
9. What is Eset Nod 32?
10. Define internet.
11. Define the term uniform resource locator.
12. Name some popular and largest ISPs.
13. Define intranet.

NOTES

4.6 ANSWERS TO 'CHECK YOUR PROGRESS'

1. A database means a place where data can be stored in a structured manner. It is a shared collection of logically related data along with their descriptions designed to satisfy the information needs of an organization.
2. A data dictionary can be treated as a special file that stores the information about the organization and usage of data contained in the database. This information is called metadata.
3. DDL compiler takes care of the processing of the schema definitions defined in the DDL. It involves metadata information, such as the name of the files, data items, storage details of each file, mapping information and constraints, etc.
4. A data model has many advantages. It ensures that all data objects provided by the functional team are represented with accuracy and in complete form. It contains enough details to be used by the technical team who build the physical database. It can also be used to communicate information within and across business organizations.
5. Channel noise is the unwanted electrical or electromagnetic energy that carries no data or information but interferes with the information or data.
6. Bit rate is the rate over network speed which is used to detect errors while transmitting data. The most popular method for detecting errors is inserting a parity bit alongside the data bits for a character. Receiving modems detect incorrect bit rate which is also called parity bit.
7. Error detection is the detection of errors caused by noise or other impairments during transmission from the transmitter to the receiver.
8. Elk Cloner was the name of the first virus which was developed by Rich Skrenta.
9. The Nod32 provides various features for fast scanning speeds for virus protection. It is used with firewall protection.
10. Internet is a 'network of networks' that links computer systems around the world.
11. Each Web page has a unique address called the Uniform Resource Locator (URL) that identifies its location on the Internet.
12. Some of the largest and popular ISPs are AT&T WorldNet, MCI, IBM Global Network, UUNet, PSINet, NetCom, etc. It is sometimes known as the Internet access provider. There are 183 ISPs in India.
13. Intranet is a private and secure internal information system within an organization. It uses internet technology to collect and disseminate information of interest to all staff.

4.7 SUMMARY

- Database means a place where data can be stored in a structured manner. It is a shared collection or batch of data that is logically related, along with

their descriptions designed to meet the information requirements of an organization.

- Database management systems, is a software system that allows users to not only define and create a database but also maintain it and control its access. It can be called a collection of interrelated data (usually called database) and a collection or set of programs that helps in accessing, updating and managing that data (which forms part of a database management system).
- A database system is a collection of application programs that interacts with the database along with DBMS (and sometimes the users who use the system).
- A logical view of data expresses the way a user thinks about data. Usually, it is expressed in terms of a data model.
- A data dictionary can be treated as a special file, which stores the information about the organization and usage of data contained in the database.
- DDL compiler takes care of the processing of the schema definitions defined in the DDL. It involves metadata information, such as the name of the files, data items, storage details of each file, mapping information and constraints, etc.
- Database administration (DBA) installs, configures, troubleshoots and maintains a database system and also separates the function of resource management from database services and technologies, for managing data.
- The relational data model represents a database as a collection of relation values or relations where a relation resembles a two-dimensional table of values presented as rows and columns.
- The network data model can be defined as a database model used to represent objects and the relationships among these objects. In this model, a record can have any number of parent records and it can also have multiple child records.
- Channel noise is the unwanted electrical or electromagnetic energy that carries no data or information but interferes with the information or data.
- Bit rate is the rate over network speed which is used to detect errors while transmitting data. The most popular method for detecting errors is inserting a parity bit alongside the data bits for a character. Receiving modems detect incorrect bit rate which is also called parity bit.
- In computer science, the error detection and correction or error controls are techniques that allow reliable delivery of digital data over unreliable communication channels.
- A computer virus refers to an executable code that is recursively attached to various other computer programs to replicate itself without the awareness of a system user.
- On the basis of their function, the primary types of viruses are: Trojan horse, email virus, macro virus, worm virus, etc.

NOTES

NOTES

- There are two methods of cleaning a virus. These are: the static detection method and the dynamic detection method.
- The antivirus software available in the market to deal with virus-related issues are: Symantec antivirus, Windows Vista Antivirus, Spyware antivirus, Norton Antivirus, Avast, Kaspersky, etc.
- Internet is a 'network of networks' that links computer systems around the world.
- Each Web page has a unique address called the Uniform Resource Locator (URL) that identifies its location on the Internet.
- Internet Service Provider (ISP) is a company that access Internet services. This service provider provides a software package in which you get registration with the providing services. Once you registered with username, password and dialing phone number, you can access ISP by paying the monthly fee.
- Surfing the Internet is quite similar to scuba diving, with regard to the sophisticated equipments deployed to access Internet. Just as we need certain equipments to dive in the deep ocean for scuba diving, we require adequate paraphernalia to successfully plunge into the huge ocean of interconnected computers and networks.
- Intranet is a private and secure internal information system within an organization. It uses internet technology to collect and disseminate information of interest to all staff.

4.8 KEY TERMS

- **Data Item:** It is the smallest unit of data that has meaning to its users.
- **File:** It refers to a collection of all occurrences (instances) of similar type of records, i.e., a group of records of a single type.
- **Database:** It refers to a shared collection of logically related data along with their descriptions designed to satisfy the information needs of an organization.
- **Database Management System:** Also called DBMS, it is a collection of interrelated data (usually called database) and a set of programs to access, update and manage that data (which form part of a database management system).
- **Data Models:** A collection of various concepts that are used to describe the structure of a database.
- **Domain:** It refers to a set of defined atomic values for an attribute.
- **Network Model:** It is a database model used to represent objects and the relationships among these objects.
- **Parallel Transmission:** It is one or more bytes of data are sent over two or more wires.

- **Serial Transmission:** The data is sent over a single wire.
- **Virus:** An executable code that is recursively attached to various other computer programs to replicate itself without the awareness of a system user.
- **Internet:** It is a 'network of networks' that links computer systems around the world.
- **Client:** A computer that requests for some service from another computer is called a client.

NOTES

4.9 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Why do we need a database?
2. What are the advantages and disadvantages of using a database?
3. Why do we store data in a database instead of a file?
4. Write the functions of DBMS.
5. What are the components of the relational data model?
6. Define the term data communication system.
7. Define a computer virus.
8. What are the symptoms of a virus infection on the system?
9. State about the net surfing.
10. Elaborate on the term domain name system.
11. What is web browser?
12. Write the equipments required for an internet connection.

Long-Answer Questions

1. Describe the approach of handling data in the early file-based systems. Discuss the disadvantages of this approach.
2. Describe any five components of a DBMS environment and discuss how they relate to each other.
3. Briefly explain the data administration giving appropriate examples.
4. Discuss about the relational and hierarchical model with the help of examples.
5. Explain the data transmission system and its modes with the help of examples.
6. Describe briefly communication parameters giving appropriate examples.
7. Explain the concept of error detection and correction in data transmission with the help of examples.
8. Describe the computer virus and its types.

NOTES

9. Briefly explain the dynamic detection method.
10. Describe the various types of data attacks.
11. Discuss briefly various types of connections and internet protocols.
12. Discuss how two ISPs can be connected to the Internet with the help of diagram.
13. Briefly explain the internet services with the help of appropriate examples.

4.10 FURTHER READING

Navathe, S.B and R. Elmasri. *Database System Concepts*, 3rd edition. New York: McGraw-Hill, 1997.

Date, C.J. *An Introduction to Database System*, 7th edition. Boston: Addison Wesley, 2000.

Stallings, William. *Data and Computer Communications*. New Jersey: Prentice-Hall 1996.

Tanenbaum, Andrew S. *Computer Networks*. New Delhi: Prentice-Hall of India 2000.

Agrawal, Rajneesh and Bharat Bhushan Tiwari. *Data Communication and Computer Networking*. New Delhi: Vikas Publishing House Pvt. Ltd 2005.

Mano, M. Morris. *Computer System Architecture*. New Delhi: Prentice-Hall of India 1992.

UNIT 5 WINDOWS, MS OFFICE AND OUTLOOK 2010

NOTES

Structure

- 5.0 Introduction
- 5.1 Objectives
- 5.2 Windows 2010
 - 5.2.1 New Ways to Communicate (Online Services and Functionality)
- 5.3 Microsoft Word 2010
 - 5.3.1 Editing of Word Document
 - 5.3.2 Formatting Paragraph and Text in MS Word
 - 5.3.3 Working with Tables, Columns and Sections
 - 5.3.4 Bullets and Numbering
 - 5.3.5 Shortcut Keys for Various Tasks
- 5.4 Microsoft Excel 2010
 - 5.4.1 Creating, Opening and Working with Excel Documents
 - 5.4.2 Working with Cells
 - 5.4.3 Sorting and Filtering Information
 - 5.4.4 Working with Arithmetic Operators and Text Functions
 - 5.4.5 Working with Charts
 - 5.4.6 Modifying and Merging Rows and Columns
- 5.5 Microsoft Access 2010
 - 5.5.1 Features of MS Access 2010
 - 5.5.2 Creating a Table
 - 5.5.3 Form Design in MS Access 2010
 - 5.5.4 Query Design and Running in MS Access 2010
 - 5.5.5 Report Design in MS Access 2010
 - 5.5.6 Using MS Access 2010 for Business Decisions
- 5.6 Microsoft PowerPoint 2010
 - 5.6.1 Creating Presentation and Slides
 - 5.6.2 Applying Animation and Transition Effects
 - 5.6.3 Adding Sound and Video in Slides
 - 5.6.4 Working with Tables and Pictures
 - 5.6.5 Working with SmartArt
- 5.7 Microsoft Outlook 2010
 - 5.7.1 MS Outlook and E-mail Accounts
 - 5.7.2 Importing and Exporting Data
 - 5.7.3 Managing E-Mail in Outlook
 - 5.7.4 Managing E-mail Accounts
- 5.8 Answers to 'Check Your Progress'
- 5.9 Summary
- 5.10 Key Terms
- 5.11 Self-Assessment Questions and Exercises
- 5.12 Further Reading

5.0 INTRODUCTION

Windows 10 is the most recent version of the operating system from Microsoft. Officially it was released in 2015 and was initially offered free of charge to legitimate users of Windows 7 and Windows 8.1. This new version combines features from

NOTES

those two previous installation to suit the users in a better way for both desktop/laptop computers as well as mobile devices. The most notable change in Windows 10 is that Microsoft replaced the Start screen tiles from Windows 8, and brought back the Start Menu. They also removed the vertical toolbars (or “charms”) that appeared from the sides of the screen. These changes make this Windows version easier to use for users of both desktop/laptops and mobile devices.

Microsoft Word (MS Word or Word) is a program which is used for creating documents. It is used by the organization or standalone user to create documents, reports and notes. Word provides the facilities of creating reports and pictures and thus, manifests your idea on the paper. If you have worked in MS Word earlier, you will notice that a lot of new features are added in MS Word 2010. If you are a beginner, there is no need to worry about learning MS Word. This unit focuses on every aspect of this software from beginning to the expert level.

Microsoft PowerPoint (MS PowerPoint or PowerPoint) is a program which is used for creating presentations. It is used by the working professionals, students and politicians for creating attractive presentations. It consists of information in the form of text and various visual aids such as graphics, pictures and video, which keeps the audience interested. This way of presenting information and sharing of ideas increases visual literacy among users. This unit emphasizes on each and every aspect of creating and designing presentations.

In this unit, you will study about the Window 2010, new ways to communicate (online services and functionality), file systems, Microsoft excel 2010, working with toolbars, formatting tools for Microsoft excel, Microsoft access 2010, features of access, Creating a table and relationship, types of function, queries, forms and reports, Microsoft PowerPoint 2010, Microsoft outlook 2010, multiple mail and new accounts, formatting message text and secure messages, reading messages and organizing mail messages.

5.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the Windows 10 operating system
- Discuss about the Microsoft edge
- Define and discuss the features of Microsoft excel 2010
- Explain the significance of Microsoft word 2010
- Describe the working with toolbars
- Elaborate on the formatting tools for Microsoft excel
- Explain the significance of Microsoft access 2010
- Discuss the various features of access
- Explain the table and their relationships
- Describe the various types of function with access 2010
- Discuss the applications of queries, forms and reports

- Understand the basics of Microsoft PowerPoint 2010
- Explain the significance of Microsoft Outlook 2010
- Work with the multiple mail and new accounts
- Discuss about the formatting message text and secure messages
- Elaborate on the reading and organizing mail messages

NOTES

5.2 WINDOWS 2010

Windows 10 is a personal computer operating system developed and released by Microsoft as part of the Windows NT family of operating systems. It was officially unveiled in September 2014 following a brief demo at Build 2014. The first version of the operating system entered a public beta testing process in October, leading up to its consumer release on July 29, 2015.

Windows 10 introduces what Microsoft described as “universal apps”; expanding on Metro-style apps, these apps can be designed to run across multiple Microsoft product families with nearly identical code—including PCs, tablets, smartphones, embedded systems, Xbox One, Surface Hub and Mixed Reality. The Windows user interface was revised to handle transitions between a mouse-oriented interface and a touchscreen-optimized interface based on available input devices—particularly on 2-in-1 PCs; both interfaces include an updated Start menu which incorporates elements of Windows 7’s traditional Start menu with the tiles of Windows 8. The first release of Windows 10 also introduces a virtual desktop system, a window and desktop management feature called Task View, the Microsoft Edge web browser, support for fingerprint and face recognition login, new security features for enterprise environments, and DirectX 12 and WDDM 2.0 to improve the operating system’s graphics capabilities for games.



Fig. 5.1 Start Menu and Action Center of Windows 10

Microsoft described Windows 10 as an “operating system as a service” that would receive ongoing updates to its features and functionality, augmented with the ability for enterprise environments to receive non-critical updates at a

NOTES

slower pace, or use long-term support milestones that will only receive critical updates, such as security patches, over their five-year lifespan of mainstream support. Terry Myerson, executive vice president of Microsoft's Windows and Devices Group, argued that the goal of this model was to reduce fragmentation across the Windows platform, as Microsoft aimed to have Windows 10 installed on at least one billion devices in the two to three years following its release.

Windows 10 received mostly positive reviews upon its original release in July 2015; critics praised Microsoft's decision to downplay user-interface mechanics introduced by Windows 8 (including the full screen apps and Start screen) in non-touch environments to provide a desktop-oriented interface in line with previous versions of Windows, although Windows 10's touch-oriented user interface mode was panned for containing regressions upon the touch-oriented interface of Windows 8. Critics also praised the improvements to Windows 10's bundled software over Windows 8.1, Xbox Live integration, as well as the functionality and capabilities of Cortana personal assistant and the replacement of Internet Explorer with Microsoft Edge.

Critics characterized the initial release of Windows 10 in July 2015, as being rushed, citing the incomplete state of some of the operating system's bundled software (such as the Edge web browser), as well as the stability of the operating system itself on launch. Windows 10 was also criticized for limiting how users can control its operation, including limited controls over the installation of updates on the main consumer-oriented edition in comparison to previous versions. Privacy concerns were also voiced by critics and advocates, as the operating system's default settings and certain features require the transmission of user data to Microsoft or its partners. Microsoft has also received criticism for how it has distributed Windows 10 to users of existing versions of Windows, which has included the automatic downloads of installation files to computers, the recurring display of pop-ups advertising the upgrade, and allegations of the installation process being scheduled or initiated automatically without expressed user consent.

Up to August 2016, Windows 10 usage was increasing, with it then plateauing, and with previous versions of Windows declining in their share of total usage as measured by web traffic. The operating system is running on more than 400 million active devices and has an estimated usage share of 27.72% on traditional PCs and 12.53% across all platforms (PC, mobile, tablet, and console).

Release

On June 1, 2015, Microsoft announced that Windows 10 would be released on July 29, 2015. Microsoft began an advertising campaign centering on Windows 10, "Upgrade Your World", on July 20, 2015 with the premiere of television commercials in Australia, Canada, France, Germany, Japan, the United Kingdom, and the United States. The commercials focused on the tagline "A more human way to do", emphasizing new features and technologies supported by Windows 10 that sought to provide a more "personal" experience to users.

System hardware and storage requirements

The basic hardware requirements to install Windows 10 are the same as for Windows 8.1 and Windows 8, and only slightly higher than Windows 7. The 64-bit versions require a CPU that supports certain instructions. Devices with low storage capacity must provide a USB flash drive or SD card with sufficient storage for temporary files during upgrades.

Some pre-built devices may be described as “certified” by Microsoft. Certified tablets must include `Power`, `Volume up`, and `Volume down` keys; `Win` and `Rotation lock` keys are no longer required.

As with Windows 8, all certified devices must ship with UEFI Secure Boot enabled by default. Unlike Windows 8, OEMs are no longer required to make Secure Boot settings user-configurable, meaning that devices may optionally be locked to run only Microsoft-signed operating systems. A supported infrared-illuminated camera is required for Windows Hello face authentication. Device Guard requires a UEFI system with no third-party certificates loaded, and CPU virtualization extensions (including SLAT and IOMMU) enabled in firmware.

Table 5.1 Hardware Requirements for Windows 10

Component	Minimum	Recommended
Processor	1 GHz clock rate IA-32 or x86-64 architecture with support for PAE, NX and SSE2 x86-64 CPUs must also support CMPXCHG16B, PrefetchW and LAHF/SAHF instructions.	
Memory (RAM)	IA-32 edition: 1 GB x86-64 edition: 2 GB	4 GB
Graphics card	DirectX 9 graphics device WDDM 1.0 or higher driver	WDDM 1.3 or higher driver
Display	800×600 pixels	N/A
Input device	Keyboard and mouse	Multi-touch display
Storage space	IA-32 edition: 16 GB x86-64 edition: 20 GB	N/A

Storage requirements

To reduce the storage footprint of the operating system, Windows 10 automatically compresses system files. The system can reduce the storage footprint of Windows by approximately 1.5 GB for 32-bit systems and 2.6 GB for 64-bit systems. The level of compression used is dependent on a performance assessment performed during installations or by OEMs, which tests how much compression can be used without harming operating system performance. Furthermore, the Refresh and Reset functions use runtime system files instead, making a separate recovery partition redundant, allowing patches and updates to remain installed following the operation, and further reducing the amount of space required for Windows 10 by up to 12 GB. These functions replace the WIMBoot mode introduced on Windows 8.1 Update, which allowed OEMs to configure low-capacity devices with flash-based storage to use Windows system files out of

NOTES

NOTES

the compressed WIM image typically used for installation and recovery. Windows 10 also includes a function in its Settings app that allows users to view a breakdown of how their device's storage capacity is being used by different types of files, and determine whether certain types of files are saved to internal storage or an SD card by default.

New Features and system security

The Windows Runtime app ecosystem was revised into the Universal Windows Platform (UWP). These universal apps are made to run across multiple platforms and device classes, including smartphones, tablets, Xbox One consoles, and other compatible Windows 10 devices. Windows apps share code across platforms, have responsive designs that adapt to the needs of the device and available inputs, can synchronize data between Windows 10 devices (including notifications, credentials, and allowing cross-platform multiplayer for games), and are distributed through a unified Windows Store. Developers can allow "cross-buys", where purchased licenses for an app apply to all of the user's compatible devices, rather than only the one they purchased on (e.g., a user purchasing an app on PC is also entitled to use the smartphone version at no extra cost).

On Windows 10, Windows Store serves as a unified storefront for apps, Groove Music (formerly Xbox Music), and Movies & TV (formerly Xbox Video). Windows 10 also allows web apps and desktop software (using either Win32 or .NET Framework) to be packaged for distribution on the Windows Store. Desktop software distributed through Windows Store is packaged using the App-V system to allow sandboxing.



Fig 5.2 Task View Feature of Windows 10

A new iteration of the Start menu is used on the Windows 10 desktop, with a list of places and other options on the left side, and tiles representing applications on the right. The menu can be resized, and expanded into a full-screen display, which is the default option in Tablet mode. A new virtual desktop system was added. A feature known as Task View displays all open windows and allows users to switch between them, or switch between multiple workspaces. Windows Store apps, which previously could be used only in full screen mode, can now be used in self-contained windows similarly to other programs. Program windows

can now be snapped to quadrants of the screen by dragging them to the corner. When a window is snapped to one side of the screen, Task View appears and the user is prompted to choose a second window to fill the unused side of the screen (called “Snap Assist”). Windows’ system icons were also changed.

Charms have been removed; their functionality in Windows Store apps is accessed from an *App commands* menu on their title bar. In its place is Action Center, which displays notifications and settings toggles. It is accessed by clicking an icon in the notification area, or dragging from the right of the screen. Notifications can be synced between multiple devices. The Settings app (formerly PC Settings) was refreshed and now includes more options that were previously exclusive to the desktop Control Panel.

Windows 10 is designed to adapt its user interface based on the type of device being used and available input methods. It offers two separate user interface modes: a user interface optimized for mouse and keyboard, and a “Tablet mode” designed for touchscreens. Users can toggle between these two modes at any time, and Windows can prompt or automatically switch when certain events occur, such as disabling Tablet mode on a tablet if a keyboard or mouse is plugged in, or when a 2-in-1 PC is switched to its laptop state. In Tablet mode, programs default to a maximized view, and the taskbar contains a back button and hides buttons for opened or pinned programs; Task View is used instead to switch between programs. The full screen Start menu is used in this mode, similarly to Windows 8, but scrolls vertically instead of horizontally.

Removed features

Windows Media Center was discontinued, and is uninstalled when upgrading from a previous version of Windows. Upgraded Windows installations with Media Center will receive the paid app Windows DVD Player free of charge for a limited, but unspecified, time. Microsoft had previously relegated Media Center and integrated DVD playback support to a paid add-on beginning on Windows 8 due to the cost of licensing the required DVD decoders, and the increasing number of PC devices that have no optical drives.

The OneDrive built-in sync client, which was introduced in Windows 8.1, no longer supports offline placeholders for online-only files in Windows 10. Functionality to view offline files is expected to be added in the future by a new Windows app.

Users are no longer able to synchronize Start menu layouts across all devices associated with a Microsoft account. A Microsoft developer justified the change by explaining that a user may have different applications they want to emphasize on each device that they use, rather than use the same configuration across each device. The ability to automatically install a Windows Store app across all devices associated with an account was also removed.

Web browsers can no longer set themselves as a user’s default without further intervention; changing the default web browser must be performed manually by the user from Settings’ “Default apps” page, ostensibly to prevent browser hijacking.

NOTES

NOTES

Parental controls no longer support browsers other than Internet Explorer and Edge, and the ability to control browsing by a whitelist was removed. Also removed were the ability to control local accounts, and the ability to scan a machine for applications to allow and block.

The Food & Drink, Health & Fitness, and Travel apps have been discontinued.

Drivers for floppy drives are no longer integrated and must be downloaded separately.

While all Windows 10 editions include fonts that provide broad language support, some fonts for Middle Eastern and East Asian languages (Arabic, Chinese, Hindi, Japanese, Korean, etc.) are no longer included with the standard installation to reduce storage space used, but are available without charge as optional font packages. When software invokes text in languages other than those for which the system is configured and does not use the Windows font fallback mechanisms designed always to display legible glyphs, Windows displays unsupported characters as a default “not defined” glyph, a square or rectangular box, or a box with a dot, question mark or “x” inside.

Windows Defender could be integrated into File Explorer’s context menu in Windows 8.x, but Microsoft initially removed integration from Windows 10, restoring it in Windows 10 build 10571 in response to user feedback.

User control over Windows Updates was removed (except in enterprise versions). In earlier versions users could opt for updates to be installed automatically, or to be notified so they could update as and when they wished, or not to be notified; and they could choose which updates to install, using information about the updates. Windows 10 Pro and Enterprise users may be configured by an administrator to defer updates, but only for a limited time. For example, in its Canadian licensing agreement, users of Windows 10 “may stop receiving updates on your device by turning off Internet access. If and when you re-connect to the Internet, the software will resume checking for and installing updates. Under the Windows end-user license agreement, users consent to the automatic installation of all updates, features and drivers provided by the service, and to the automatic removal or changes to features being modified is not required; it is implicit “without any additional notice”.

System security

Windows 10 incorporates multi-factor authentication technology based upon standards developed by the FIDO Alliance. The operating system includes improved support for biometric authentication through the *Windows Hello* and *Passport* platforms; devices with supported cameras (requiring infrared illumination, such as Intel RealSense) allow users to log in with iris or face recognition, similarly to Kinect. Devices with supported readers support fingerprint-recognition login. Credentials are stored locally and protected using asymmetric encryption. The Passport platform allows networks, software and websites to authenticate users using either a PIN or biometric login to verify their identity, without sending a password.

The enterprise version of Windows 10 offers additional security features; administrators can set up policies for the automatic encryption of sensitive data, selectively block applications from accessing encrypted data, and enable *Device Guard*—a system which allows administrators to enforce a high security environment by blocking the execution of software that is not digitally signed by a trusted vendor or Microsoft. Device Guard is designed to protect against zero-day exploits, and runs inside a hypervisor so that its operation remains separated from the operating system itself.

NOTES

5.2.1 New Ways to Communicate (Online services and functionality)

Windows 10 introduces a new default web browser, Microsoft Edge. It features a new standards-compliant rendering engine forked from Trident, annotation tools, and offers integration with other Microsoft platforms present within Windows 10. Internet Explorer 11 is maintained on Windows 10 for compatibility purposes, but is deprecated in favour of Edge and will no longer be actively developed.

Windows 10 incorporates Microsoft's intelligent personal assistant, Cortana, which was first introduced with Windows Phone 8.1 in 2014. Cortana replaced Windows' embedded search feature, supporting both text and voice input. Many of its features are a direct carryover from Windows Phone, including integration with Bing, setting reminders, a *Notebook* feature for managing personal information, as well as searching for files, playing music, launching applications and setting reminders or sending emails. Cortana is implemented as a universal search box located alongside the Start and Task View buttons, which can be hidden or condensed to a single button.

Microsoft Family Safety is replaced by Microsoft Family, a parental controls system that applies across Windows platforms and Microsoft online services. Users can create a designated family, and monitor and restrict the actions of users designated as children, such as access to websites, enforcing age ratings on Windows Store purchases, and other restrictions. The service can also send weekly e-mail reports to parents detailing a child's computer usage. Unlike previous versions of Windows, Child accounts in a family must be associated with a Microsoft account—which allows these settings to apply across all Windows 10 devices that a particular child is using.

Windows 10 also offers the Wi-Fi Sense feature originating from Windows Phone 8.1; users can optionally have their device automatically connect to suggested open hotspots, and share their home network's password with contacts (either via Skype, People, or Facebook) so they may automatically connect to the network on a Windows 10 device without needing to manually enter its password. Credentials are stored in an encrypted form on Microsoft servers, and sent to the devices of the selected contacts. Passwords are not viewable by the guest user, and the guest user is not allowed to access other computers or devices on the network. Wi-Fi Sense is not usable on 802.1X-encrypted networks. Adding “_optout” at the end of the SSID will also block the corresponding network from being used for this feature.

Universal calling and messaging apps for Windows 10 are built in as of the November 2015 update: Messaging, Skype Video, and Phone. These offer built-in alternatives to the Skype download and sync with Windows 10 Mobile.

NOTES

Editions and pricing

Windows 10 is available in four main editions for personal computer devices, of which the *Home* and *Pro* versions are sold at retail in most countries, and as pre-loaded software on new computers. *Home* is aimed at home users, while *Pro* is aimed at small businesses and enthusiasts. Each edition of Windows 10 includes all of the capabilities and features of the edition below it, and add additional features oriented towards their market segments; for example, *Pro* adds additional networking and security features such as BitLocker, Device Guard, Windows Update for Business, and the ability to join a domain. The remaining editions, *Enterprise* and *Education*, contain additional features aimed towards business environments, and are only available through volume licensing.

As part of Microsoft's unification strategies, Windows products that are based on Windows 10's common platform but meant for specialized platforms are marketed as editions of the operating system, rather than as separate product lines. An updated version of Microsoft's Windows Phone operating system for smartphones, and also tablets, was branded as *Windows 10 Mobile*. Editions of *Enterprise* and *Mobile* will also be produced for embedded systems, along with *Windows 10 IoT Core*, which is designed specifically for use in small footprint, low-cost devices and Internet of Things (IoT) scenarios and is similar to Windows Embedded.

5.3 MICROSOFT WORD 2010

Microsoft Word is used for the creation of documents. It is the most commonly used format for storing information and sending it through e-mails. There are several interactive features in MS Word 2010 that help in making the documents productive and innovative. Here, we will explore some features such as the management of various files in the word documents, management of the information in a secured way and various shortcuts and tips to access different commands.

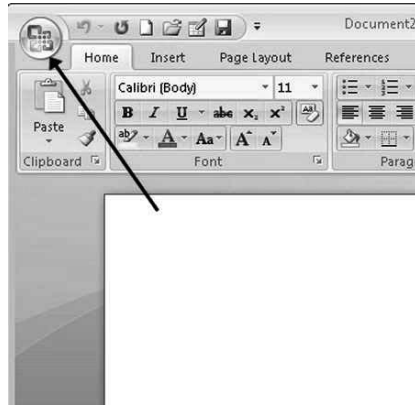
File Management

File management provides the facility to organize the files by saving them in the computer and drives such as disk drive, USB drive, etc. The saved file can then be searched and opened later on. Also, the files can be moved from one location to another as the need be. To better understand the features and functions included in file management, it is essential for you to know about the **File** menu of MS Word 2010.

To make you comfortable with the various aspects of file management, a detailed description of the **File** menu is given below:

File Menu

In the previous version of Microsoft Office there was no **File Menu**. Instead, there was an **Office Button**. This is shown in the following screenshot.



NOTES

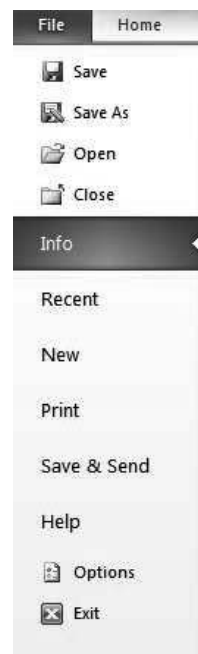
The Office Button has been replaced by the **File Menu** in Microsoft Office 2010. Whenever **File Menu** is clicked, a backstage view of Microsoft Office is visible. In MS Office 2010 **File Menu**, allows users to easily access frequently used commands.

A screenshot displaying **File Menu** is shown below:



When you click on the **File** menu, various options related to the file are displayed.

A screenshot displaying the options is shown below:



Creating a new document

To create a new document in MS Word. Follow the following steps.

1. Go to the file menu in the ribbon interface.
2. Click on New option, you will see the new blank document template.

Inside **File** tab you will find six main sections and six buttons.

The six sections present inside **File** are as follows:

- **Info: Info** section gives important information related to file, by specifying its properties. It enables you to protect documents, check for issues and manage versions.
- **Recent: The Recent** section is by default selected is. This section shows all the **Recent Documents** and **Recent Places**. **Recent Documents** are the word documents which can be opened just by clicking. **Recent Places** are the places where these recent documents are being saved. If a user wants to quickly access recent documents, he should check the option **Quickly access this number of Recent Documents** by selecting the number from a drop down box. All documents will appear on the left hand side. You can recover documents which have not been saved by clicking **Recover Unsaved Documents** button.
- **New: New** section helps you to create new documents. By default, the **Blank Document** is the template selected, however the user can choose any template,.
- **Print: Print** section enables you to print the files with the available settings and a preview of the **Print** section is shown on the right hand side of the window.
- **Save & Send:** There are different options present inside **Save & Send** option. **Send Using E-mail** allows user to send the file with the help of email. In **Save to Web** option you need to login with the help of Windows Live ID. After that, you click on **Save As** option to save the file at the location which you want. The **Save to SharePoint** option is similar to **Save to Web** option. In **Publish as Blog Post option**, user can create blog post by incorporating the present document. This helps to publish blogging contents online. In **Change File Type** option, the user can change the extension of the file type. In **Create PDF/XPS Document** option user can convert word document to PDF or XPS format.
- **Help:** In case of any need, the user can click on **Help** button. The **Help** button also lets you know about updates related to the Office.

The six buttons which are present inside **File** are as follows:

- **Save: Save** button helps in saving the document which you are, currently working upon. Suppose, you are saving the document for the first time then it will ask you the location where you want to save it..

NOTES

- **Save As:** **Save As** button is used when you want to rename the document and save it in another location.
- **Open:** **Open** button is used when you want to open a document which has already been saved.
- **Close:** **Close** button is used when you want to close the documents which you are currently working upon.
- **Options:** **Options** button is used when the user wants to apply various Word options related to the file.
- **Exit:** **Exit** button is used to exit documents. Before you exit, you are asked whether you want to save the document or not.

Saving a document

In MS Word 2010, files are saved in the default format, which is .docx. This is because .docx format is more secured and damages can be easily recovered. Some other formats of saving the files in Microsoft Word are .docm, .dotx and .dotm, which are explained as follows:

.docm file format is used when the document is to be saved in macro-enabled documents.

.dotx file format is used when the document is to be saved in document template.

.dotm file format is used when the document is to be saved in macro-enabled template.

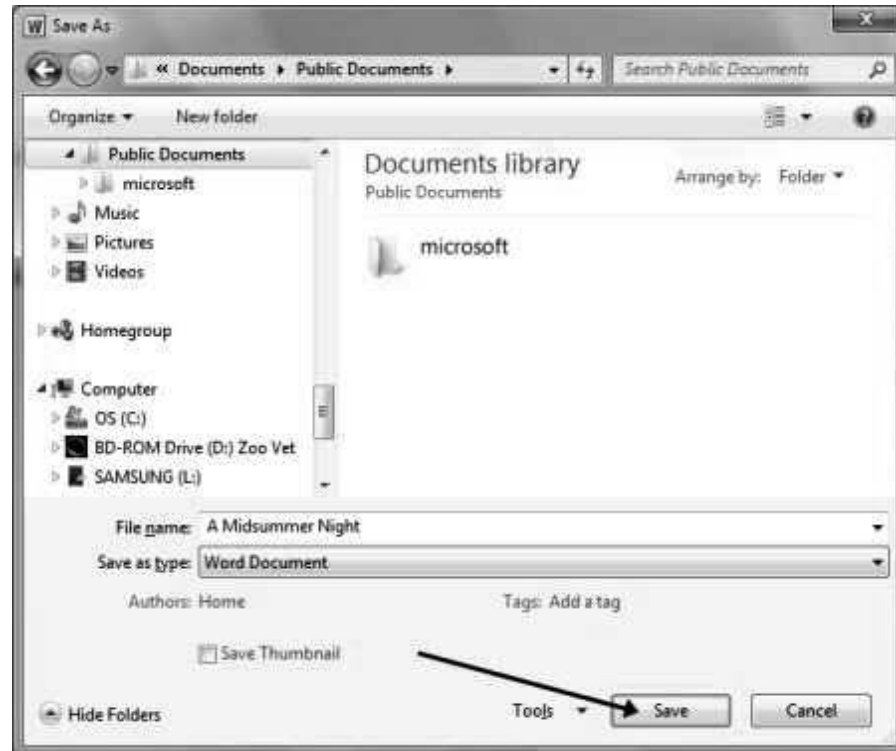
There are many ways of saving the files but these are optimized for the specific uses.

Steps to save the documents are as follows:

1. Click the **File** tab which is present on the left hand side of the ribbon.
2. Click the **Save As** option.
3. Select the drive and folder where you want to save the file.
4. Go to the **File name** textbox to write the name of the file.
5. Select the format in which you want to save the File. By default, the file format in which the file is saved is .docx.
6. Click on **Save** button to save the file.

NOTES

NOTES



This can also be done by clicking the **Save icon** present on the **Quick Access Toolbar**. The keyboard shortcut for saving a word document is Ctrl+S.

Saving a document in PDF or XPS format

There are two methods for saving a document in PDF or XPS format, both are given below:

1. One method of saving the file in Pdf or Xps format is with the help of **Save As** button.
2. Another method of saving the file in Pdf or Xps format is with the help of **Save & Send** section.

Here are the steps for saving a document in PDF or XPS format using **Save As** option:

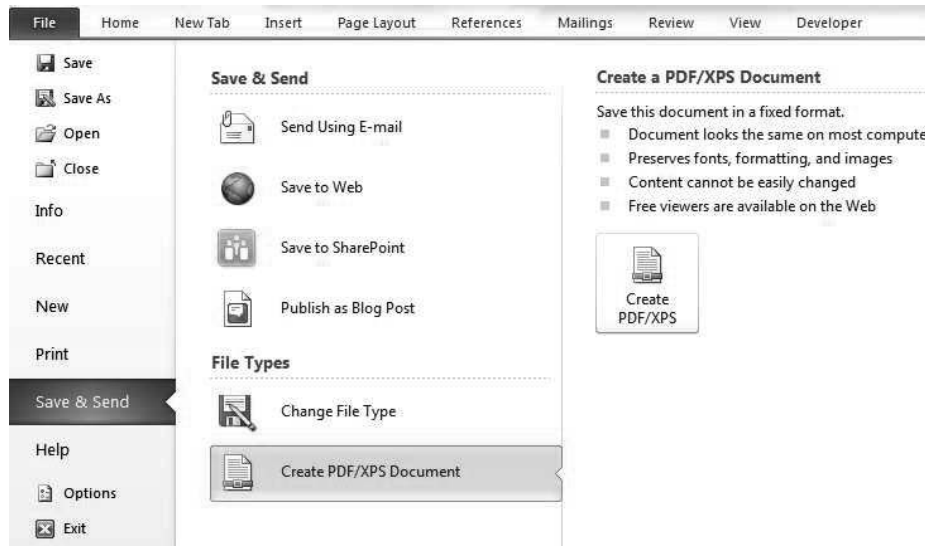
1. Go to the **File** tab which is present in the ribbon.
2. Inside **File** tab, click on the button **Save As**. The **Save As** dialog box appears.
3. Choose PDF extension from **Save as type** which will save the word file in PDF format. In order to save the file as an XPS Document, choose XPS Document extension.

To save a document in PDF or XPS format using **Save & Send** option, follow the following steps:

1. Go to the **File** tab which is present in the ribbon.
2. Inside **File** tab click on the section **Send & Save**.
3. Click on the button **Create PDF/XPS Document**.

The screenshot after clicking on **Create PDF/XPS Document** is given below:

Windows, MS Office and Outlook 2010



NOTES

4. Click on **Create PDF/XPS** button, present on the right hand side of the dialog box. After clicking on **Create PDF/XPS** button, the **Publish as PDF or XPS** dialog box appears.
5. After that Click on **Publish** button to publish the document in PDF format. If the user wants to publish the document in XPS format then user should choose **XPS Document** from the **Save as type** drop box.

Moving file into a specific folder

Steps to move the file into a specific folder are as follows:

1. Click the file which you want to move.
2. Hold the left click of the mouse on the file and drag that file to the specific folder.

5.3.1 Editing of Word Document

Opening the document

Steps to open documents are as follows:

1. Click the **File** tab which is present on the left hand side of the ribbon.
2. Click the option **Open**.
3. Select the location of the file, that is, the appropriate drive and folder from where you want to open the file.
4. Select the file which you want to open.
5. Click on **Open** button to open the file.

Deleting Text

You can delete or remove the text from the document that is no longer required. To do so follow the given steps.

NOTES

- For deleting text to the Left of the insertion point, press the Backspace Key on the keyboard.
- For deleting text to the Right of the insertion point, press the Delete Key on the keyboard.
- For deleting a Range of Text, first select the text that you want to delete or remove and then press the Delete Key on the keyboard.

Copying and Moving Text

You can copy and paste the same text that exists in your opened document to other regions or locations of the same document. Moreover, you can copy and paste the text to another document also. This saves your time. MS Word also allows you to move the selected text from one region of the document to another region as per your requirement. To do this just cut and paste or drag and drop the selected text to the desired location or region.

Copying and Pasting Text

By copying a text you can create duplicate or replica of the selected text. To do this follow the steps given below.

1. Select the text that is to be copied.
2. On the Home Tab, click on the Copy command or simultaneously press the CTRL + C keys on the keyboard. Alternatively, you can right-click on the selected text and then select Copy.
3. Position the insertion point in the text where the text has to be copied.
4. On the Home Tab, click on the Paste command or simultaneously press the CTRL + V keys on the keyboard.
5. The copied and pasted text will appear at the position specified by you.

Cut and Paste Text

1. Select the range of text that you have to cut and move to the new specified position. On the Home Tab, click on the Cut command or simultaneously press the CTRL + X keys on the keyboard. Alternatively, you can right-click on the selected text and then select Cut command.
2. Position the insertion point in the document where the text has to be moved. On the Home Tab, click on the Paste command or simultaneously press the CTRL + V keys on the keyboard. The text will be moved to the specified location in the document.

Find and Replace

In the document let us use the Find and Replace command for changing the word 'Word' to 'MS Word'. To do so, follow the steps given below.

1. On the Home Tab, click on the Replace command.
2. When you click on the Replace command, then the Find and Replace dialog box will appear.

3. In the Find what: field, type the text (in our case 'Word') that you want to find in the document.
4. In the Replace with: field, type the text (in our case 'MS Word') with which you want to replace. Now click on Find Next. Now press 'Replace' to change the word.
5. MS Word will quickly find the first instance of the text 'Word' and highlight it in gray color.
6. Now check your document to **Review** the text and confirm that the word you wanted to replace is done successfully. If there are more than one word to be replaced in the document then select the command '**Replace All**'. All the words in the document will be replaced and an update will be displayed on the screen that how many words are changed and replaced.

NOTES

Printing a new document

After creating the document in MS Word 2010, you can **save, share and print your document** to view it offline. MS Word provides various options for printing the document. You can preview your document before printing using the Print pane so that you can make necessary modifications.

The MS Word provides the very significant feature 'Custom Printing' which helps you in printing only the part text or selected text of the document when you do not want to print the whole or entire document, because by printing unnecessary pages you will be wasting the printing paper. Using the Custom Printing option you can print either several or numerous individual or separate pages or a specified range of pages. Word permits you to specify or identify and select the exact pages from the document for printing.

Printing Document using Custom Print

To print a specific range of pages and individual pages you have to enter the individual page numbers and the range of page numbers. Remember that each entry has to be separated with a comma, for example 2, 5, 8, 15, 20-25, 41-54, and so on.

1. Click on the **File Tab** to open the **Backstage view**. Click on the **option Print**.
2. The **Print pane dialog box** will appear. Click on the **arrow** that appears next to the **field 'Print All Pages'** in the '**Settings**' option. A **menu** will be displayed, select the **option 'Custom Print'**.
3. To do **Custom Printing**, in the **field 'Pages:'** enter the page numbers (individual page numbers or range of page numbers) that has to be printed.
4. Click on '**Print**' to print the preferred pages.

Print Preview

The 'Print Preview' feature of MS Word helps you to view on the screen how the printed version of the document would look like before printing a hard copy. In latest versions of the MS Word (MS Word 2010 Onwards), there is 'NO' Print Preview Tab or Option. As an alternative, click on the **File Tab** then on the **option**

NOTES

Print from the **menu** that appears on the **left side**. The **Print pane** will be displayed. In the **Print pane**, click on the **'Page Setup' option** to open **Page setup dialog box**. In the **Page setup dialog box** define the **Margins, Paper Size** and **Layout options** for **setting the document text** so that the text remains within the printable area.

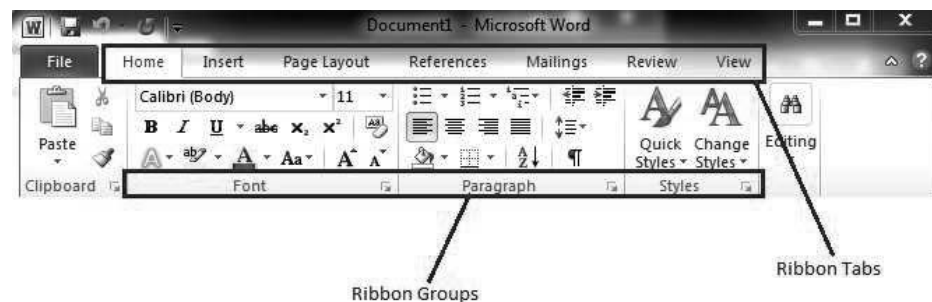
The Ribbon Interface

Ribbon View is a substitute that provides the facilities of accessing the commands for applications. It is organized using a horizontal bar. It provides an easy way to access the commands, as all the commands are organized using tabs and groups. With the help of Ribbon, performing various actions to the Office application is easy. Ribbon is dynamic in nature, which implies that behavior changes as the size of Microsoft Office changes.

'Ribbon' is a set of toolbars which displays the commands and tools for performing various tasks. In Microsoft Office 2003, user worked with the classic drop down menu and toolbar and users were able to work with them easily and quickly. With the release of Microsoft Office 2007, the new term 'Ribbon' was proposed that refers its own implementation of tabbed toolbars bearing heterogeneous controls, known as 'The Fluent User Interface', which swapped the menus and toolbars with a single 'Office Menu'. MS Word, PowerPoint, Access and Excel implemented 'Ribbon' in MS office 2007.

With the release of MS Office 2010, however, 'ribbon' was implemented in the rest of the MS Office applications and added end user customization support for its end user. MS Office 2010 ribbon is great, flexible and easy to use.

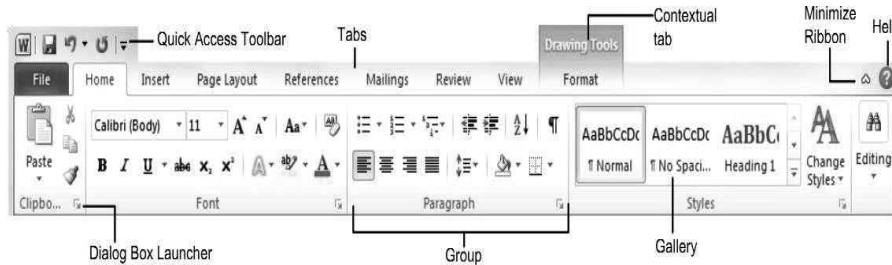
Screenshot for **Ribbon** is as shown:



If you click on any tab or groups, each button and each dropdown menu will perform different actions

For example: When one clicks on the **Home** tab or any other tab, various options are displayed. With the help of these options, one can change the Formatting, Orientation, Layout, Caption, and Proofing of the MS Office Application.

As an example in MS Word 2010, the labeling of the commands and menus is as shown in the screenshot:



NOTES

- **Tab:** It shows different types of commands for creating MS Offices applications. You can see the commands by clicking on the particular tab.
- **Quick Access Toolbar:** It is a way to provide quick access view to the most frequently used commands. By default, **Save**, **Undo** and **Redo** buttons are visible on the **Quick Access Toolbar**.
- **Contextual Tabs:** It displays the Commands for a particular selected object. When the user draws any shape, a Contextual tab called **Drawing Tools** appears.
- **Minimize Ribbon:** You can minimize the ribbon in the following ways:
 - o By clicking on **minimize ribbon** button.
 - o By double clicking the tab on the **ribbon**.
 - o By right clicking the tab from the **contextual menu** and selecting **minimize ribbon** button.
 - o By pressing **Ctrl+F1** button.
- **Help:** For getting any help, you can click on **Help** button or by clicking on the **File** tab and selecting **Help** from the menu.
- **Dialog Box launcher:** You can see the **Dialog box launcher** or the **Task pane** by clicking on the **Clipboard Task Pane**.
- **Group:** Set of all correlated commands which appear in each tab.
- **Gallery:** Gallery contains the option list and additional choices are displayed as thumbnail preview.

5.3.2 Formatting Paragraph and Text in MS Word

Formatting word document is necessary in order to create an effective document. In case of formatting, you format a text, letters, words, pictures, paragraphs, and so on. It can be applied to any portions of word documents. There are four types in which formatting are classified, which are character, document, section and paragraph. There are various tools which let you format the word documents. You can apply font and paragraph formatting; create bullet and numbering to show the information in point. You can also apply various styles to enhance the text. In this section, you will learn to format a word document by applying different formatting effects.

Paragraph and Font Formatting

Font and paragraph formatting is one of the most important styles of formatting. If you do not apply any style, then font and paragraph formatting are applied by default. It is applied to keep the documents simple and appealing.

NOTES












Paragraph Formatting




Paragraph formatting is used to change the appearance of paragraph. It is present in the **Home** tab of ribbon view. You can apply the **Paragraph Formatting** by selecting the paragraph. It contains various controls for providing a quick access to many frequently used functions. All options of paragraph formatting are found in the paragraph group.

The screenshot of **Paragraph Formatting** is as shown:



The available **buttons and icons** of the Paragraph Formatting toolbar are shown:

-  **Bullets** button of the formatting toolbar helps in creating and removing the bullets of the selected paragraph.
-  **Numbering** button of the formatting toolbar helps in creating and removing the numbers of the selected paragraph.
-  **Multilevel list** button of the formatting toolbar provides you the facility of adding numbered lists of items, using multiple levels.
-  **Decrease Indent** button decreases the indentation level of the paragraph.
-  **Increase Indent** button increases the indentation level of the paragraph.
-  **Sort** button helps in sorting the text, number and date to either ascending or descending order.
-  **Show/Hide** shows or hides the formatting symbols of various sections of a document. Arrow symbol denotes tab and dot denotes space. It is set from left to right direction.
-  **Align Text Left** button of the formatting toolbar helps in alignment of text, number, paragraph and object to left.
-  **Align Center** button of the formatting toolbar helps in alignment of text, number, paragraph and object to center.
-  **Align Text Right** button of the formatting toolbar helps in alignment of text, number, paragraph and object to right.
-  **Align Justify** button of the formatting toolbar helps in alignment of text, number, paragraph and object to both left and right margins.

-  **Line and Paragraph spacing** of the dropdown menu helps in changing the amount of space between the paragraphs when selected. You can choose any space by clicking on the dropdown menu.
-  **Shading** dropdown box helps you to add colored background to the area of text which is selected.
-  **Bottom Border** dropdown box helps in selecting the borders from options in the dropdown menu. By default, “Bottom Border” is selected. You can change border to Top Border, Left Border, Right Border etc. You can add and remove the border around any texts, pictures or paragraphs.

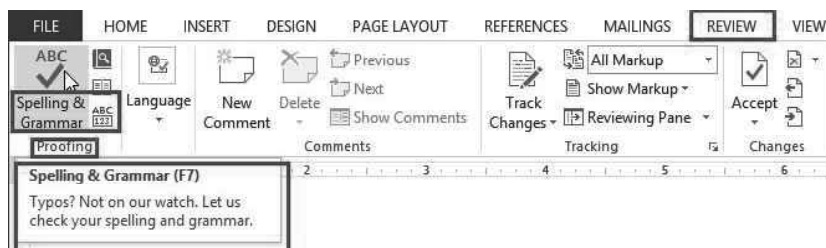
NOTES

Spelling and Grammar Checking

MS Word provides significant feature ‘**Proofing**’ for checking the text in the document by using the option ‘**Spelling & Grammar**’. If you have made mistakes in the document while typing the text then you can use the numerous proofing features provided by MS Word to produce error-free and professional documents. You can check the spellings in the document and after making necessary correction you can change the corrected spelling in the whole document.

Running a Spelling and Grammar Check

1. To start the proofing of the document text, on the **Review Tab** go to the **Proofing group option** and then select the command **Spelling & Grammar**.



2. The **Spelling pane** will be displayed on the right side of the screen. MS Word will provide **one or more suggestions** for every error in the **document text**. Select the correct suggestion to **modify the error** if you want and then click on ‘**Change**’ to **change this specific error** or click on ‘**Change All**’ to **modify the similar error in the whole document**. If you **do not want to modify or change the error** then click on ‘**Ignore**’ for **ignoring this error** or click on ‘**Ignore All**’ for **ignoring the similar errors in the whole document**. Generally, the **error is highlighted** by a **red line** in the document text.
3. MS Word will go through each error in the document till you review all the errors in the current document. When you **review the last error** in the document then the dialog box will be displayed on the screen to confirm that the checking of spelling and grammar is now completed. Click on **OK**.
4. You can also manually modify or correct the spelling errors in the document when no suggestions are provided by MS Word.

NOTES

Ignoring Errors

Do not entirely depend upon the Spelling and Grammar check as it may not be correct always. Essentially while checking errors in grammar, it is possible that MS Word may not notice various grammatical errors. It is also possible that at times the Spelling and Grammar check option may mark any spelling as incorrect while actually it is correct, for example name of people or place which is not included in the MS Word predefined dictionary. Use the following options for checking a spelling or grammatical error.

Checking 'Spelling Errors'

- **Ignore:** When you click on '**Ignore**' option then the MS Word will **skip checking the marked word**, i.e., it will not be changed.
- **Ignore All:** When you click on '**Ignore All**' option then the MS Word will not only **skip checking the marked word**, i.e., it will not be changed, but it will also **skip checking all other occurrences or instances** of the **similar word** in the current MS Word document.
- **Add:** When you click on '**Add**' option then MS Word will **add the marked word** to the **existing predefined dictionary** so that it will **not be marked as error** again. Before clicking on the **Add option** check the **spelling of the word** to ensure that the **marked word** is **spelled** accurately.

Checking 'Grammar Errors'

MS Word provides detailed explanation for each marked 'Grammar Error' that why it is incorrect. It also displays some related example sentences or phrases to explain the concept, as shown in the given Illustration. These examples will help you in determining whether to change the marked sentence or phrase by clicking on the Change option or to ignore it by clicking on the Ignore option.

Change: When you click on the Change option then MS Word will modify or change the marked sentence or phrase in the document with the suggestion that you have selected.

Ignore: When you click on the Ignore option then MS Word will skip checking the word or phrase in the document without making any change or modification to it.

Automatic Spelling and Grammar Checking

By default, MS Word will automatically check the Spelling and Grammar errors in the document. Hence, there is no need to perform a separate checking in the document for spelling and grammar. MS Word indicates these errors by distinct colored and wavy lines.

- If the **word** is **marked** with the **red wavy line** then it indicates that it is a misspelled word, i.e., the spelling of the marked word is incorrect.
- If the **word or sentence** is **marked** with the **blue wavy line** then it indicates that it is a grammatical error, i.e., either the usage of word is grammatically incorrect or the sentence includes misused words.

The misused word is also termed as a '**contextual spelling error**'. This term is used when the spelling of the marked word is correct but it is used incorrectly. For

example, while writing a letter if you **start** with the **phrase ‘Deer Mr./Ms. XYZ’** then the **word ‘Deer’** will be a misused word or a contextual spelling error. The spelling of the word Deer is correct but in the letter it is used at incorrect place. The correct word to be used is ‘Dear’.

Correcting the Spelling Errors

To correct the Spelling Errors, follow the steps given below.

1. Place the insertion point at the word marked as error by MS Word. Now right-click the marked underlined word. The menu with spelling suggestions will appear.
2. Select the spelling that is correct for your document from the list of suggested spellings as shown below in the Illustration.
3. MS Word will change the marked word with the corrected spelling in the document.

Word Count

Total number of words in a Word document is displayed on the status bar at the bottom of the Word window. If it is not displaying then right-click the status bar and then click Word Count.

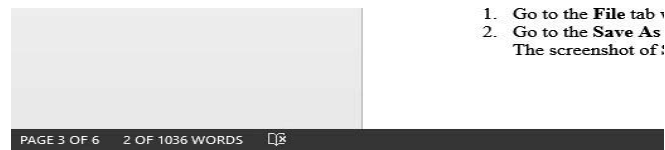


Table of Contents (TOC)

Table of Contents gives a quick reference point or an abstract overview to the reader where the reader can find data or content.

Inserting Table of Contents:

TOC is inserted in the document at the position, where the cursor is placed and it is not inserted from the starting place of the document.

Steps to insert table of content are as follows:

1. Create a blank document.
2. Click on **References** tab from the **Ribbon** view.
3. Click on the **Table of Content** option under **Table of Content** group. A dialog box opens which shows different type of **Table of Content** options.

Table of Content Fields: This is another method of adding fields in Table of Content.

Steps to add entries in Table of Content are as follows:

1. Click on the **Insert** tab.
2. Click on the **Quick Parts** which comes under **Text** group and select the option **Field**.

NOTES

NOTES

3. Then, a **Field** dialog box opens.
4. Select the option **TC**. Then type some text in **Text Entry** text box and click **OK** button.
5. For displaying the entries, click on the **Table of Content** option in the **Reference** tab. Then, click on the option **Insert Table of Contents**.
6. Then, a **Table of Content** dialog box opens.
7. Click on the **Options** in this dialog box, then, a **Table of Content Options** dialog box opens.
8. Uncheck the **Styles** and check the **Table entry fields** options and then click **OK** button.

Updating the Table of Contents: If you want to change the some content in the document then you can update it by using following steps:

1. Click on the **Reference** tab. Select **Table of Contents** under it and choose any one style for creating TOC. An example of **Table of Contents** is as shown:

Contents

9. Updating The Table Of Contents	7
12. Adding Entries To A Table Of Contents	7
14. Removing Entries From A Table Of Contents	8

2. Then click on **Update Table**, a dialog box **Update Table of Contents** will be opened click on the **OK** button of that dialog box.
3. Update the content of table according to your need.

Removing entries from the Table of Content: You can also remove the entries from the Table of Contents.

Steps to remove entries from the Table of Content are as follows:

1. Click on the **Reference** tab.
2. Click on **Add Text** option under the **Table of Contents** group.

5.3.3 Working with Tables, Columns and Sections

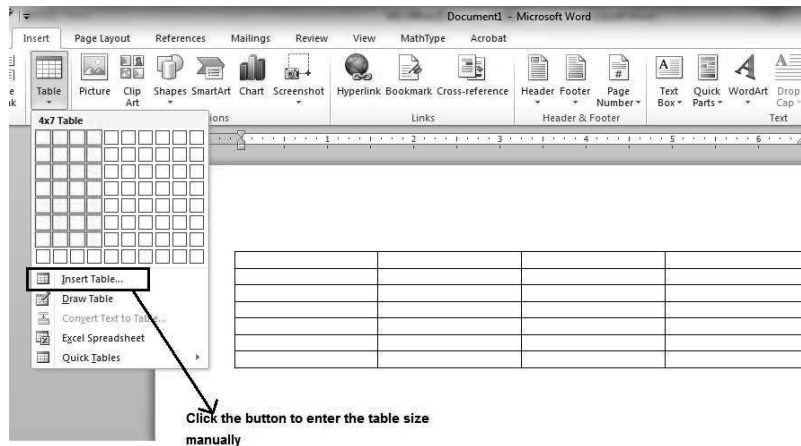
Table

Table consists of the rows and columns. Rows are horizontal and columns are vertical. At the intersection of the rows and column there is a cell which contains the text. We can format the tables according to our requirements.

Creating a table

- Go to the **Insert** tab and click on the **Table** option.
- Move your cursor over the grid, the table will appear in the document. We can add the desired number of rows and columns in the table. Left click on the grid when the size of the table is appropriate.

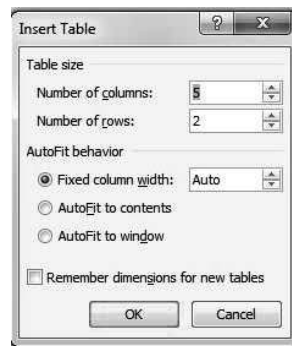
The screenshot of the table created using the grid is as shown:



NOTES

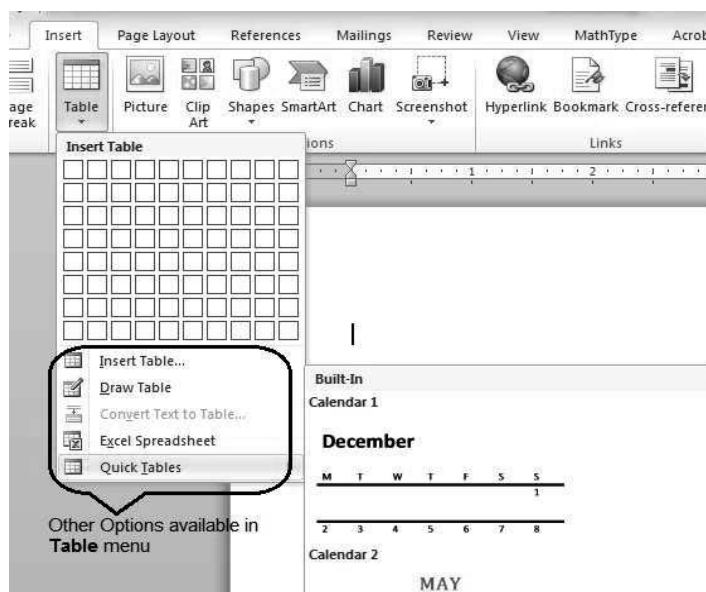
- Sometimes, the required table might be greater in size than that is available in the grid. In that case, click on the **Insert Table** button and enter the size manually.

The screenshot of the table created manually is as shown:



- Below the grid, there are other options that can be used to design the table according to the requirement.

The screenshot showing the other options available for creating a table is as shown:

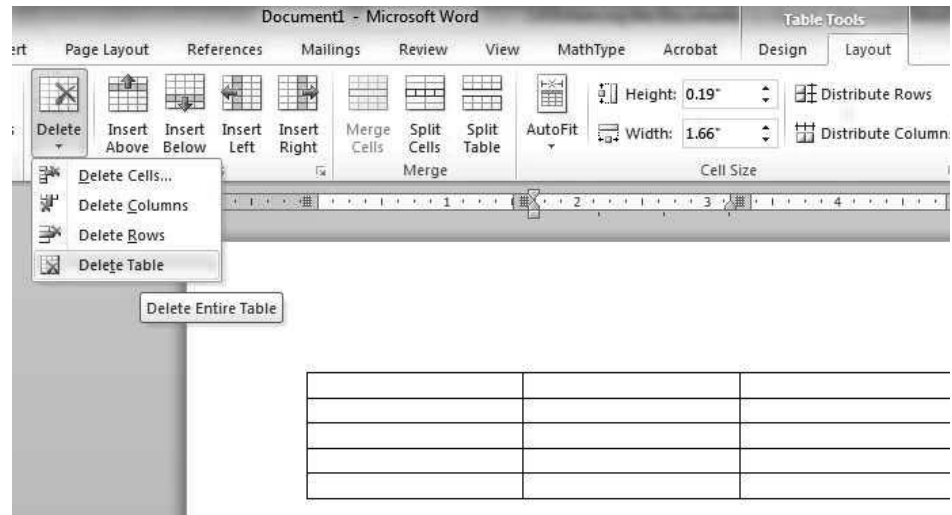


NOTES

Deleting the table

- Click on the table. Table layout will be displayed on the top.
- Click on the **Delete** button.
- Select the **Layout** tab if this is not already selected.
- Click on the **Delete Table** button. This will delete the table. Other actions can also be performed on the table according to the requirement.

The screenshot of **Delete Table** is as shown:

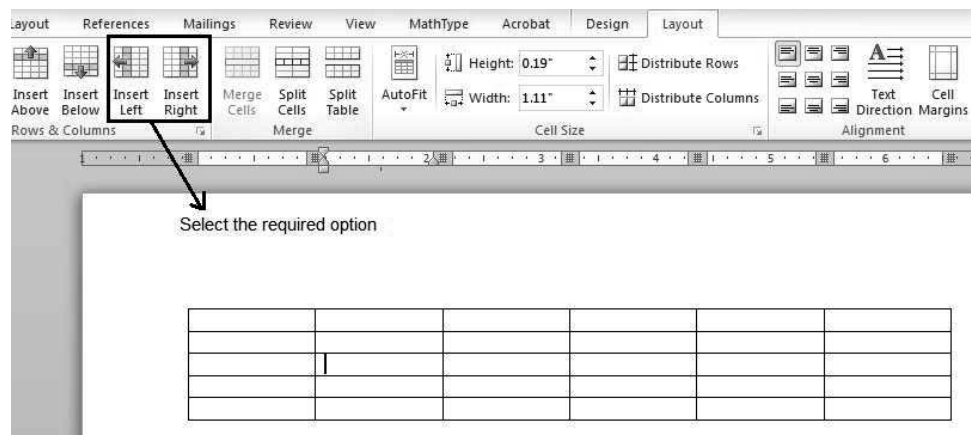


Column

Insert Column

- Click on the table. Go to the **Layout** tab.
- In the **Layout** tab, go to the **Rows and Columns** section.
- Click on **Insert Right** or **Insert Left** for inserting a column on the left and right side of the particular column.

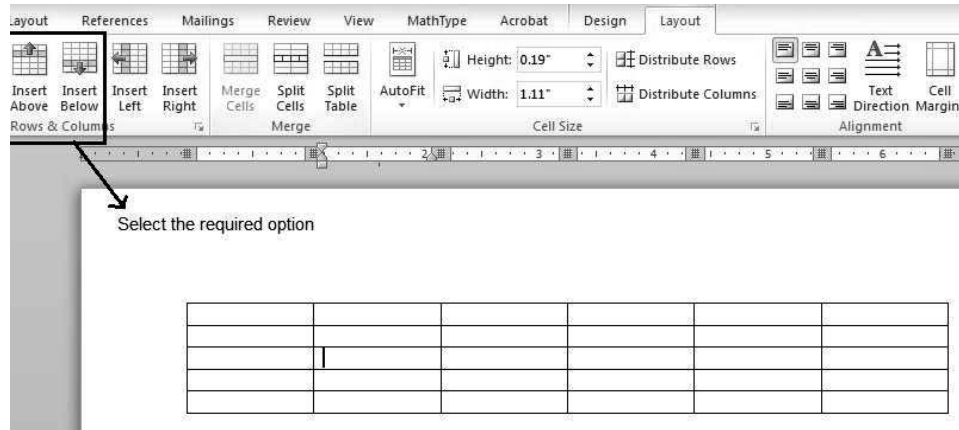
The screenshot of inserting a column on left or right of a selected column in a table is as shown:



Insert Row

- Click on the table. Go to the **Layout** tab.
- In the layout tab, go to the **Rows and Columns** section.
- Click on **Insert Above** or **Insert Below** button for inserting a row above or below the row where the cursor is present.

The screenshot of inserting a row above or below a selected row in a table is as shown:



NOTES

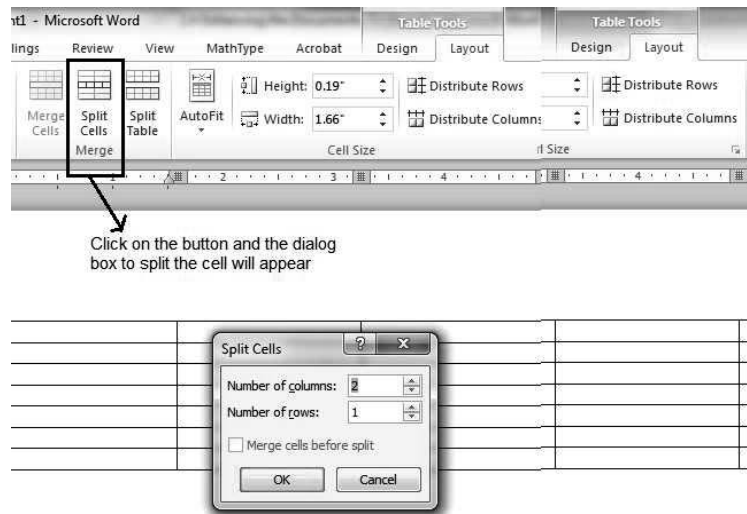
Sections

Section means splitting the cells or table into different parts or sections.

1. Splitting the cells

- Go to the particular cell in the table. Then, go to the layout menu.
- Click on **Split Cells** button. **Split cells** dialog box appears that prompts to enter the number of rows and columns.
- Specify the rows and columns. The cell will be divided into sections.

The screenshot of **Split Cells** is as shown:

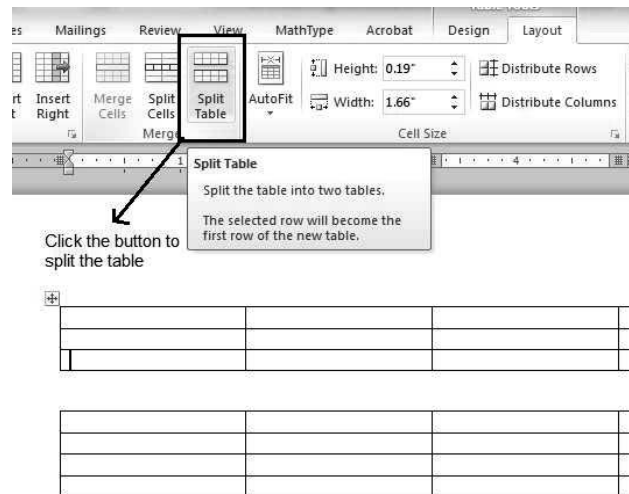


NOTES

2. Splitting the table

- Go to a particular row in the table from where you want to divide the table. Go to the **Layout** bar.
- Click on **Split Table** Button. The table will be divided into sections. The row that was selected will be the first row of the second table of the split tables.

The screenshot of **Split Table** is as shown:



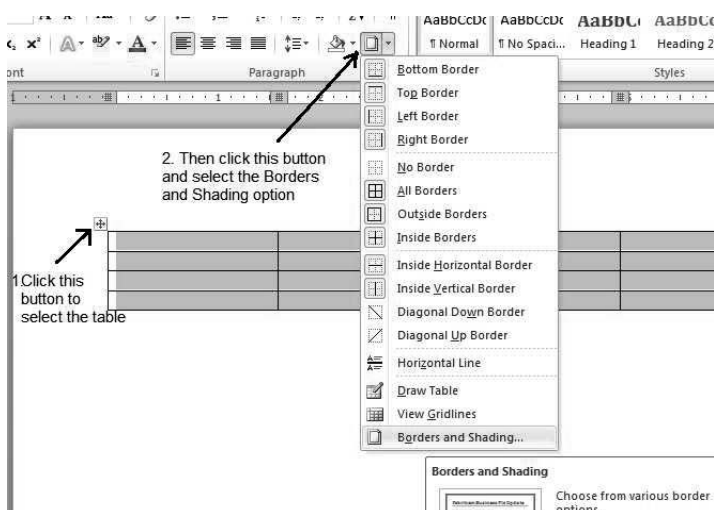
Borders and Shading

With Microsoft, the user can apply border to a table. You can also provide different shades to the table of your choice.

Applying Border to a table

- Select the table to which you want to apply the border.
- Click the **Border** button to select the border from different types of border.
- You can delete the border by selecting **No Border** option.

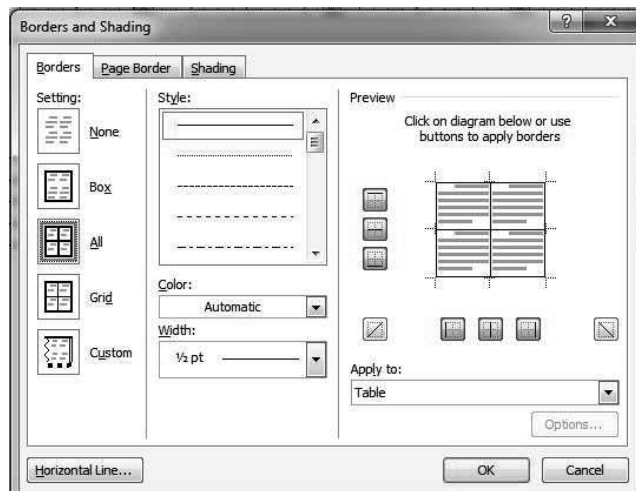
The screenshot of **Border** option is as shown:



Use the 'Border' option

- Select the table. Click the **Border** button, a list will appear. Click on border and shade option.
- A dialog box will appear, select border from that dialog box.
- Choose the border of your choice and color and apply it to the table.

The screenshot of **Borders** options is as shown:

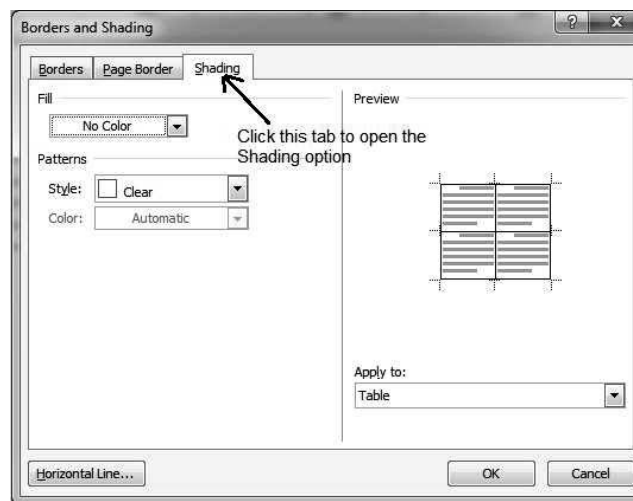


NOTES

Adding shades to the table

- Select the row or column you want to shade. Go to **Border** button.
- Select **Borders and Shading** option from the list. A dialog box will appear. Select **Shading** tab and select the color of your choice.
- Apply the shade to the row or column. The row or column will get shaded.

The screenshot of **Shading** option is as shown:



Page Setup and Break

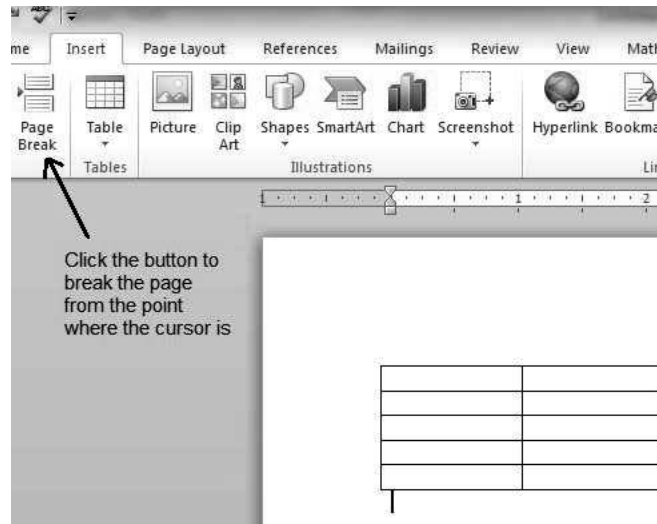
Break: In word, a new page begins after the previous page is completely filled but with the help of break, the user can force the text to start from the new page.

NOTES

Insert page break

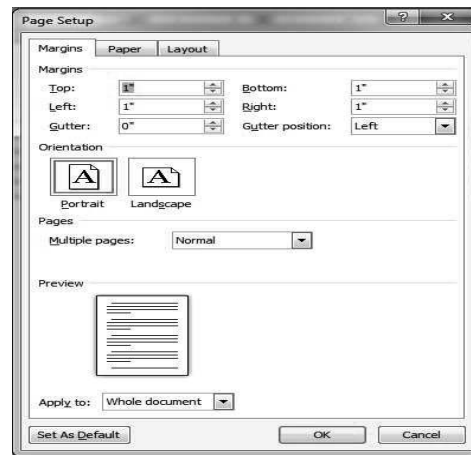
- Open the document. Go to the location from where you want to start a new page.
- Go to the **Insert** tab and click **Page Break**. New page will start from the insertion point.

The screenshot of **Insert page break** is as shown:



Page Setup: This option is present in **Page Layout** tab. It helps in setting the orientation of the page in the document. It also includes options for setting margins on a page and page break. By default, the orientation of a page is selected as Portrait. You can also choose landscape mode. Let's learn about **margin**.

The screenshot of **Page Setup** dialog box is as shown:

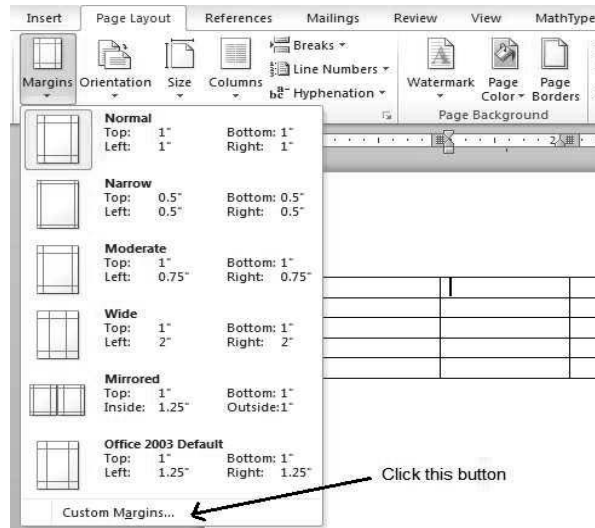


Margin

- Open the document; go to **Page Layout** option and click on **Margin** tab. A list will appear. Click on **Custom Margin** option available.
- A dialog box will appear that specifies the margins you want to keep for the document.

The screenshot of selecting the **Margins** option is as shown:

Windows, MS Office and Outlook 2010

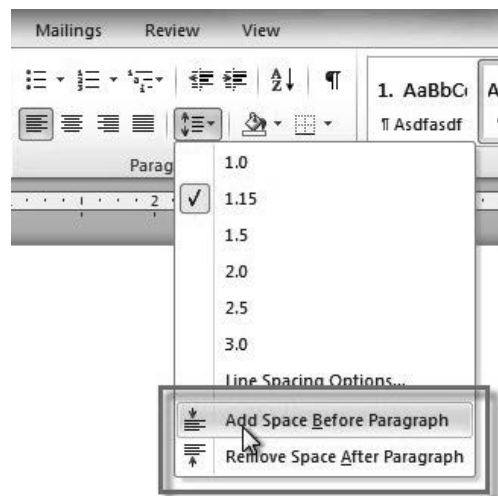


NOTES

Line spacing

To format line spacing, follow the steps given below.

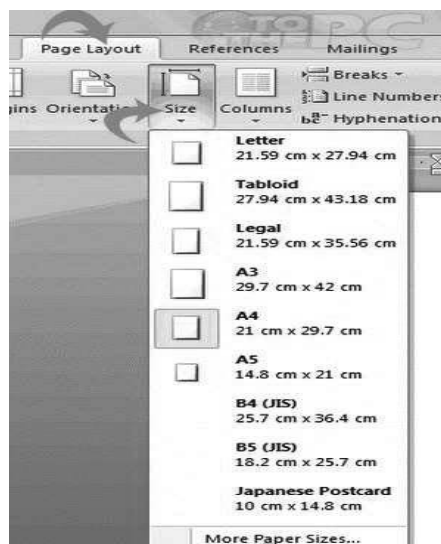
1. Select the text in the document.
2. Click the Line and Paragraph Spacing command in the Paragraph group on the Home tab.
3. Select the desired spacing option from the drop-down menu.
4. From the drop-down menu, you can also select Line Spacing Options to open the Paragraph dialog box.



Page length

You can set the page length using the size command of page setup group of page layout menu in the ribbon interface as shown below.

NOTES



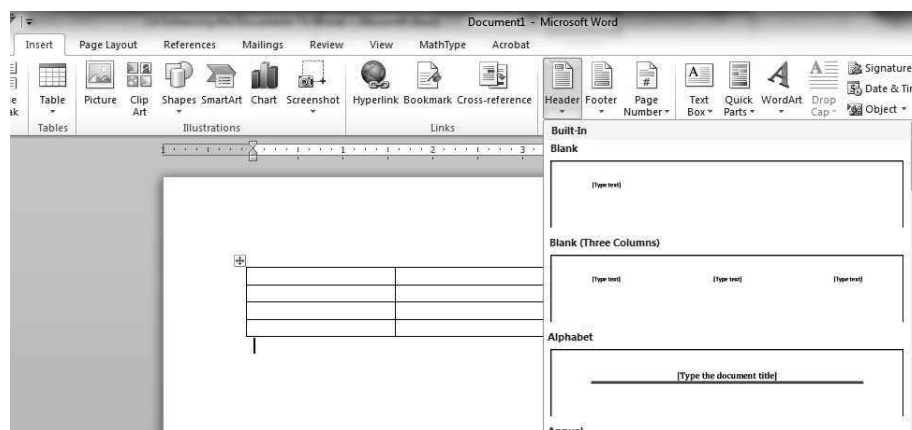
Header or Footer

Header and footer are the important parts of any document. They are used for representing information about the document such as the page number, heading of the document etc. Header is present on the top of the document and footer is present at the bottom of the document.

Adding Header or Footer

- Click on the **Insert** tab. After this, click on either **Header** or **Footer** whichever you want to apply to the document. A list appears from which you can select the header you want to apply to the document.
- When you select the header, it will appear in the editable form in which you can write whatever you want to add.

The screenshot of inserting the header in the document is as shown:

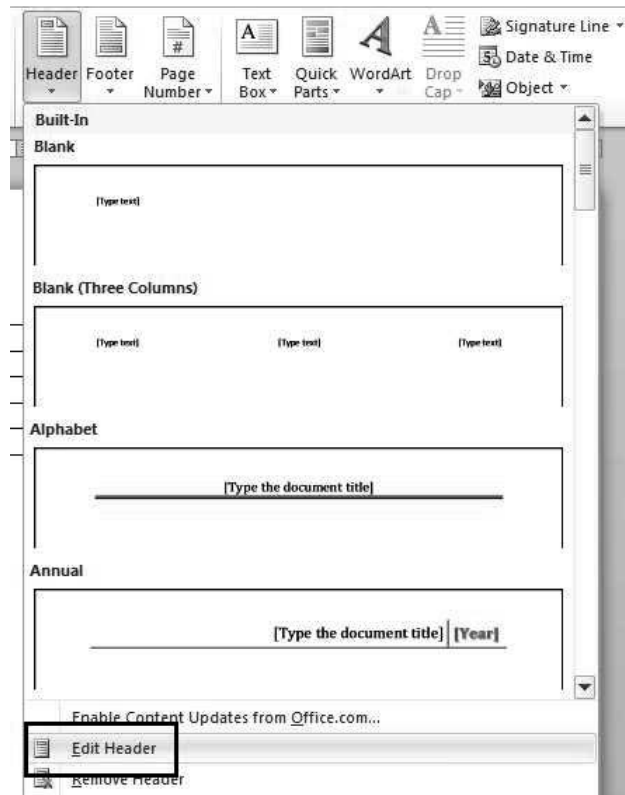


Editing the Header or Footer

- Click on the **Insert** tab. Select header or footer whichever you want to edit. A list will appear with the option **Edit Header**.

- Click the **Edit Header** button. The header is now available in editing form so that you can edit it.

The screenshot of **Edit Header** is as shown:



NOTES

Inserting special symbols and pictures

You can insert the symbols in a Word document using the symbol command in the insert tab in the ribbon interface. When you click on the symbol command, it shows the various types of symbols which can be selected as per the requirements.

You can also insert the pictures in the document from the available pictures using the pictures command in the insert tab. Click on the Insert tab on the Ribbon interface and select pictures. On the window that appears, go to the location where the picture is located and click Insert.

To wrap text around an image follow the given steps.

1. Select the image. The Format tab will appear.
2. Click the Format tab.
3. Click the Wrap Text command in the Arrange group.
4. Select the desired menu option. The text will adjust based on the option you have selected. ...
5. Move the image around to see how the text wraps for each setting.

Hyphenation Off

To turn off auto hyphenation in Word 2010, follow the steps given below

1. Go to the Page Layout tab > Page Setup group.
2. Click Hyphenation.
3. Select None.

NOTES



Types of justifications

There are several types of justification:

- **Left-justification:** All lines in the paragraph butt up against the left text margin. No extra spaces are added to the line.
- **Center-justification:** All lines in a paragraph are centred between the left and right text margins. No extra spaces are added to the line.
- **Right-justification:** All lines in a paragraph butt up against the right text margin. No extra spaces are added to the line.
- **Fill-justification.** All lines in a paragraph are expanded so they butt up against both the left and right text margins. Space is added, between words and characters, as necessary to fill out the line.

Cursor movement and control

You can move the cursor to another location by moving the pointer and then clicking, or by using the keyboard. To type anywhere else in the document, you need to move the cursor to that place. Here are a couple of ways to do that:

- With your mouse, move the cursor just to the left of “During,” and then click to insert the cursor. Then start typing.
- Press the UP ARROW () on your keyboard to move the cursor up one line at a time. Then press the LEFT ARROW () to move the cursor left, one character at a time.

5.3.4 Bullets and Numbering

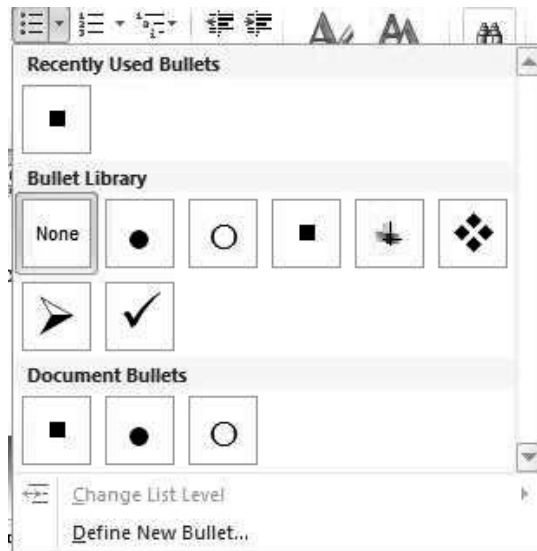
Bullets and numbering allow you to present information in the simple ways. With the help of bullets and numbering, information is separated easily in the form of list. This helps in saving the space in the document. Bullet and numbering option are available in the paragraph group on the **Home** tab. Microsoft Word lets you to create bulleted and numbered lists, to specify different lines or topics. Once you create bulleted and numbered list, each time you press enter, same bullets are added to your list. If you want to add another bullet, then you can choose different bullets and numbering from the **Home** tab. If you hit enter twice, then you will return to the earlier level. If you hit enter twice at the end of the list, then the list will be terminated.

The screenshot of bullet and numbering list is as shown:



Bulleted list

The screenshot of different types of bullets that can be used to create a bulleted list are as shown:

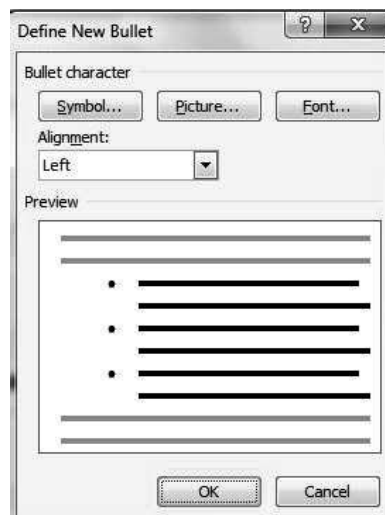


NOTES

Custom bullet

You can define your own bullet using **Define New Bullet** option. When you click on this option, following dialog box appears. This holds option for setting picture, symbol and font of the bullet character. It also contains the options for alignment and for seeing the preview of the page.

The screenshot of **Define New Bullet** dialog box is as shown:

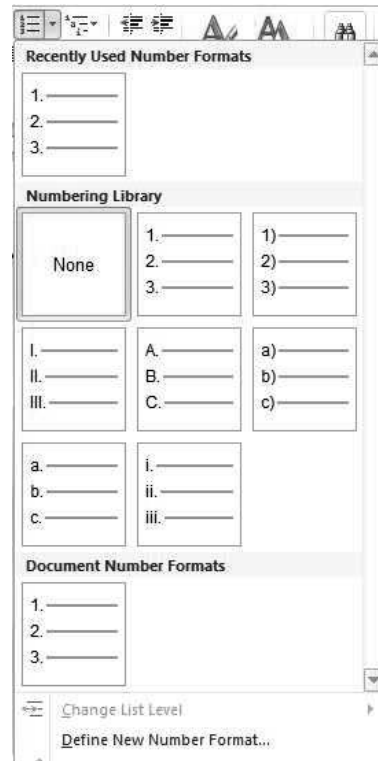


Numbered list

Numbered list dialog box contains different types of numbered lists like Recently Used Numbered Format, Numbering Library and Document Number Formats.

The screenshot of different types of numbered list is as shown:

NOTES



Custom Numbering

You can define your own numbering using **Define New Number Formats** option. When you click on this option, a dialog box appears containing options for setting number format. It also contains the options for alignment and for seeing the preview of the page.

Font Formatting

Font formatting is used to change the appearance of words or a single character. It is present in the **Home** tab of ribbon view. You can apply the Font formatting by highlighting the text.

Font Formatting contains the various controls for providing a quick access to many frequently used functions.

The screenshot of **Font Formatting** is as shown:



The available **buttons and icons** of the Font Formatting toolbar are as follows:

- **Calibri (Body)** Font dropdown box of the Font Formatting helps in changing the font of text or number when selected. You can choose any font by clicking on the dropdown menu. The default font in Word is Calibri (Body).

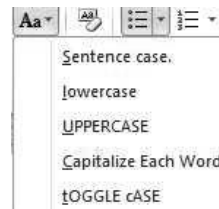
- **11** ▾ **Font Size** dropdown box of the Font Formatting helps in changing the size of the text when selected. You can choose any size by clicking on the dropdown menu. The default font size in Word is 11.
- **A⁺** **Grow Font** button of the Font Formatting helps in increasing the size of the text when selected.
- **A⁻** **Shrink Font** button of the Font Formatting helps in decreasing the size of the text when selected.
- **Change Case** dropdown menu of the Font Formatting helps in changing the case of the selected text.

NOTES

For example:

Suppose you have written “Hello”. You want to change the uppercase letter to lowercase, then you can choose **Change Case** dropdown menu to do that.

The screenshot of **Change Case** is as shown:



- **Aa** **Clear Formatting** menu of the Font formatting helps in clearing the formatting by selecting that particular area. It is useful in situation where you want to remove entire formatting at once.
- **B** **Bold** button of the Font Formatting toolbar helps in making selected text or numbers bold and in removing bold formatting when the text or numbers are already bold.
- **I** **Italics** button of the Font Formatting toolbar helps in making selected text or numbers italics and in removing italics formatting when the text or numbers are already italics.
- **U** ▾ **Underline** dropdown box of the Font Formatting toolbar helps in making selected text or numbers underline and in removing underline formatting in already underlined text or numbers. You can choose any type of line from Underline dropdown menu.
- **abc** **Strikethrough** button draws one straight line throughout the selected word.

For example:

Consider the following sentence “See the effect”. If you apply Strikethrough effect by selecting the words then the same sentence will look as shown: “~~See the effect~~”.

- **x₂** **Subscript** button helps in making selected text or character look smaller without changing its Font size. It places them below the other characters.

NOTES


For example:

Consider the following set of character “a5b.” If you apply Subscript effect by selecting number 5, then the following set of character will look as shown: “a₅b.”

-  **Superscript** button helps in making selected text or character look smaller without changing its Font size. It places them above the other characters.



For example:

Consider the following set of character “a5b.” If you apply Superscript effect by selecting number 5, then the following set of characters will look as shown: “a⁵b.”

-  **Text Effects** dropdown menu of the Font Formatting toolbar helps in drawing colored line around the text. You can choose any color from dropdown menu.

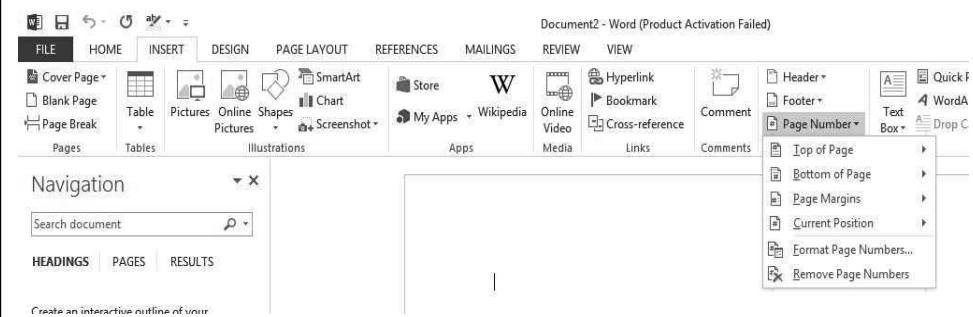
For example:

Consider the following sentence “See the effect”. If you apply Text Effect by selecting the text then the same sentence will look as shown: “**See the effect**”

-  **Highlight** button of the Font Formatting toolbar highlights the text or number with the specified color. By default, “Yellow” color is selected. You can choose any color from the drop down menu. You can add and remove the color by clicking the specific color.
-  **Font Color** button of the formatting toolbar helps in changing the color of the text. By default, “Red” color is selected. You can choose any color from the drop down menu.

Page Numbering

For adding Page Numbers to the Header section or Footer section simply click on the Page Number command to select the option Top of Page or Bottom of Page as shown below in the Illustration.



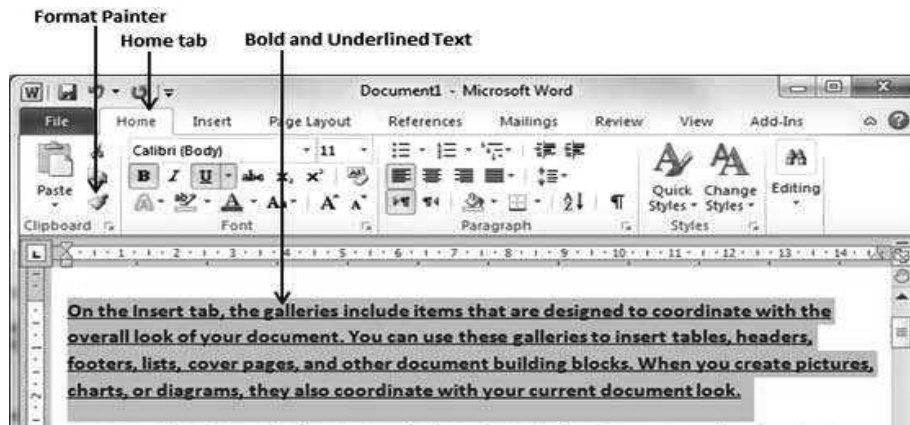
Inserting date and time

You can insert date and time in a document by simply click on date and time command in the insert tab. Follow the steps Insert → Date and Time. After using this you will see date and time in the document.

Format Painter

To use the format painter, use the following steps.

1. Select the text or graphic that has the formatting that you want to copy.
2. On the Home tab, click Format Painter.



3. Use the brush to paint over a selection of text or graphics to apply the formatting. This only works once. To change the format of multiple selections in your document, you must first double-click Format Painter.
4. To stop formatting, press ESC.

5.3.5 Shortcut Keys for Various Tasks

Some of the shortcut keys for various tasks in MS Word are shown in the table below:

Table 5.2 MS Word Shortcuts

Using Keyboard Shortcut	
Action	Keystrokes
Creation of new document	CTRL+N
Saving a document	CTRL+S
Opening a document	CTRL+O
Printing a document	CTRL+P
Closing a document	CTRL+W
Selecting an entire document	CTRL+A
Copying a text	CTRL+C
Cut specific text	CTRL+X
Paste specific text	CTRL+V
Undo	CTRL+Z
Redo	CTRL+Y
Bold	CTRL+B
Italics	CTRL+I
Underline	CTRL+U
Alignment of text to the left	CTRL+L

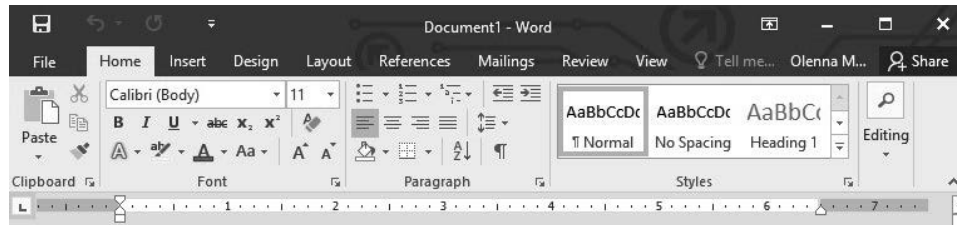
NOTES

NOTES

Alignment of text to the right	CTRL+R
Alignment of text to the justify	CTRL+J
Alignment of text to the center	CTRL+E
Finding the text	CTRL+F
Replacing the text with particular text	CTRL+H
Adding and removing 6 point of spacing before a paragraph	CTRL+0
Opening font preferences windows	CTRL+D
Inserting link	CTRL+K
Indentation of a paragraph	CTRL+M
Creating an hanging indent	CTRL+T
Creating a bullet point	CTRL+SHIFT+L
Changing the font	CTRL+SHIFT+F
Increasing selected font to 1 point to 12 point	CTRL+SHIFT+>
Increasing selected font to 1 point	CTRL+]
Decreasing selected font to -1 to -12 point	CTRL+SHIFT+<
Decreasing selected font to -1 point	CTRL+[
Viewing or hiding non printing characters	CTRL+SHIFT+*
Moving contents to the beginning	CTRL+'!
Moving contents to the end	CTRL+'"
Deleting Word to the right of the cursor	CTRL+Del
Deleting Word to the left of the cursor	CTRL+Backspace
Moving cursor to the end	CTRL+END
Moving cursor to the beginning	CTRL+HOME
Resetting default font to the highlighted text	CTRL+SPACEBAR
Changing text to heading one	CTRL+ALT+1
Inserting time	SHIFT+ALT+T
Inserting date	SHIFT+ALT+D
To Save As file	F12
Spell Checker	F7
Opening Help	F1

New Features in Word 2016/ Difference between Word 2010 and 2016

The **Tell Me** box in Word, Excel and PowerPoint 2016 helps in performing important tasks. Rather than using Help or just poking around as in 2010 version, you can now type what you want to do in this box. Office shows you a list of commands for completing your task. If you have selected something in your document, you will see commands specific to that object.



NOTES

Insights lets you search for information from within Word, Outlook, Excel, and PowerPoint. Right-click a word or phrase and choose **Smart Lookup**.



If several people are working on a Word document that is stored in OneDrive for Business, **real-time co-authoring**, the feature which was not available Word 2010, allows multiple users to collaborate, edit and update simultaneously. They can see changes as they are made and even where cursors are currently placed.

Word 2016 has the new editing feature of adding and viewing threaded comments, which was missing in Word 2010. There are five view modes in Word 2016 as compared to four in Word 2010. Smart Look Up feature of Word 2016 was missing in Word 2010 and 2013. You can send a link to a document saved on OneDrive, if you are using Word 2016, which was not possible in 2010 version.

5.4 MICROSOFT EXCEL 2010

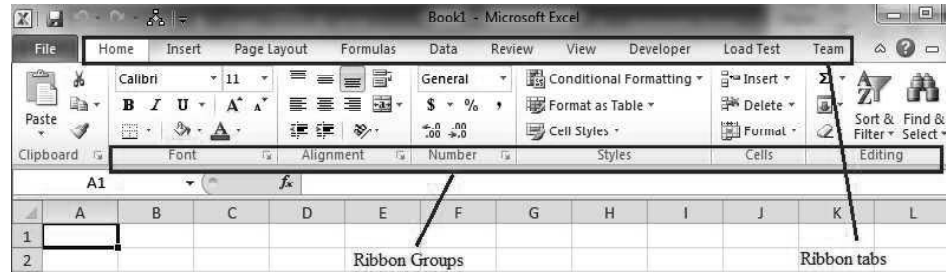
Let us analyse the different features of MS Excel.

The Ribbon Interface

Ribbon interface in Microsoft Word 2010 is same as that of Microsoft Excel 2010, the only difference lying with the changes in commands. With the help of ribbon performing various tasks with the spreadsheet becomes easy.

Screenshot for Ribbon is shown as:

NOTES



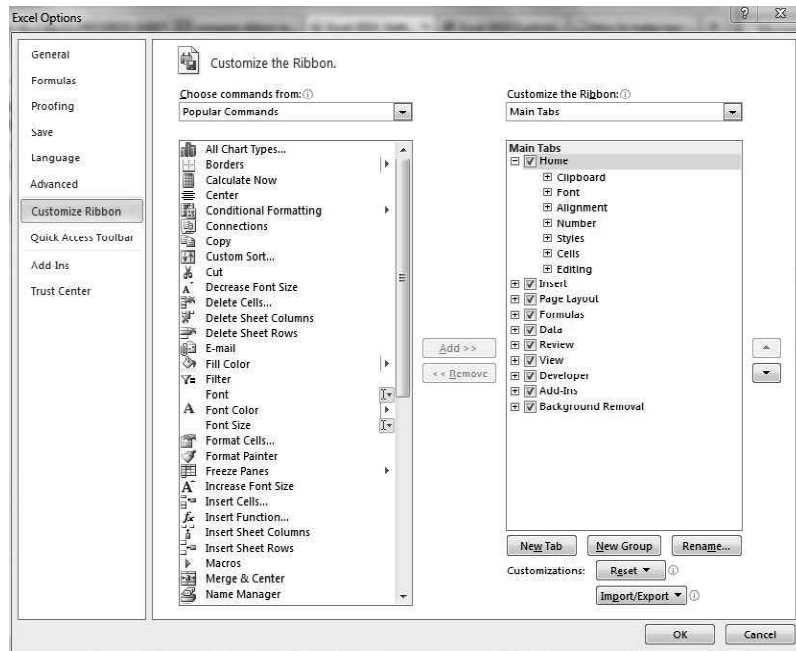
If you click on any tab or groups, each button and dropdown menu will perform different actions

Customizing the Ribbon

Follow these steps to customize the ribbon:

1. Right click on any tab or group and click on the **Customize the Ribbon** option. **Excel Options** dialog box will appear.

Screenshot displaying **Excel Options** is as shown:



2. Choose the command which you want to add. Here, **Copy** command has been selected. Next, click on **New Group** button (in order to add commands the user needs to create a custom group first). Click on **Add** button in order to add selected commands to the **New Group**.
3. Click on **OK** button. **Copy** command will be added to **New Group** which is seen under **Home** Tab after the **Editing** option.

Screenshot displaying **New Group** is as shown:



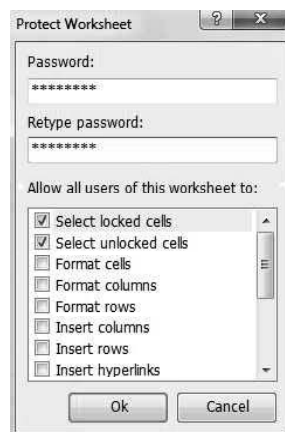
Excel 2010 Backstage View

Backstage view has been previously introduced in MS Word 2010. Most of the features of backstage view remain the same. Some modifications are discussed as follows:

Protect Workbook under Permission tab of Info option includes two more choices. They are:

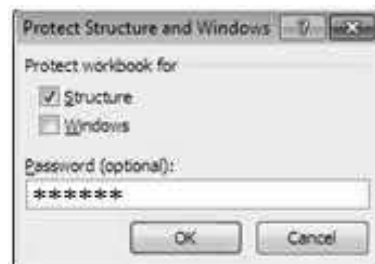
- **Protect Current Sheet:** This option protects current sheet from unauthorized access. Here you set a password for protecting current worksheet. You can also specify which type of changes people can make in this worksheet by checking them on allow all users of this workbook to list.

Snapshot of Protect Worksheet is as shown:



- **Protect Workbook Structure:** By using this option you can protect entire workbook to be changed by unauthorized users. It provides two levels of protection. These are as follows:
 1. **Protect Workbook for Structure:** It protects the structure of the worksheet from unwanted modifications. For example unauthorized users can not add new worksheet or delete an old and cannot change the order of the sheets.
 2. **Protect Workbook for Windows:** It is not checked by default. If you check Windows then the book is protected also from resizing any windows you have inserted in your workbook like charts, pictures, shapes etc.

Screenshot of Protect Structure and Windows is as shown:



NOTES

As shown, in both the above cases Password is optional.

In order to unprotect, click on Unprotect Worksheet or Unprotect Workbook. You are asked for password, if set.

NOTES

5.4.1 Creating, Opening and Working with Excel Documents

Microsoft Excel is used for storing the data in the form of table. It is used by several organizations to perform complex calculations, statistical analysis and tracking income and expenses. It uses pie charts, slicers, sparklines to easily analyze the data. Excel 2010 provides very interactive features which help in analyzing and visualizing the data in productive and innovative ways. In this chapter, you will learn about the essential features of MS Excel, various ways of building Worksheets, methods of formatting and filtering data, management of Excel worksheets and various shortcut commands to perform the task easily.

File Management

File management is the task of maintaining the newly created folders or the existing folders in the computer. It involves various tasks such as saving the folder, maintaining the data and many more.

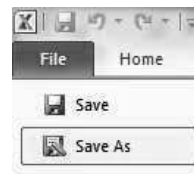
Saving a document

In Microsoft Excel 2010, files are saved in .xlsx format. There are different ways of saving the files in the computer depending upon the requirements. When the files are saved in .xlsx format, you can use new features provided by Excel such as Sparkline and Slicers which are not supported by other file formats.

Steps to save the documents in Excel are as follows:

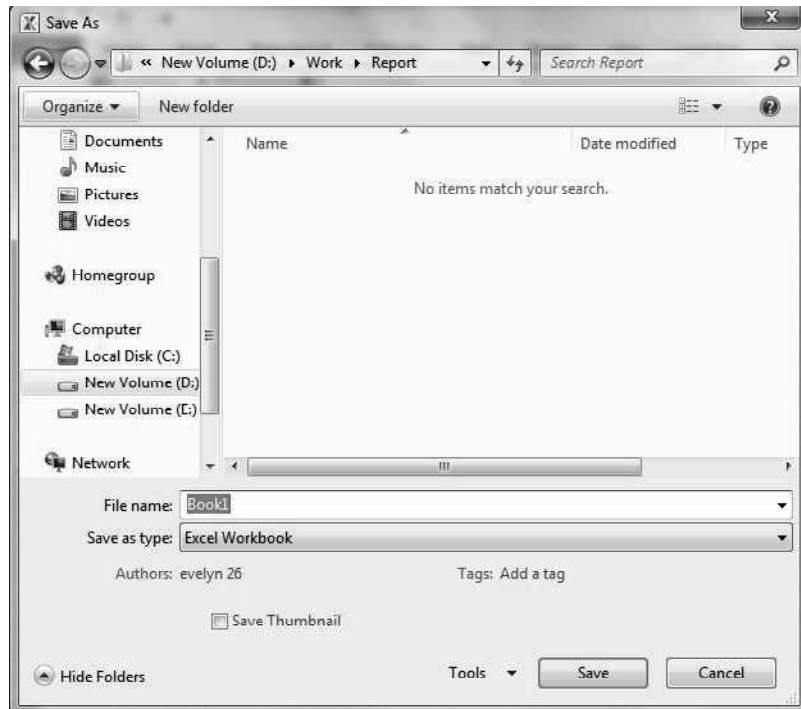
1. Go to the **File** tab which is present on the left hand side of the ribbon.
2. Go to the **Save As** options on the left most side of the screen.

The screenshot of **Save As** option is as shown:



3. Choose the drive and folder where you want to save the file.

The screenshot displaying specific **Drive and Folder** is as shown:



NOTES

4. Go to the **File name** textbox to write the name of the file.
5. Choose the format in which you want to save the File. By default, the file format is .xlsx.
6. Click the **Save** button to save the file.

Saving a Document as PDF

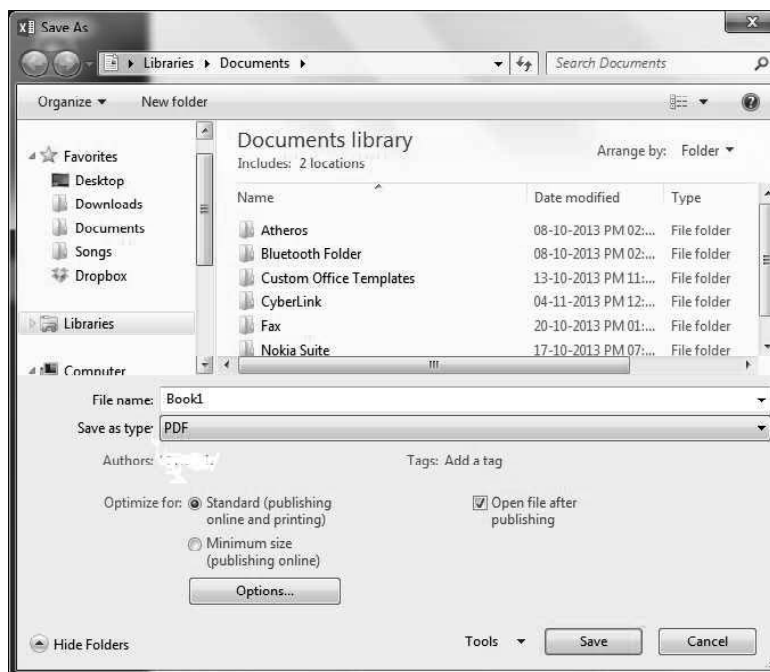
Excel 2010 provides enables the user to store documents as PDF files, while sharing them with others. PDF stands for Portable Document Format.

To save an Excel file in PDF format follow these steps:

- First click on the **File** tab present on the top left corner of MS Excel or press Alt+F as a shortcut to **File** tab.
- After this click on **Save As** option.
- Now, enter the filename. When you are asked to select the type of file, choose **PDF**. You can select **Standard** for high quality printing or select **Minimum size** if you desire a file of small size.

Screenshot displaying the **Save as type** option is as follows:

NOTES



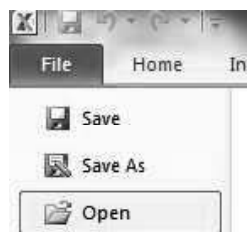
- Finally, click on **Save** button to save file in PDF format.

Opening a document

Steps to open the document are as below:

1. Go to the **File** tab which is present on the left hand side of the ribbon.
2. Go to the **Open** option on the left most side of the screen.

The screenshot showing **Save As** option is as shown:



3. Choose the appropriate drive and folder from where you want to open the file.
4. Choose the file which you want to open.
5. Go to the **Open** option to open the file.

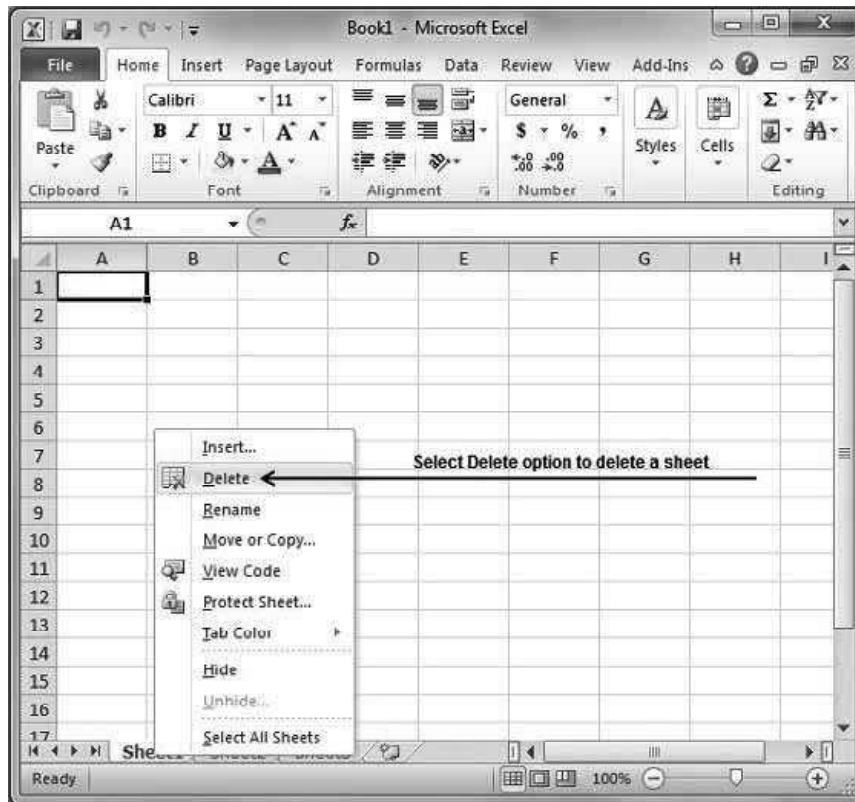
You can save the important files in the **Libraries** so that you can access the specific folder and files easily. The **Libraries** option appears on the left hand side of the screen.

Deleting a spreadsheet

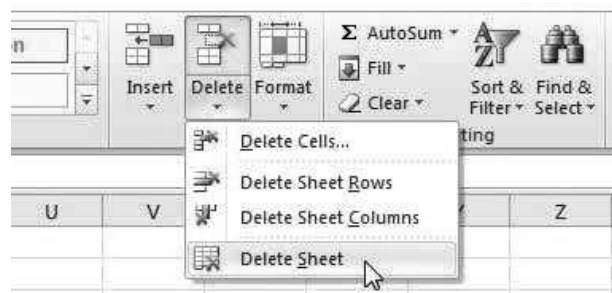
Follow the steps given below to delete a spreadsheet.

1. Open the Excel workbook containing the sheet that you want to delete.
2. Click the tab at the bottom of the window for the worksheet that you want to delete.

NOTES



3. Click the **Home** tab at the top of the window.
4. Click the arrow under the **Delete** button in the **Cells** section of the ribbon at the top of the window, then click the **Delete Sheet** button.



5. Click the **Delete** button to confirm that you want to delete the sheet.

Reordering spreadsheets

You can change the order of sheets as per your need in a file by simply click, hold and drag on the sheet tab at the bottom of the window.

Renaming a spreadsheet

You can follow the steps given below to rename a spreadsheet:

- Double-click on one of the existing worksheet names.
- Right-click on an existing worksheet name, then choose Rename from the resulting Context menu.

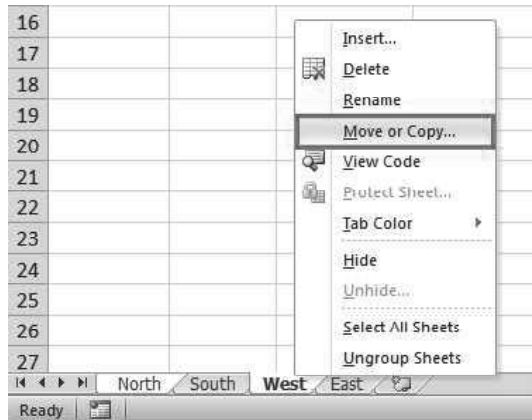
- Select the worksheet you want to rename (click on the worksheet tab) and then select the Sheet option from the Format menu. This displays a submenu from which you should select the Rename option.

NOTES

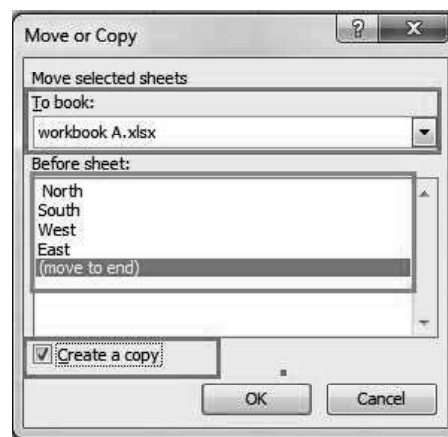
Duplicating/Copying sheets in a workbook

Using **Move or Copy** command, you can make one copy of a worksheet, multiple specific worksheet or all worksheets into active workbook or another workbook.

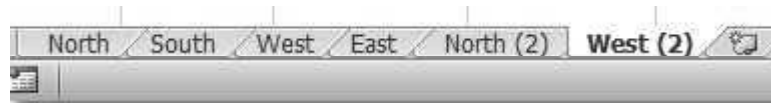
1. Select the workbook you need to copy.
2. Right click on sheet tab and select **Move or Copy** option. An illustration is shown below.



3. In the **Move or Copy** dialog box, specify the following settings.
 - From **To book**, you can specify the workbook where the worksheets will be copied into.
 - Specify the position of the copied sheets, you can choose it after all of the existing sheets.
 - Check **Create a copy** option, if you don't check this option, the selected worksheets will be moved into another workbook.



4. And then click **OK**, it will copy the selected worksheets one time into the specified workbook. The screenshot is given below.



5.4.2 Working with Cells

Excel file is a workbook that contains one or more worksheets. Formatting Excel sheet improves the presentation of the worksheet and makes it easy to use. By default, each Excel file has three worksheets.

Entering data and Text formatting

To insert data in the spreadsheet, you can select the row or column where you want to insert the data or insert new rows and column in between.

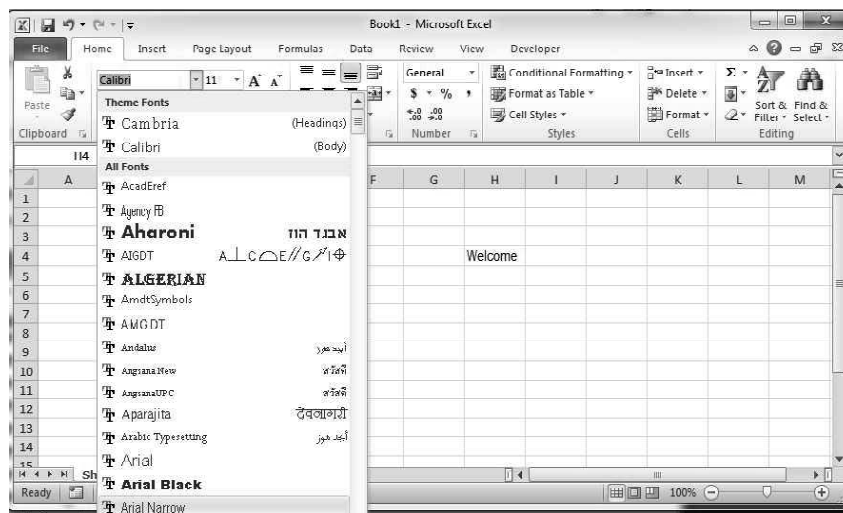
Formatting Cells

Microsoft Excel has 16,384 columns and 1,048,576 rows. Cell is intersection of rows and columns where you can enter data in the cell. There are many options to format cell and text. We are going to learn how to change colour, style of cell, alignment of text, etc. There are four ribbon groups such as **Font**, **Alignment**, **Number and Styles**. **Font** command is used to change font, font size, and style of text such as **Bold**, **Italic** and **Underline**, change border, change the cell color and text color. **Alignment** command is used to set alignment of text and merging the cells. **Number** command is used to select cells and set number format such as percentage, date, etc. **Cell styles** option in **Styles** group is used to set specific color to cell.

Steps to change the **Font** are as follows:

1. Select a cell in which you want to set a specific font.
2. Click on **Font** option from the **Font** on **Home** tab.
3. Select the desired font from drop down list to apply the font.

The screenshot of selecting the **Font** is as shown:



NOTES

NOTES

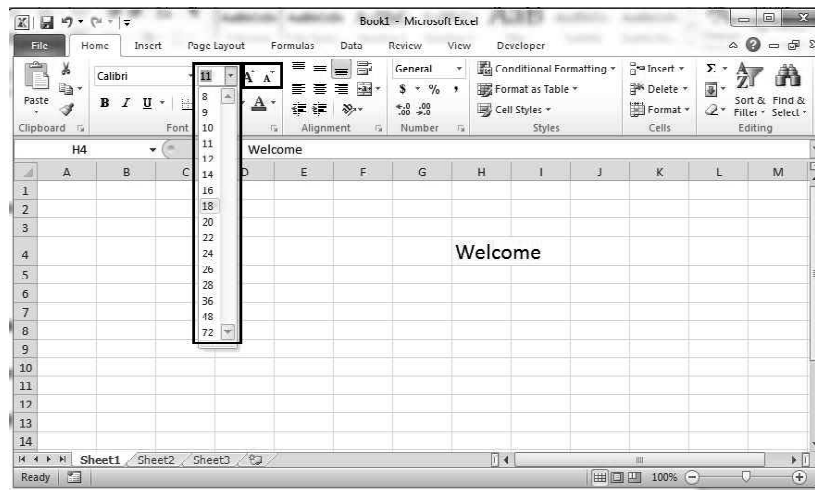
To change a font size

Steps to change a **font size** are as follows:

1. Select a cell in which you want to set specific font size.
2. Click on **Font Size** from the **Font** on **Home** tab. There are two options for increasing and decreasing fonts.
3. Select font size from drop down list. Otherwise, click on increasing font or decreasing font.

If you click on increasing font option, text size increases according to the selection. If you click on decreasing font option, text size decreases according to the selection.

The screenshot of selecting the **Font Size** is as shown:



To set Bold, Italic and Underline option

Steps to set **Bold, Italic and Underline** are as follows:

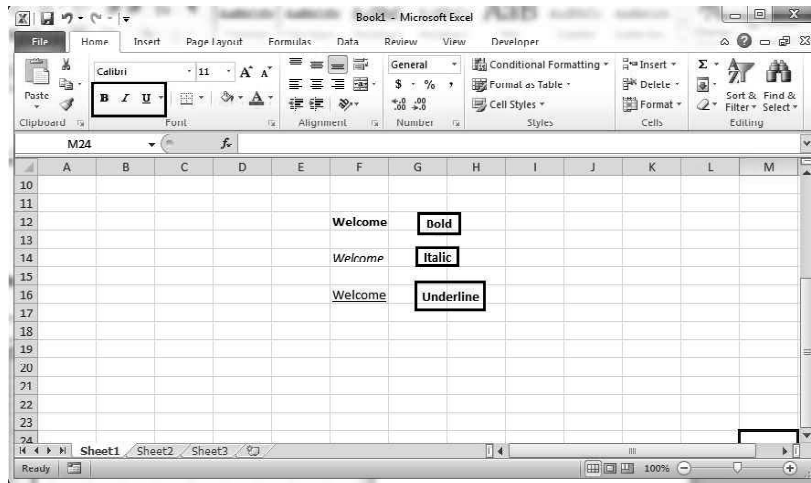
1. Select a cell in which you want to set style such as **Bold, Italic and Underline**.
2. Click on **Style** option in **Font** group on **Home** tab. There are three options such as **Bold, Italic and Underline**.

An example of **Bold, Italic and Underline** is as shown:

Bold	B	Welcome
Italic	<i>I</i>	<i>Welcome</i>
Underline	<u>U</u>	<u>Welcome</u>

3. Text appears according to selection.

The screenshot of selecting the font style is as shown:



NOTES

To change text alignment

Steps to change **text alignment** are as follows:

1. Select the cell.
2. Click on the **Home** tab.
3. Click on vertical Alignment and horizontal Alignment commands in **Alignment** group.

In vertical Alignment, there are three options:

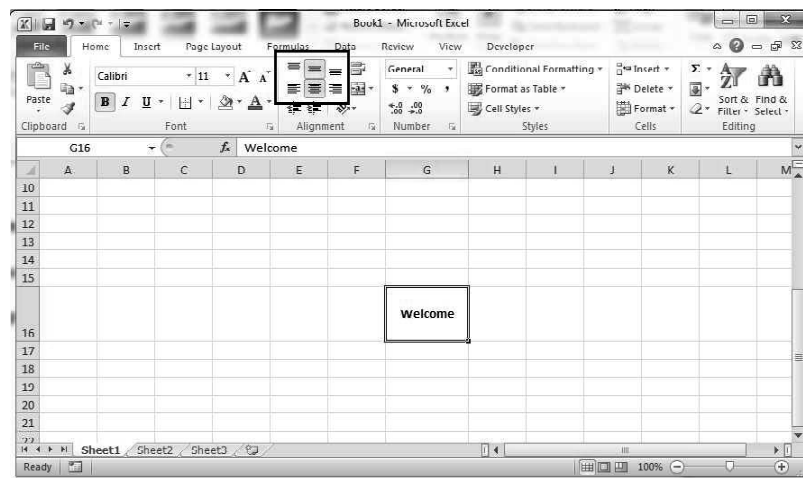
1. **Top Align:** Text set to the top of the cell.
2. **Middle Align:** Text set to the middle of the cell.
3. **Bottom Align:** Text set to the bottom of the cell.

In horizontal Alignment, there are three options:

- **Align Text Left:** Text set to the left of the cell.
- **Center:** Text set to the center of the cell.
- **Align Text Right:** Text set to the right of the cell.

Text aligns according to the selection.

The screenshot of selecting the **Alignments** is as shown:



The given screenshot shows vertically Middle Align and horizontally Center alignment.

Applying cell styles

NOTES

Cell styles contain the formatting of font and cell. You can directly use cell styles to set the range of cells. Excel 2010 has in-built cell styles that you can use. You can also create your own cell style.

Steps to apply **cell style** are as follows:

1. Select the cell.
2. Click on cell styles in **Styles** group on **Home** tab.
3. Select cell style that you want to apply on the cell. Cell is now displayed according to the style selected.

The screenshot of the **cell style** is as shown:



Selecting Cells or Range

When you edit sheets in MS Excel 2010, you often need to select cells, copy cells, set a range for deleting cells, and also search for data within the cells of a worksheet. To select a particular cell, select that cell or use arrow keys to reach that cell.

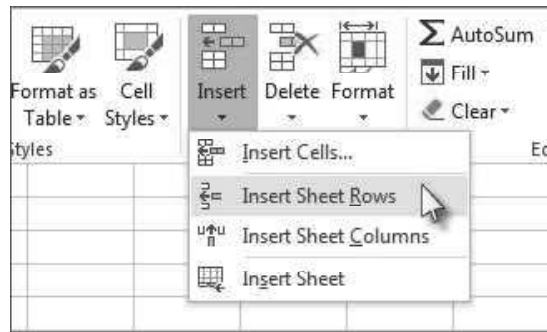
For the selection of range, click on first cell and then drag to the last cell or instead of dragging, just hold SHIFT key and use arrow keys to reach the last cell.

For selecting all cells in worksheet, press CTRL+A.

Inserting and Deleting Rows and Columns

To insert rows and column in a worksheet, follow the steps given below.

1. Select the cell or row or column where you want to insert row/column.
2. Click the drop-down arrow attached to the Insert button in the Cells group of the Home tab. The screenshot is given below.



NOTES

3. Select the one of the option in the drop down menu as per requirement.

5.4.3 Sorting and Filtering Information

In MS Excel, you can store enormous data. If you want to arrange the data, you can use **Sort** command. **Filter** command is used to find any specific information from specific column. In Excel 2010, **Sort** and **filter** command are available together in **Editing** group on **Home** tab. Both the commands are described in further subsections.

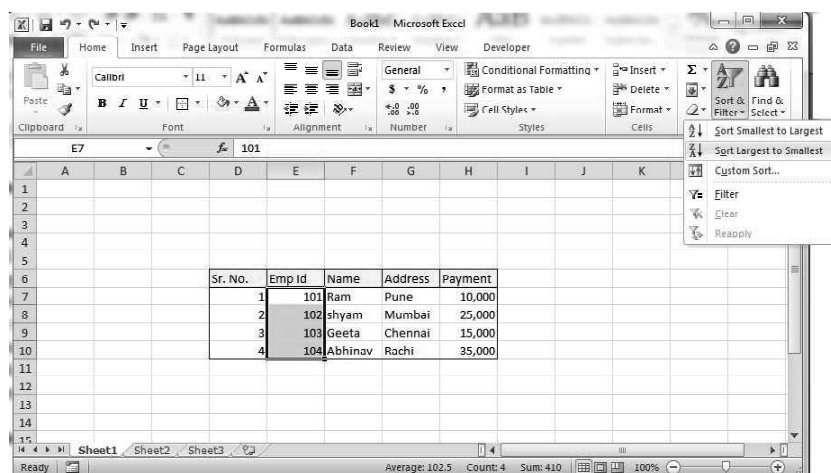
Sort: **Sort** is most commonly used command to arrange the list of data in alphabetical order. In Excel, there are two basic sorting methods such as A to Z and Z to A. For example, if you want to arrange employee name in ascending order, click on A to Z in **Sort & Filter** option. Sort command is used to visualize the data such that it increases understanding and makes data easily searchable.

Sort Data in single column

Steps to **Sort Data** in single column are as follows:

1. Select a column in which data needs to be sorted.
2. Click on **Sort & Filter** option in **Editing** group on **Home** tab. It displays drop down list.

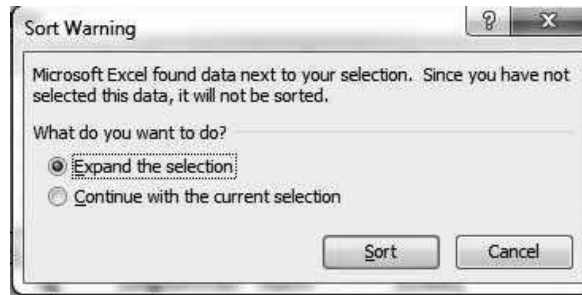
The screenshot of selecting **Sort & filter** is as shown:



3. Click on **Sort Largest to Smallest** option. It displays **Sort Warning** dialog box.

NOTES

The screenshot of **Sort Warning** dialog box is as shown:



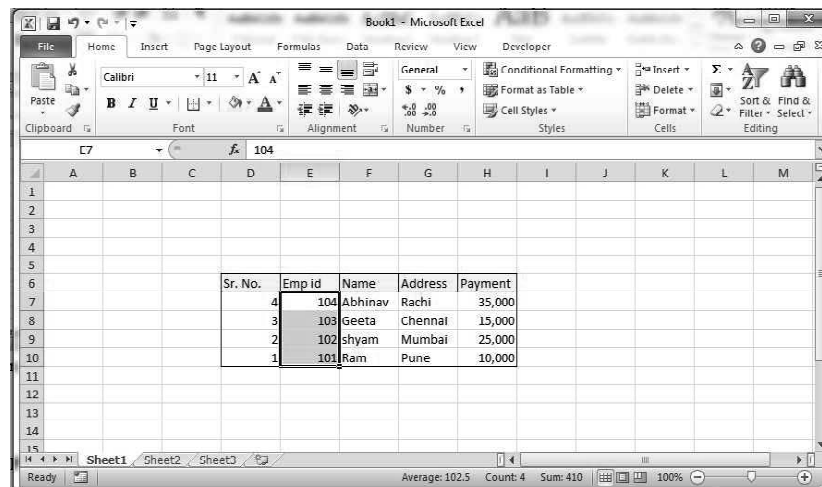
There are two options which are as follows:

- **Expand the selection:** Sort selected column with other columns data.
- **Continue with the current selection:** Sort only selected single column.

4. Click on **Expand the selection**.

5. Click on **Sort** button.

The screenshot of selecting **Sort** button is as shown:



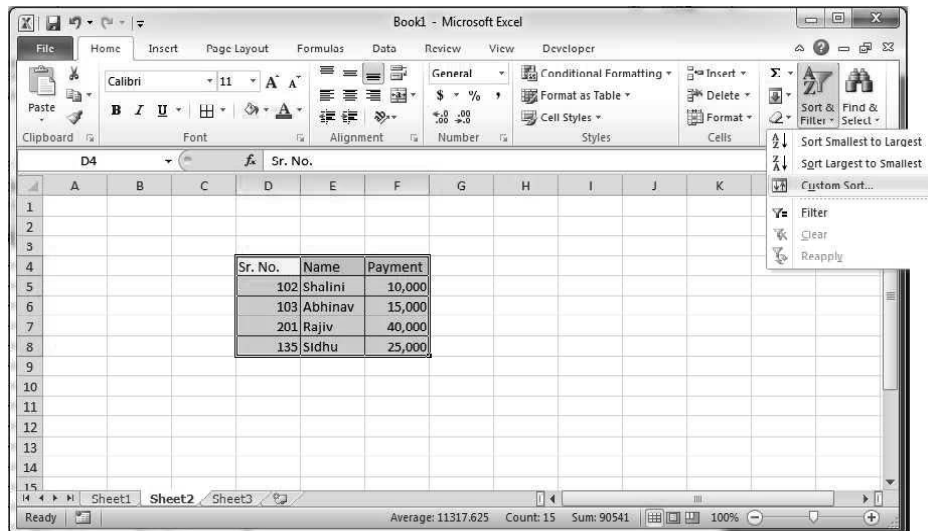
To Sort multiple data

Steps to **Sort multiple data** are as follows:

1. Select whole table in Excel sheet.
2. Click on **Sort & Filter** option in **Editing** group on **Home** tab. It displays drop down list.

The screenshot of selecting **Sort & Filter** is as shown:

NOTES

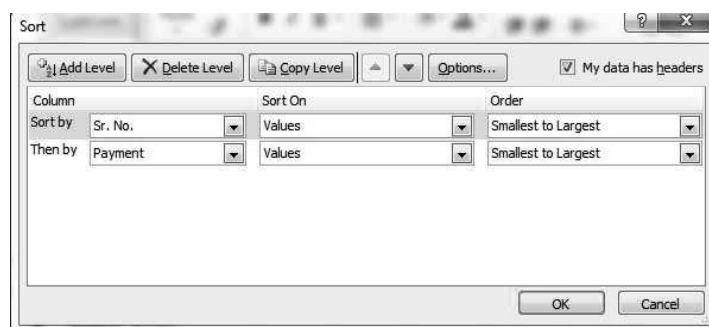


3. Click on **Custom Sort**. It shows **Sort** dialog box as shown in the screenshot:



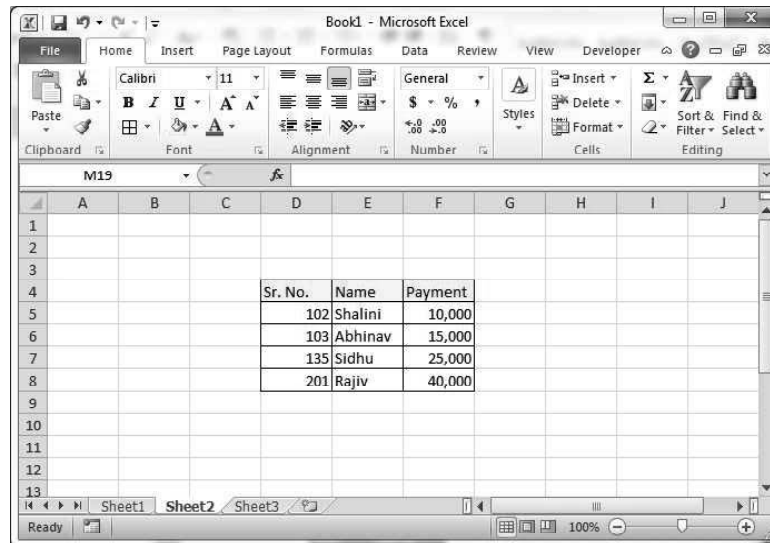
4. Select the column that you want to sort in **Sort by** drop-down box.
5. Select type of sort in Sort On list.
6. Select Values option if you want to sort text, number or date and time.
7. Select Cell color, Font color, Cell icon option if you want to sort data by format.
8. Select order in Order list.
9. **Add Level** is used to add another column to sort by. **Copy Level** is used to copy a column to sort by. **Delete Level** is used to delete selected entry. Up or down arrows are used to change the order of the **Column**.

The screenshot displaying the options in **Sort** dialog box is as shown:



NOTES

10. Click **OK** button. It displays a table which is as shown in the screenshot:



Filter

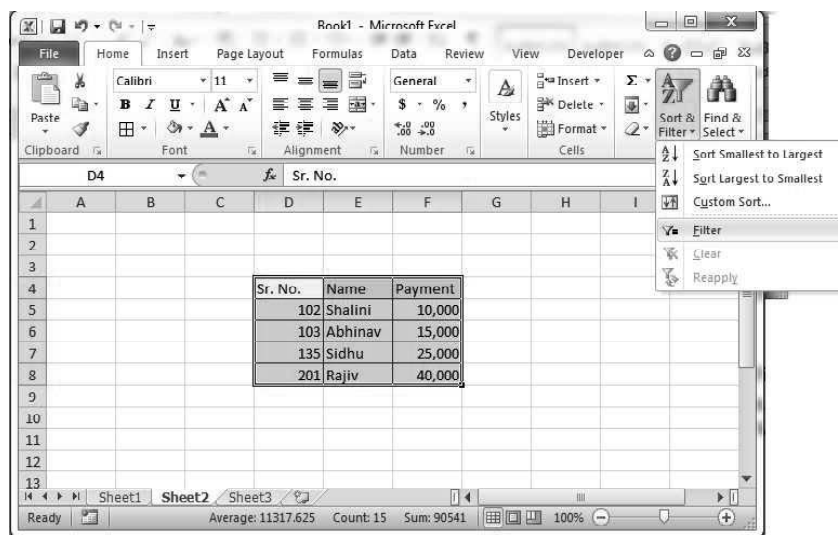
You can find any value, information quickly in worksheet using **Filter** command. You can filter more than one column of data. **Filter** command creates a list of data. Therefore, you can see any information from a list. Let's learn how to use filter command.

To filter data

Steps to **filter data** are as follows:

1. Select a table.
2. Click on **Sort & Filter** option in **Editing** group on **Home** tab. It displays the drop-down list.

The screenshot of selecting **Sort & Filter** is as shown:



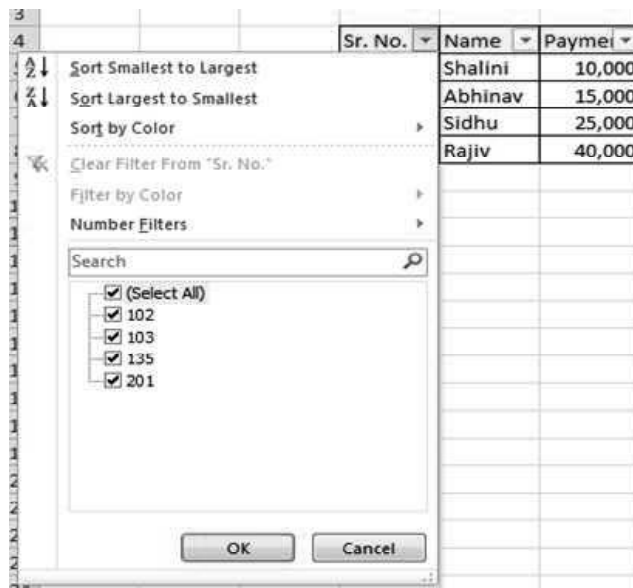
3. Select **Filter** option. It displays a table with drop-down arrows in the header of each column.

The screenshot of table after applying **filter** option is as shown:

Sr. No.	Name	Payme
102	Shalini	10,000
103	Abhinav	15,000
135	Sidhu	25,000
201	Rajiv	40,000

4. If you click on drop-down arrows, it displays the list of data with check box.

The screenshot of table displaying drop-down arrows is as shown:



When you uncheck any check box of data, that data will not be displayed in the table.

AutoSum in spreadsheet

Follow the given steps to apply AutoSum

1. Click the cell where you want to display the calculation.
 - To sum with a range of numbers, select the range of cells you want.
 - To sum with only some of the numbers in a range, select the cells or range you want using the Ctrl key. Excel inserts the sum in the first empty cell below the selected range.
 - To sum both across and down a table of number, select the range of cells with an additional column to the right and a row at the bottom.
2. Click the Formulas tab.
3. Click the AutoSum button.

NOTES

Working with Basic Formulae and Charts

There are various formulas and functions available in Microsoft Excel 2010, which make Excel simpler and easier to use. Let us explore these functions in the next sections.

NOTES

5.4.4 Working with Arithmetic Operators and Text Functions

Arithmetic Operators involve four major operations. These are: addition, subtraction, division and multiplication. These are known as the **basic operators**. These basic operators can be combined so as to perform more complex mathematical operations.

For example, if it is required to add cells A1 and A2 and then multiply their sum by a third cell (A3), then in Excel, this operation can be performed using the formula: $= (A1 + A2) \times A3$

Similarly, if you want to add cells from A1 to A5 and then multiply their sum by a third cell (A6), then, write the formula as: $=SUM (A1:A5) \times A6$

This formula adds numbers in the cell from A1 to A5 and the answer obtained is multiplied by number in cell A6. The answer is displayed in the cell where formula is written. While performing more complex operations on an arithmetic operator, operator precedence plays a very important role. Excel follows the BODMAS (Bracket Order Division Multiplication Addition Subtraction) rule of order of execution in arithmetic operations. The order precedence will be clear with the help of following example:

1. Open a new Excel sheet. In cell A1, enter 25. In cell A2, enter 50 and in cell A3, enter 2.
2. Click in cell A5 and enter the formula $= (A1 + A2) \times A3$ and hit enter. The answer displayed in A5 is 150.
3. Now, enter the formula $= A1 + A2 \times A3$ and hit enter. The answer displayed in A5 is 125. Clicking on cell means that the result will be shown in this cell.

In this example, the bracket is solved first where addition takes place. Thereafter, the result is multiplied by third number outside bracket.

The screenshot representing the discussed example is as shown:

A5		fx: =(A1+A2)*A3			
	A	B	C	D	E
1	25				
2	50				
3	2				
4					
5	150				
6					

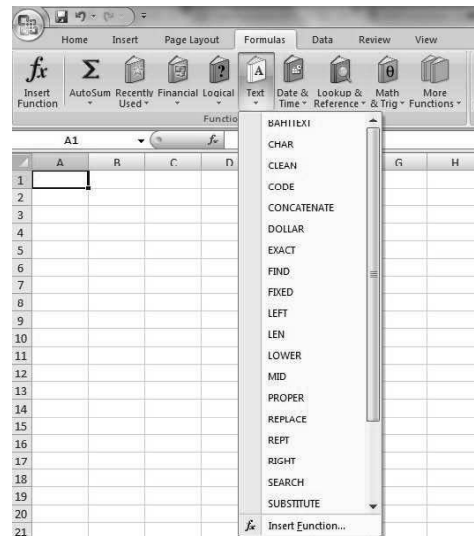
A5		fx: =A1+A2*A3			
	A	B	C	D	E
1	25				
2	50				
3	2				
4					
5	125				
6					

Text Functions

Text functions, which are the part of **Function Library** group, are available at the **Formulas** tab. Some of the functions available under **Text** functions are CHAR, CONCATENATE, REPLACE, LOWER, UPPER, DOLLAR, TEXT, etc.

The screenshot of **Text** functions is as shown:

Windows, MS Office and
Outlook 2010



NOTES

Some of the commonly used **text** functions are explained as follows:

- **CONCATENATE Function**

This function is used to combine or merge two values. Suppose, you have two columns and you want to combine both of them, then this function is used.

The syntax of **CONCATENATE** function is—**CONCATENATE (Text1, Text2, ...)**

Where, **Text1** represents first text string and **Text2** represents text string for concatenation with the previous text string. The text strings must be separated by commas.

For example, if the excel has “first name” of a person in it as the first column and the other column contains the “last name” then, to combine both the first and last name in a single column, **CONCATENATE** function can be used.

Use the **CONCATENATE** function as—**CONCATENATE (“John”, “, Smith”)**. Thus, on applying this function the output will be the full name of the person like—John Smith

- **REPLACE Function**

REPLACE function is used to remove the sequence of unwanted characters from the Excel sheet with another set of characters. These unwanted characters appear in the cell whenever the data is copied from one Excel sheet to another.

The syntax of **REPLACE** function is: **REPLACE (Old_text, Start_num, Num_chars, New_text)**.

Old_text represents the original string of characters required to be replaced.

Start_num identifies the position from where the substitution of the characters starts. **Num_chars** denotes the number of characters to be replaced in the **old_text**.

New_text is the set of characters that will be used for replacement.

NOTES

To access the REPLACE Function, go to the **Text** function drop-down menu, navigate to the **REPLACE** function and click on it. Also, enter the data in the respective fields of **Function Arguments** dialog box of the **REPLACE** function.

For example: Enter the original string as “**ab123**” (as string must always be enclosed in double inverted commas) in the **Old_text** and enter numeric value **2** in the **Start_num** field. In **Num_chars** field, enter the value as **4**. At last, in the **New_text** field, enter **cde** as the string which will be used for substitution.

After filling the data, when you click on the **OK** button, the final result or output of the **REPLACE** function will be obtained as:

acde

- **UPPER, PROPER and LOWER Function**

In Microsoft Excel, some functions are available for converting the entered data into uppercase, proper case and lowercase. If you want to make all the characters of a string uppercase, then enter the command = **UPPER (Place of string)**. In the similar way, the functions **LOWER** and **PROPER** are used.

Thus, if you enter the name as **William** and apply Upper function on it as— **UPPER (William)**, then the result obtained or displayed will be **WILLIAM**.

Similarly, if the name entered in the Excel sheet column is **RICKY PONTING** and **LOWER** function is applied on it as— **LOWER (RICKY PONTING)**, then the result obtained or displayed is **ricky ponting**.

And, in case you apply **PROPER** function with syntax as— **PROPER (RiCky PoNting)**, then you will get the output — **Ricky Ponting**.

- **TEXT Function**

The **TEXT** function helps in converting a numeral value or number to text in a required number format. This function is helpful when you want to display numbers in a more readable format. It also helps in combining numbers with text or symbol and in displaying the number as a text containing several places of decimal digits.

The syntax of **TEXT** function is

—**TEXT (Value, Format_text)** where, **Value** represents a numeric value or is a reference to a cell containing a numeral or numeric value. **Format_text** represents a field for entering the text string to be converted into a numeric value.

To access the **TEXT** Function, go to the **Text** function drop-down menu, navigate to the **TEXT** function and click on it. Now, enter the data in the two fields under the **Function Arguments** dialog box of the specific function.

For example, enter the number **143** as the **Value** and “**0.0000**” as the **Format_text** field to convert the number in the decimal format. After filling the data apply the **Text** function as— **TEXT (143, “0.0000”0)**. When you click on **OK**, then the final result of the **TEXT** function i.e. after converting a numeric value to a text containing 4 decimal digits, will be displayed in the Excel sheet as:

143.0000

• **Summing up Text Functions:**

Function	Description
BAHTTEXT	It is used for transforming or converting a number to a text using the currency format BAHT.
CHAR	It is used for providing the character of the specified code number or value.
CLEAN	It is used for clearing out all those characters from the text that are non-printable and will not be printed with the operating system.
CODE	It is used for extracting the first character from the text string and returning the numeric code for that.
DOLLAR	It is used for converting a numeric value to text using the \$ currency format.
EXACT	It is used for finding or comparing two text values.
FIND	It is used for locating the position of one text or substring within another text or string.
FIXED	It is used for rounding off the number and fixing the number of decimal places.
LEFT	It is applied to the text value for returning the leftmost characters from it.
LEN	It is used for calculating the value of the length of the string.
MID	It is used for returning the number of required characters using the provided position, from the text string.
REPT	It is used for repeating the text as per the number of times mentioned or given
RIGHT	It is applied on a text value for returning the rightmost characters from it.
SEARCH	It is used for searching or locating the required value or text within another value or text
SUBSTITUTE	It is used for replacing the old text with new or current text.
T	It is used for checking whether the passed value is text or not. If the value is text then it returns that and if it is not then double quotes are returned.
TRIM	It is used for replacing and trimming out the extra spaces present in the text, leaving behind only single spaces between words
VALUE	It is used for converting the argument or value passed as number to text.

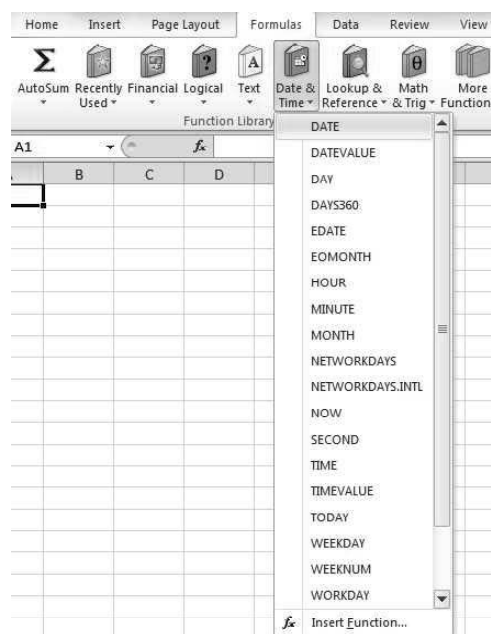
NOTES

NOTES

Date and Time Functions

Date & Time functions are used for making the modifications in date and time and also, to know the current date and time. There are various **Date & Time** functions which are collectively present or available in the MS Excel under **Date & Time** icon of the **Formula** tab.

The screenshot of **Date & Time** functions is as shown:



The **Date & Time** functions are explained as follows:

• DATE Function

The **DATE** function is used to represent the specific date. The syntax of **DATE** function is

“=DATE (Year, Month, Day)”. This function can be accessed from the **Date & Time** drop-down menu of **Function Library** group under **Formulas** tab.

For Example: In the **Function Arguments** dialog box of **DATE** function if you enter valid entries for **Year**, **Month** and **Day** for the respective fields, then the **DATE** function will be applied as: =DATE(2013,10,2) and you will get the output as 10/2/2013.

• TIME Function

The **TIME** function returns the time in **TIME** format (Hour, Minute and Second). The syntax of **TIME** function is — **TIME (Hour, Minute, Second)**.

This function is also accessed from the **Date & Time** drop-down menu of **Function Library** group of **Formulas** tab.

For Example: In the **Function Arguments** dialog box of the **TIME** function, if you fill the valid entries of **Hour**, **Minute** and **Second** in the corresponding fields and use the **TIME** function on the entered data as— **TIME**

(12,34,78), then the output in the cell will be shown as **12:35PM**.

• **Summing up Date & Time Functions**

Function	Description
DATEVALUE	It is used for representing the data in the Excel date-time format, by converting the text to numeric value or numbers
DAY	It is used for returning or calculating the day of the month by converting the provided date or serial number
DAYS360	It is used for finding or returning the days present between two provided dates. It calculates the number of days on the basis of either a year of 360 days or twelve-30 days-months.
EDATE	It is used for returning the same date in the serial format, by using the supplied start date. The specified months are added or subtracted from the start date to calculate the past or future dates.
EOMONTH	It is used for calculating or finding out the last day of the month using the given start date and specified month before or after the start date.
MONTH	It is used for extracting or returning the particular month from the supplied or provided date.
NETWORKDAYS	It is used for finding or returning the total number of working or business days present between two dates. Weekends and holidays are excluded while calculating the number.
NETWORKDAYS.INTL	This is a New Function in Excel 2010. It is used for calculating the total number of working or business days that are present between the two dates. It also gives out the number of weekends and holidays separately while calculating the number of workdays.
NOW	It is used for calculating or returning the present date and time
TIME	It is used for returning the time in the time zone format by using the supplied values of Hour, Minute and Second.
TIMEVALUE	It is used for returning the current date in a predefined format without taking any arguments or values.
TODAY	It is used for returning the current date in the predefined format without taking any arguments or values.
WEEKDAY	It is used for representing or finding the day of the week by using the supplied date value.

NOTES

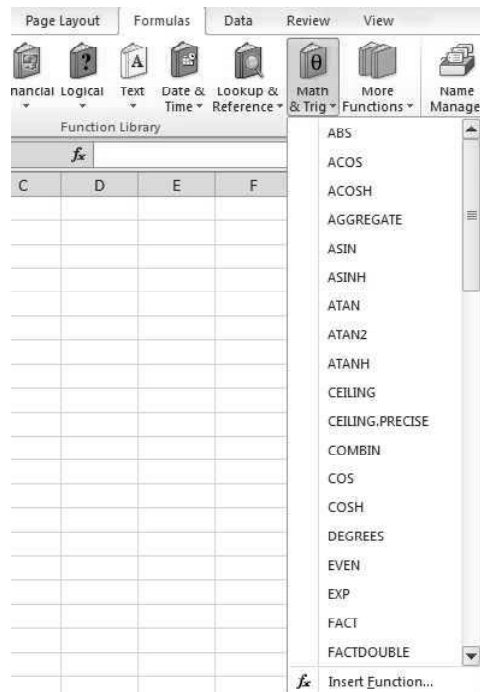
NOTES

WEEKNUM	It calculates or finds out the week number from the given date.
WORKDAY	It is used for calculating or knowing the date before or after given number of working days and using a specific date as start date.
WORKDAY.INTL	This is a New Function which has been added to Microsoft Excel 2010. It is used for providing the valid date before or after the supplied date, by calculating the given number of workdays on the basis of weekends.
YEAR	It is used for representing or extracting the year from a given date.
YEARFRAC	It is used for calculating the year in fractional form, on the basis of two supplied dates.

Math Functions

Math functions are those functions that are used to perform mathematical operations in Microsoft Excel. In Microsoft Excel 2010, you can access the math functions from the **Math & Trig** drop-down menu of the **Function Library** group. PRODUCT, QUOTIENT, POWER, AGGREGATE, MOD, SUM, FACT, EXP, SIN, COS etc. are some of the useful functions available in the **Math & Trig** functions.

The screenshot displaying **Math & Trig** drop-down menu is as shown:



The four basic math operations are PRODUCT, MOD, QUOTIENT and SUM are discussed as follows:

- **Product:** This function calculates the product or multiplication of the numbers in the given cells. **For example:** The resultant product, of two numbers **10** and **20**, using the **PRODUCT** function — **PRODUCT (10, 20)** will be **200**.
 - **Quotient:** This function calculates the quotient when one number is divided by another in the specified cells. **For example:** The Quotient obtained by dividing the number **34** by the numeric value **2** using the **QUOTIENT** function— **QUOTIENT (34, 2)** will be **2**.
 - **Mod:** This function calculates the remainder when one number is divided by another in the given cells. **For example:** The value of the remainder or Mod obtained by dividing the number **56** by the numeric value **23** using the **MOD** function — **MOD (56, 23)** will be **10**.
- **Summing up Math and Trig Functions:**

NOTES

Function	Description
ABS	It is used for converting the given or supplied number into its absolute value.
ACOS	It is used for providing the value of a number in arccosine.
ACOSH	It is used for representing the inverse hyperbolic cosine of a number.
AGGREGATE	New function that has been added to Microsoft Excel 2010. It is used for calculating the aggregate value of a list of numbers or multiple arguments.
ASIN	It is used for providing the value of a number in arcsine.
ASINH	It is used for or representing the inverse hyperbolic sine of a number.
ATAN	It is used for providing the value of a number in arctangent.
ATAN2	It is used for representing the arctangent of the supplied x coordinate and y-coordinate.
CEILING	It is used for returning the number after rounding it up to the nearest possible integer number or value. The number is rounded up or down on the basis of the sign the number has.
CEILING PRECISE	It is used for returning the number by rounding it up to the nearest possible integer number or value. The number is always rounded up irrespective of the number's sign.
COMBIN	It is used for providing all the possible combinations of objects or group of objects for the specified or given number of values or items.
EVEN	It is used for returning the positive or negative number by rounding it up or down to the even number that is nearest to the supplied number.

NOTES

EXP	It is used for returning the number in exponential form by raising the power to the number supplied within the function.
FACT	It is used for calculating or finding out the factorial of a number.
FLOOR	It is used for returning the number by rounding it down to the nearest possible integer number or value. The number is rounded up or down on the basis of sign of the number.
FLOOR PRECISE	It is used for returning the number by rounding it down to the nearest possible integer number or value. The number is always rounded up irrespective of the number's sign.
GCD	It is used for calculating or providing the greatest common divisor of a given or supplied number.
INT	It is applied or used on a real or fractional number to convert it into the nearest possible integer.
LCM	It is used for calculating the least common multiple or smallest common factor of a given or supplied number.
MMULT	It is used for returning the product of two arrays as a matrix product.
MINVERSE	It is applied on a supplied array matrix to calculate or find out the inverse matrix.
ODD	It is used for returning the positive or negative number by rounding it up or down to the odd number that is nearest to the supplied number.
RAND	It is used for returning a number ranging between 0 to 1, randomly.
SQRT	It is applied to a given number to calculate or return the positive square root of that number.
SUMSQ	It is applied on the list of numbers to return to determine the sum of the squares of the numbers passed or referred.
TRUNC	It is used for truncating or trimming the number by removing or extracting the decimal points or fractional part from the number. Thus, it converts the given number to an integer number or value.

Statistical Functions

Statistical Function is present inside **More Functions** dropdown box. The user needs to browse **More Functions** in order to access **Statistical Function**. **Statistical** function plays an important role in Microsoft Excel. There are various functions available in **Statistical** function like **AVERAGE, MAX, MIN, COUNT, MEDIAN, SLOPE, SKEW, FREQUENCY**, etc.

The basic statistical functions are described as follows:

- **AVERAGE:** It gives the average of the given numbers.

The screenshot representing **AVERAGE** function is as shown:

A1		fx =AVERAGE(34,55)				
	A	B	C	D	E	
1	44.5					

NOTES

- **MAX:** It is a function used to return the maximum value from the values passed in arguments. The screenshot representing **MAX** function is as shown:

A1		fx =MAX(10,16,24,20,7)					
	A	B	C	D	E	F	
1	24						

- **MIN:** It is a function used to return the minimum value from the values passed in arguments. The screenshot representing **MIN** function is as shown:

A1		fx =MIN(10,16,24,20,7)					
	A	B	C	D	E	F	
1	7						

- **COUNT:** It returns the count of the number of arguments.

The screenshot representing **COUNT** function is as shown:

A1		fx =COUNT(10,16,24,20,7)					
	A	B	C	D	E	F	
1	5						

- **MEDIAN:** This function returns the middle number of the set of numbers. Syntax of **MEDIAN** function is—**MEDIAN** (number1, number2, ...), where, number 1 represents the first number and number 2 represents the second number. Each number is separated from the other with the help of a comma operator.

The middle number can be easily found in case of odd numbers but when the number is even, it is found by calculating the average of the middle two numbers.

- **AVEDEV:** This function calculates the average deviation of the given set from mean. Syntax of **AVEDEV** function is—**AVEDEV** (number1, number2, ...)

The equation for the average deviation is as:

$$\frac{1}{n} \sum |x - \bar{x}|$$

Where,

NOTES

n is the total number of values,

x takes each and every value which is present in the set,

\bar{x} bar is the average of groups of values.

Suppose, you have 5 numbers which are 1, 2, 3, 4 and 5. First the mean of five numbers is calculated, which is 3. Then the average deviation is calculated using above formula.

$$\begin{aligned}\frac{1}{n} \sum |x - \bar{x}| &= \frac{1}{5} (|(3-1)| + |(3-2)| + |(3-3)| + |(3-4)| + |(3-5)|) \\ &= \frac{1}{5} (2 + 1 + 0 + 1 + 2) \\ &= 1.2\end{aligned}$$

- **CORREL:** It calculates the Pearson Product- Moment Correlation Coefficient for the two sets of given values. Syntax of **CORREL** function is—**CORREL**(array1,array2,...), where, **array1** is cell range of the first cell and **array2** is cell range of the second cell.

The equation is as follows:

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Where,

\bar{x} and \bar{y} bar are mean of the group of values.

- **RANK:** This function gives statistical rank of the given value. In case there are duplicate values, it gives the same rank. The syntax for the calculation of **RANK** of the number is as:

RANK (number, ref, [order])

- **STDEV:** This is the function which calculates the standard deviation for a given set of values.

The equation is like:

$$\text{Sample Standard Deviation} = \sqrt{\frac{\sum (x - \bar{x})^2}{(n-1)}}$$

Where,

x is the group of values and n is the total number of values.

• Summing up Statistical Functions:

Functions	Description
AVERAGEIF	It is used to calculate the average according to certain criteria.
AVERAGEIFS	It is used to calculate the average according to a set of criteria.
BETA:DIST	It is used for returning beta cumulative distribution function.

BETA:INV	It is used for returning the inverse of the beta cumulative distribution function.
BINOM:DIST	It is used for returning binomial distribution probability.
BINOM:INV	It is used for returning the inverse of the binomial distribution probability.
CHISQ:DIST	It is used for returning chi-squared distribution.
CHISQ:DIST.RT	It is used for returning chi-squared distribution of right tailed probability.
CHISQ.INV	It is used for returning the inverse of chi-squared distribution.
CHISQ.INV.RT	It is used for returning the inverse chi-squared distribution of right tailed probability.
CHISQ.TEST	It is used for returning the values of chi-squared distribution.
COUNTA	It is used for returning total number of non-blanks in the cells.
COUNTBLANK	It is used for returning total number of blanks in the cells.
COUNTIF	It is used for returning the total number of cells on the basis of certain criteria.
COUNTIFS	It is used for returning total number of cells on the basis of set of criteria.
COVARIANCE.P	It is used for returning population covariance.
DEVSQ	It is used to perform the summation of the square of the deviations.
EXPON.DIST	It is used for returning exponential distribution.
F.INV	It is used for returning the inverse of F probability distribution.
FDIST.RT	It is used for returning the right tail F probability distribution.
FORECAST	It is used for predicting future values from a given value.
FREQUENCY	It is used for finding the repetition of values within a range.
GEOMEAN	It is used to perform the geometric mean.
HARMEAN	It is used to perform the harmonic mean.
KURT	It is used to perform kurtosis of the given data set.
LARGE	It is used for finding the largest element present in the data set.
LINEST	It is used for returning statistical information determining the graph line on the basis of values of x and y coordinates.
LOGEST	It is used for returning the exponential trend parameter.
LOGNORM.DIST	It is used for returning log normal distribution.
MAXA	It is used for returning greatest value from the argument list.
MINA	It is used for returning smallest value from the argument list.
MODE.MULT	It is used for returning frequently used values.

NOTES

NOTES

NORM.DIST	It is used for returning normal cumulative distribution of mean and deviation.
NORM.INV	It is used for returning inverse of the normal cumulative distribution.
PERCENTRANK	It is used to find the percentage rank of any element, present in the data set.
PERCENTILE	It is used to find data set percentile.
POISSON.DIST	It is used for returning position distribution.
QUARTILE	It is used to find data set quartile.
STDEV.P	It is used to find data set population standard deviation.
STDEV.S	It is used for computing standard deviation on the basis of population sample
SMALL	It is used to find the smallest element which is present in the data set.
TREND	It is used for returning linear trend values.
TRIMMEAN	It excludes the smallest and largest element in the data set and finds the mean.
VAR	It is used to find the variance of the given data set.
VARP	It is used to find the population variance of the given data set.

5.4.5 Working with Charts

Let us analyse the different tools of working in a chart.

Chart

A chart in Excel refers to the tool that helps in representing data in the graphical form. It makes it easier for the audience or a user to understand the task in simple language. It helps in representing the comparison between the same types of data. There are many functions that can be performed with charts. Let's learn to create a chart.

Steps for **creating a chart** are as follows:

1. Select the cells (rows and columns) containing data for which the chart is to be created.

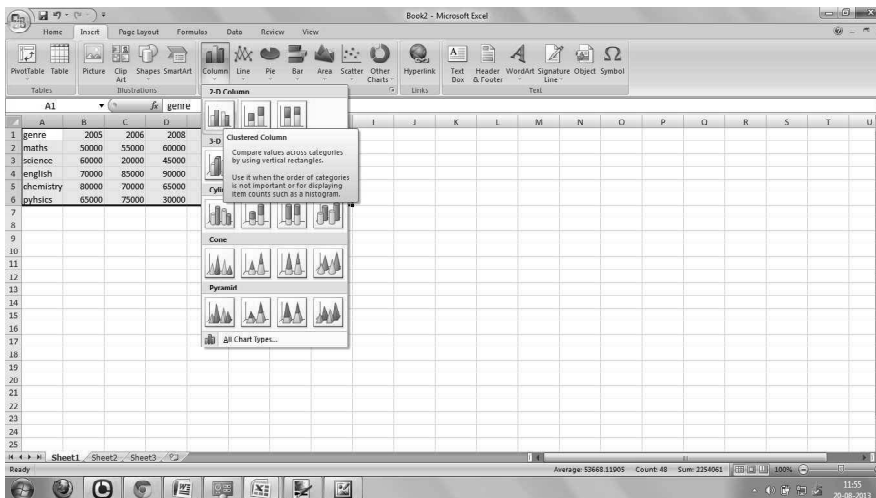
The screenshot to **select all the cells** of the table is as shown:

NOTES

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	genre	2005	2006	2008	2009	2010	2011	2012													
2	maths	50000	55000	60000	70000	75000	80000	85000													
3	science	60000	20000	45000	50000	55000	60000	65000													
4	english	70000	85000	90000	80000	70000	60000	55000													
5	chemistry	80000	70000	65000	60000	75000	80000	55000													
6	physics	65000	75000	30000	35000	45000	70000	75000													

2. Go to **Insert** tab, select the chart type which you want to insert from the **Charts** tab.

The screenshot displaying **different chart types** is as shown:



3. Select the chart of your choice. It gets displayed in the worksheet.

The screenshot to **display the chart** of the table is as shown:

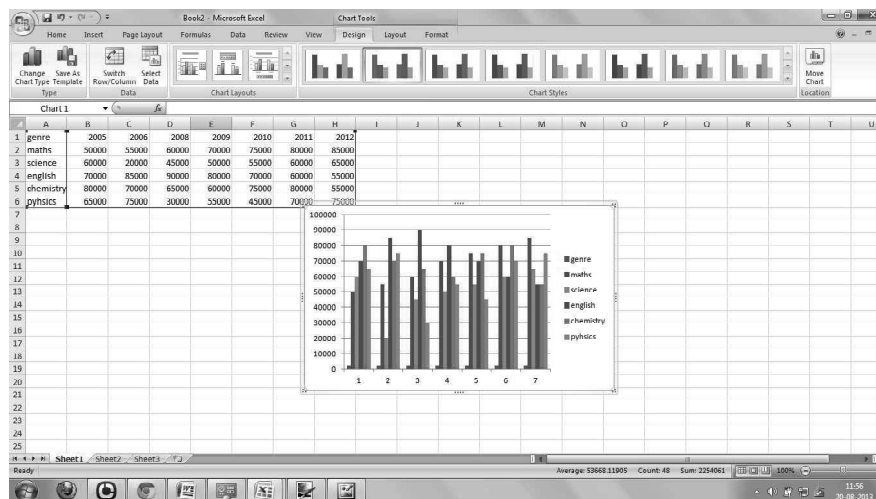


Chart Tools

Chart tools are the tools that can be used to modify the inserted chart. It appears when the chart is selected in the worksheet.

NOTES

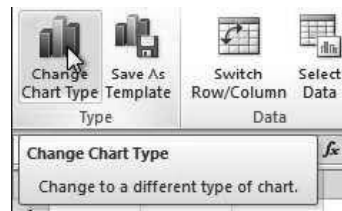
The screenshot displaying the **Chart Tools** is as shown:



Steps for **changing the chart type** are as follows:

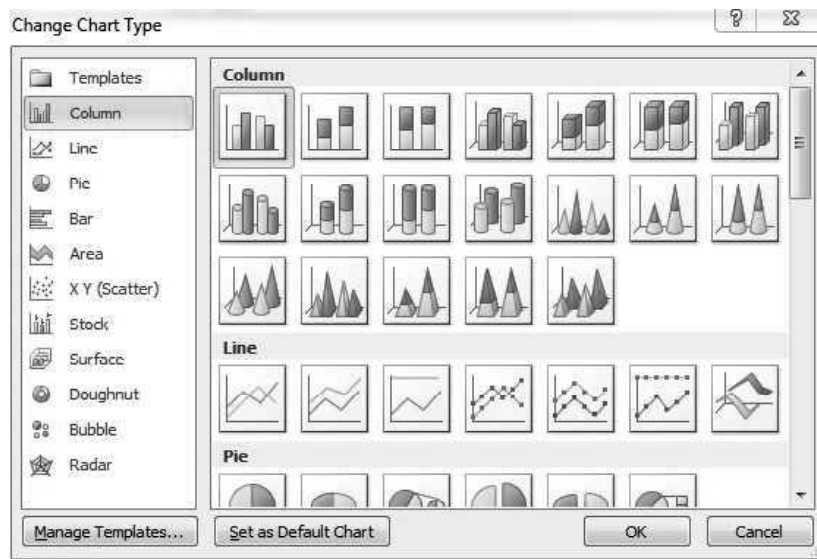
1. Go to **Design** tab and select **Change Chart Type** option.

The screenshot displaying the **Change Chart Type** option is as shown:



2. A **Change Chart Type** dialog box gets displayed on the screen.
3. Choose any chart from the list of chart type, and then press **OK**.

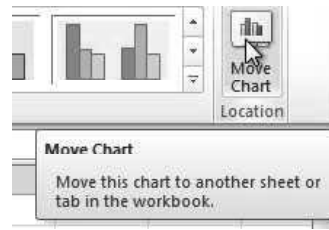
The screenshot to **choose a chart** from the **Change Chart Type** dialog box is as shown:



Steps to **move chart** to different worksheets are as follows:

1. Go to **Design** tab, select **Move Chart** command. A dialog box appears.

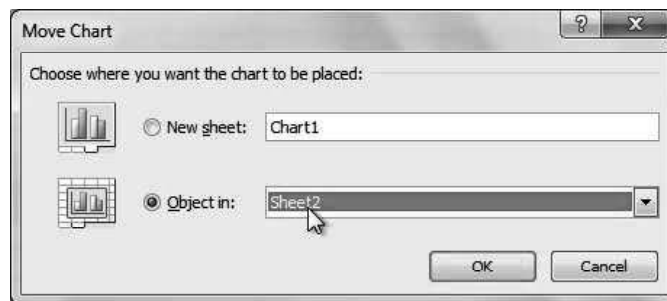
The screenshot to display **Move Chart** option is as shown:



NOTES

2. Select the location where you want to move the chart by filling in the text box **Object in**.

The screenshot to place the chart to the location as per the user choice is as shown:

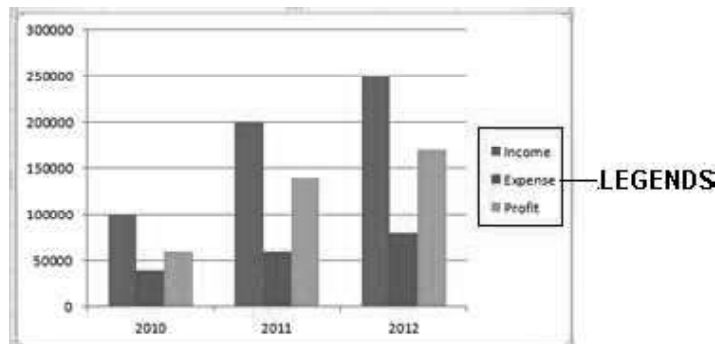


Legends

Legend refers to information about the elements of chart. Legends enable us to understand what information is given by different colors and trend lines.

For example: If a record set has three fields, then each field in chart is represented by a different color.

Snapshot of a chart which has a legend is as shown:



There are three fields shown in this chart: Income, Expense and Profit. Each is represented by a different color.

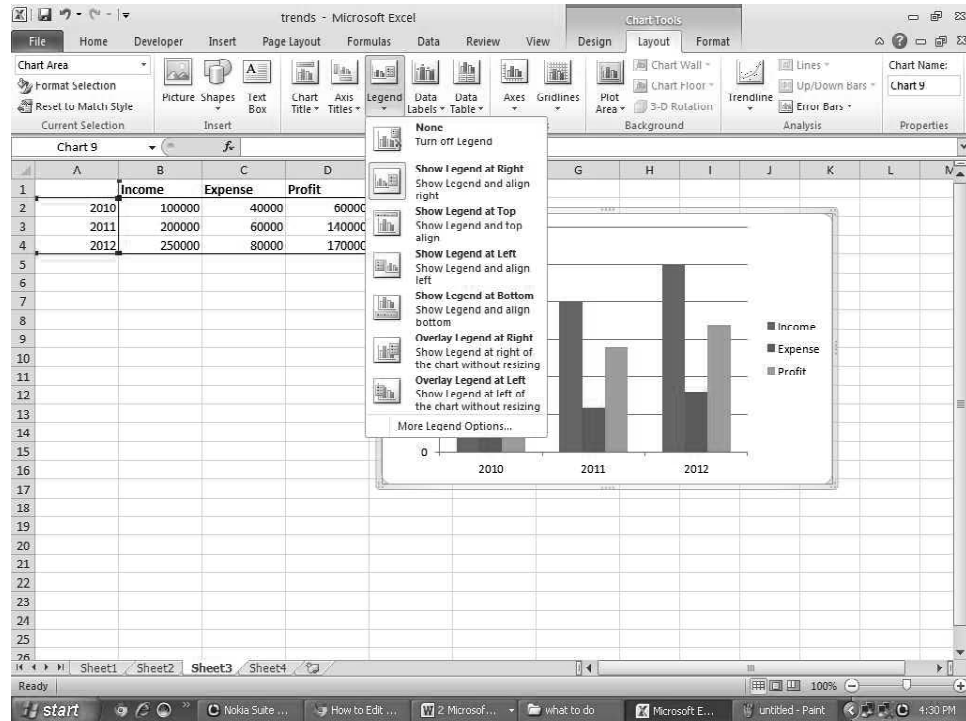
Legend Options

Steps for formatting or setting legend as per requirement, are as follows:

1. Select chart.
2. Click on **Layout** to switch on **Layout** tab.
3. In this tab, you will find **Legend** option under **Labels** group.

NOTES

Snapshot of **Legend** in **Labels** group is as shown:



In this option, you will find various choices to set the legend at the left, right, top and bottom of the chart or as **Overlay Legend**. You can change the style as well, by clicking on **More Legend Options** under the same popup menu.

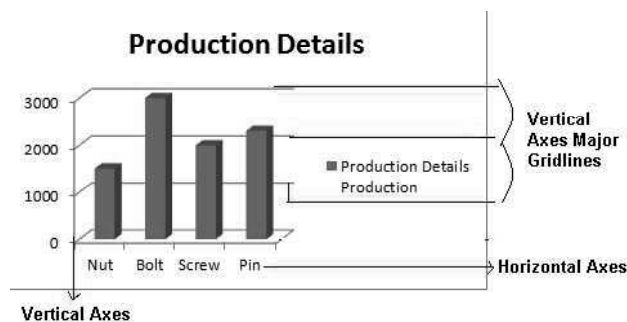
You find the **Format Legend** window also on the right click popup menu of **Legend**.

Formatting Axes and Gridlines Within Charts And Graphs

A chart is a pictorial view of the numeric data contained by a sheet. In Excel 2010, you will find various types of charts in the **Charts** group of **Insert** tab. These can be easily inserted in the worksheet.

Excel provides 2D charts as well as 3D charts. You can insert them according to your need.

Snapshot of a **Column** type **3D** chart is as shown:



Formatting Axes

Steps needed to format the axes of the chart are as follows:

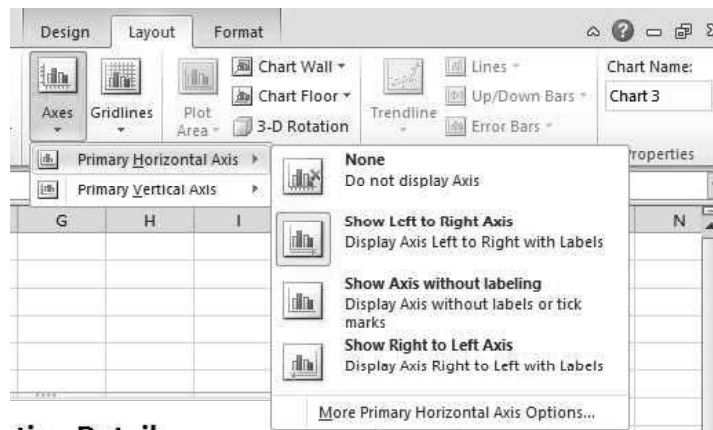
1. First select the chart by clicking on it.
2. Click on **Layout** under the menu bar in order to show the **Layout** tab.

Screenshot of **Layout** tab is as shown:



3. Click on **Axes** option under **Axes** group.

Screenshot of sub-options included under **Axes** option are as shown:



4. Within **Axes** option, you will find various choices under both **Primary Horizontal Axes** and **Primary Vertical Axes**.

Primary Horizontal Axes

Within this head, you will find the following options:

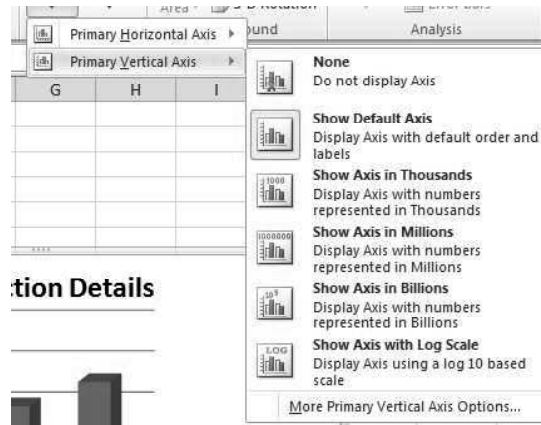
1. **None:** It is used when there is no need for the axes to be displayed. If you click on this option, the horizontal axes will be removed.
2. **Show Left to Right Axes:** The chart presented above shows this option, as the **Vertical Axes** is on the left moving to the right side.
3. **Show Axes without Labeling:** By applying this option, you can remove only the labels of the axes. It has almost the same effect as the **None** option.
4. **Show Right to Left Axes:** Vertical axis is moved to the right of the chart and the labeling of horizontal axes starts from right, moving to left.
5. **More Horizontal Axes Options:** By clicking on this you will find a **Format Axes** dialog box, enabling you to change the appearance of the axes. In left pane of this dialog box you will find various formatting options for applying on the axes.

Primary Vertical Axes

Screenshot of sub-options of **Primary Vertical Axes** is as shown:

NOTES

NOTES



Within this head, you will find following options:

1. **None:** It is used when there is no need for the axes to be displayed. If you click on this option, the vertical axes will be removed.
2. **Show Default Axis:** It will show the default labels and orders.
3. **Show Axis in Thousands:** It will show number format in thousands.
4. **Show Axis in Millions:** It will show number format in millions.
5. **Show Axis in Billions:** It will show number format in billions.
6. **Show Axis with Log Scale:** It will show the log 10 based scale format of the numeric values.
7. **More Vertical Axis Options:** On clicking on this you will find a **Format Axis** dialog box, to set the values and apply formatting options.

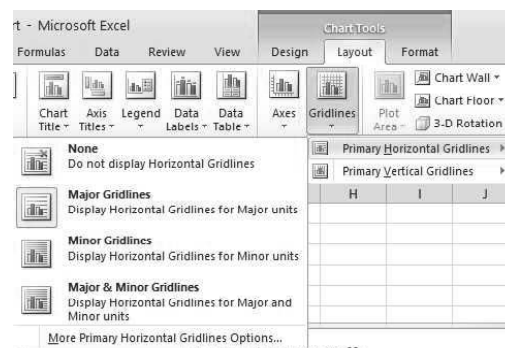
Formatting Gridlines

A chart can have Horizontal as well as Vertical Gridlines. You can customize both of them.

Steps needed to customize the gridlines are as follows:

1. Select the chart by clicking on it.
2. Click on **Gridlines** from **Axes** group of the **Layout** tab. You will find two options: **Primary Horizontal Gridlines** and **Primary Vertical Gridlines**.

Snapshot of sub-options of **Gridlines** is as shown:

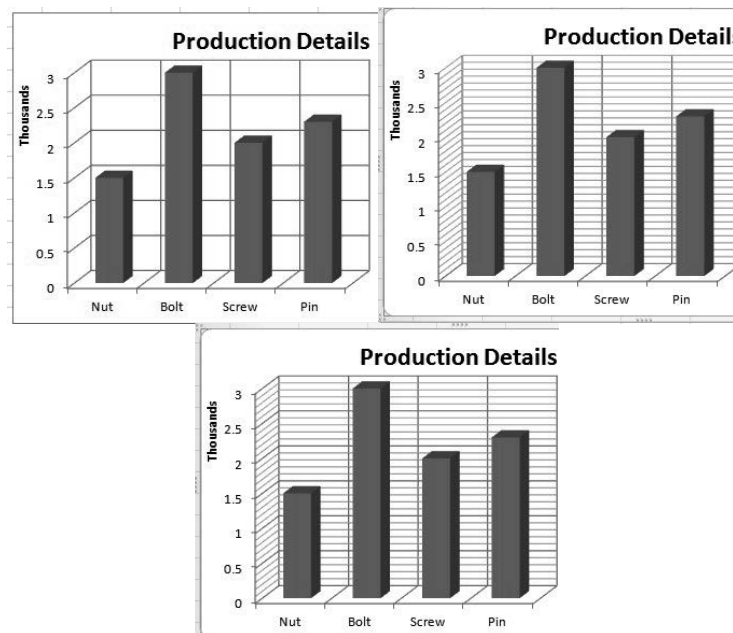


3. In both of the gridlines you will find same options. The options contained under these heads are:

- **None:** Click on it if you don't want to display gridlines.
- **Major Gridlines:** This option helps display gridlines with major units. The gap between gridlines is more than that between minor gridlines.
- **Minor Gridlines:** Click on it to display gridlines with minor units. The gap between gridlines is less.
- **Major and Minor Gridlines:** This option helps display Major as well as Minor Gridlines. Major Gridlines will be shown with a slightly dark colour.

NOTES

Snapshots of **Major**, **Minor** and **Major and Minor Gridlines** for Horizontal Gridlines are as shown:



- **Major Primary Gridlines Options:** Click on it to show a **Major Format Gridlines** Dialog, enabling you to format and color the gridlines.

Setting Up Live Charts

Live charts are charts which display, changing situations graphically. They help to easily understand the current situation. Live Charts are used to compare situations in different scenarios, to enable analysis.

For example: The rise and falls of securities listed in stock market could easily be displayed on the charts, to understand the actual condition.

These charts are prepared for displaying values of different fields, for any small or large record-set. If new values are added to that record set, the charts automatically show these newer changes. Hence, the end users need not concern themselves with the record-set.

NOTES

For understanding this better, let's take an example. Suppose, there is a Cricket record-set containing the runs of Team 2, after the completion of each over. The first team has already completed its turn and information about the runs they have scored has already been inserted in that record-set.

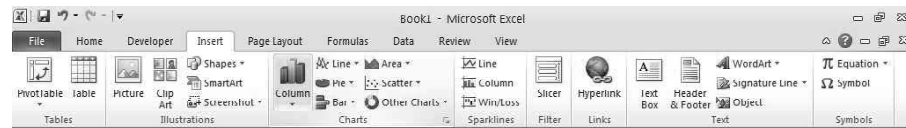
Screenshot of that database is as shown:

	A	B	C
1	Overs	Team 1	Team 2
2	1	5	7
3	2	8	6
4	3	11	5
5	4	14	4
6	5	12	
7	6	10	
8	7	8	
9	8	9	
10	9	10	
11	10	11	

Only ten overs play has been recorded in this example, but as you know the actual game will produce a huge database. Therefore, a pictorial representation of this information would make it easier for viewers to compare and understand the positions of both teams.

Excel provides various types of charts for showing any record-set graphically. Different types of charts are found under **Charts** group of **Insert** tab.

Screenshot of **Charts** group under **Insert** tab is as shown:



For comparing two related terms, the **Column** and **Bar Charts** are considered the best.

Steps to insert these charts are as follows:

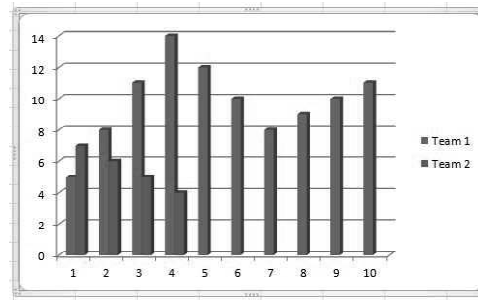
1. Select the fields to be displayed in the chart.

Screenshot of the selected fields is as shown:

	A	B	C
1	Overs	Team 1	Team 2
2	1	5	7
3	2	8	6
4	3	11	5
5	4	14	4
6	5	12	
7	6	10	
8	7	8	
9	8	9	
10	9	10	
11	10	11	

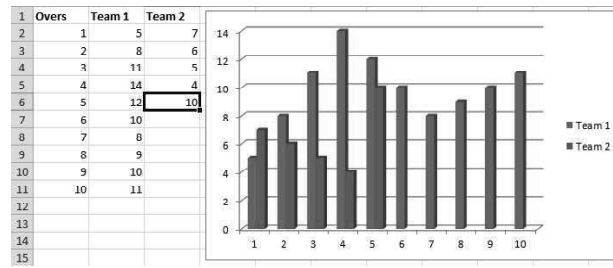
2. Click on either **Column** or **Bar** chart under the **Insert** tab. (You can also insert **Line** chart)
3. Click on any of the charts displayed in this category.

Screenshot of a **3-D Column** chart is as shown:



A live chart presents the recent changes in the record-set as well. In above example, when second team would complete its 6th over, the runs will be added to the record-set. As soon as the runs are added, the chart will automatically display it.

Screenshot for showing automatic modification in chart, as per the changes in record-set is as shown:



When you will select a chart, three new tab options will appear: **Design**, **Layout** and **Format**. You can use any of these options for modifying and formatting the inserted chart. Some of the important options for making the live charts more understandable are discussed below:

1. **Chart Layout:** For adding a chart title, you can go to **Design** tab and click on any of the pictures displayed under **Chart Layout** group.
2. **Move Chart:** For displaying a chart in different locations, you can click on **Move Chart** button of **Layout** group under **Design** tab.
3. **Select Data:** For selecting any other data range or modifying legends and axes labels for a particular chart, you can click on **Select Data** of **Data** group under **Design** tab.

Find, Replace and Go to

Excel 2010 provides some options to find specific data, replace previous data with new data, and go to specific cell. These are done by **Find**, **Replace** and **Go to** command. **Find** command is used to find quickly specific text, number in worksheet. **Replace** command is used to find and replace any data with new data. **Go to** command is used to jump on a specific cell or range of cells. These commands are available on **Editing** group on **Home** tab. Let us learn these commands one by one in detail. First, we will learn to find text or number in worksheet.

NOTES

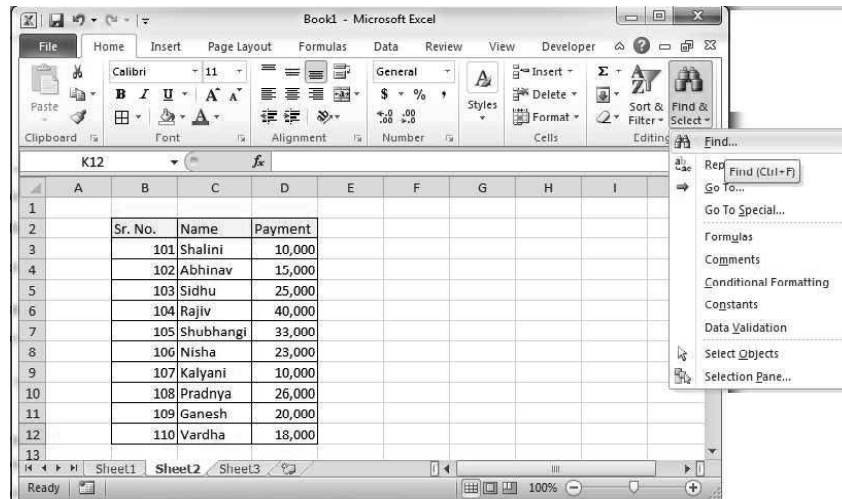
NOTES

Find:

Steps to **Find** text or number are as follows:

1. Click on **Editing** group on **Home** tab
2. Click **Find & Select** on **Editing** group.
3. Click on the **Find** command.

The screenshot of **Find & Select** is as shown:



4. On clicking, **Find and Replace** dialog box appears. Type a text or number that you want to find.

The screenshot displaying **107** number in **Find what** box is as shown:



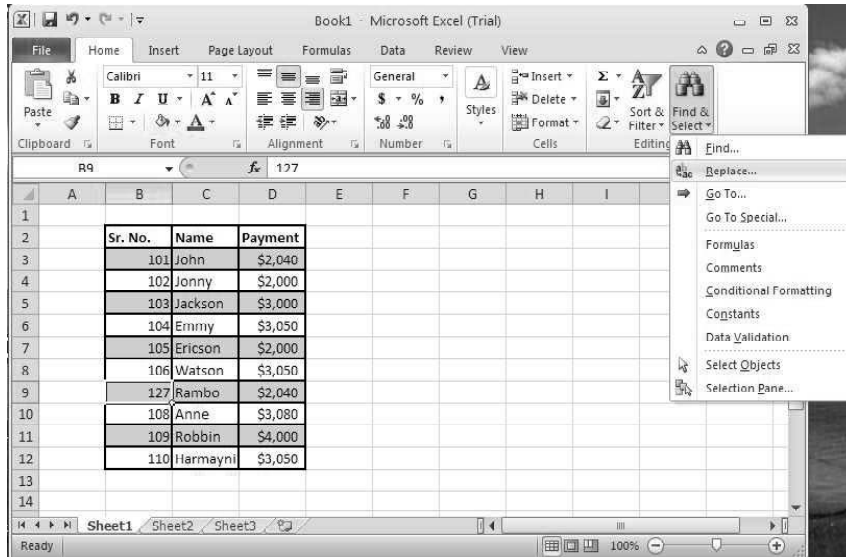
5. Click on **Find Next** button. It selects the cell where 107 numbers placed.
6. If you click on **Find All** button, the location of all the occurrences of that number or text is displayed.

Replace:

Steps to **Replace** text or number are as follows:

1. Click on **Editing** group on **Home** tab
2. Click **Find & Select** on **Editing** group.
3. Click on the **Replace** command.

The screenshot of **Replace** is as shown:



NOTES

- On clicking **Replace**, **Find and Replace** dialog box appears. Type a text or number in **Find what** and **Replace with** box.

The screenshot of displaying **107** number in **Find what** and **127** number in **Replace with** textbox is as shown:



- Click on **Find Next** button. It will select cell where **107** number is placed.
- Click on **Replace** button. The number **107** is replaced with **127**.

The screenshot of table after replacing **107** with **127** using **Replace** button is as shown:

Sr. No.	Name	Payment
101	John	\$2,040
102	Jonny	\$2,000
103	Jackson	\$3,000
104	Emmy	\$3,050
105	Ericson	\$2,000
106	Watson	\$3,050
127	Rambo	\$2,040
108	Anne	\$3,080
109	Robbin	\$4,000
110	Harmayni	\$3,050

- If you click on **Replace all**, it will replace all the occurrences of the **107** and replace it with **127**.

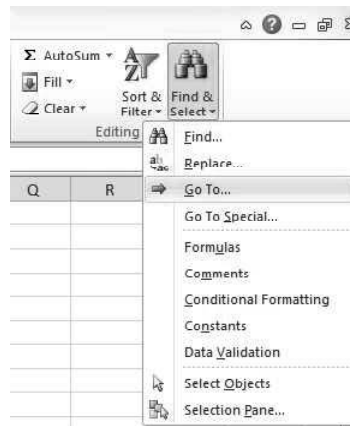
NOTES

Go To:

Steps to use **Go To** command are as follows:

1. Click on **Editing** group on **Home** tab.
2. Click **Find & Select**.
3. Click on the **Go To** command.

The screenshot of **Find & Select** is as shown:



4. On clicking **Go To**, **Go To** dialog box appears that prompts to enter reference. Enter the cell reference where you want to jump, say **D3:D7**.

The screenshot of **Go To** dialog box is as shown:



5. Click **OK** button. The cells **D3 to D7** are selected.

Screenshot of table after clicking **OK** button is as shown:

	A	B	C	D
1				
2		Sr. No.	Name	Payment
3		101	John	\$2,040
4		102	Jonny	\$2,000
5		103	Jackson	\$3,000
6		104	Emmy	\$3,050
7		105	Ericson	\$2,000
8		106	Watson	\$3,050
9		127	Rambo	\$2,040
10		108	Anne	\$3,080
11		109	Robbin	\$4,000
12		110	Harmayni	\$3,050

5.4.6 Modifying and Merging Rows and Columns

In the MS Excel worksheet, by default the height and width of the cells in every row and column is similar. The default Column Width is '8.43' and Row Height is '15'. MS Excel provides different methods for modifying column width and row height including wrapping of text and merging of cells.

NOTES

Modifying Column Width

It becomes essential to modify the width of a column in the worksheet when the complete cell content is not visible or displayed. To make it visible you have to increase the width of the column. Additionally, if the width of the cell is more or extra and the cell content is less then you can decrease the width of the cell to give an appropriate width to the worksheet column.

1. To change or modify the column width, position the mouse pointer or cursor over the column line in the column heading. The mouse pointer or cursor will now change from the White Cross to Double-Arrow.
2. Click, hold and drag the mouse pointer for increasing or decreasing the column width.
3. Now release the mouse pointer. See the column width of the selected column is increased and changed and the complete cell content is visible. Similarly increase the column width also so that the complete cell content will be visible. You can also change the column width of range of columns together by selecting them and applying the command.
4. While entering numeric data or date if the cell displays the pound signs (#####) then it specifies that the column width is not sufficient or appropriate for displaying the cell content. To view the complete cell content, increase the width of the column.
6. Alternatively, you can also increase or decrease the width of the column by selecting Format command in the Cells group on the Home Tab. If you select the option 'Column Width...' then the Column Width dialog box will appear. Enter the new value for changing the column width and press 'OK'. You can also select the option 'AutoFit Column Width' or 'Default Width...'. When you select the option 'AutoFit Column Width' then MS Excel will automatically increase or decrease the Column Width as per the cell content.
7. When you press 'OK' then the column width will be increased.

AutoFit Column Width

As discussed, the AutoFit Column Width feature automatically set's the column width to fit its content. One method we have already discussed that is using the Format option. Alternatively, select the cell for formatting and position the mouse cursor or pointer above the column line in the column heading. The White Cross Cursor will now become a Double-Arrow Cursor. Now when you double-click the mouse then the Column Width will automatically change to fit the cell content. Similarly, you can change the Row Height.

Modifying Row Height

You have learnt how to modify the Column Width of a cell. Similarly, the Row Height of a cell can also be modified.

NOTES

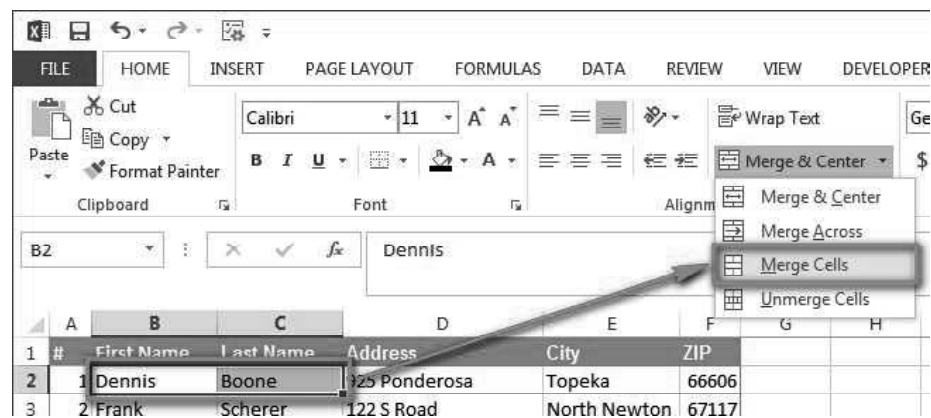
1. To change or modify the row height, position the mouse pointer or cursor over the row line in the row heading. The mouse pointer or cursor will change from the White Cross to Double Arrow.
2. Click, hold and drag the mouse pointer for increasing or decreasing the row height.
3. Now release the mouse pointer. See the row height of the selected Row 2 is increased and changed. Similarly you can increase the row height of the other rows in the worksheet.
4. Alternatively, you can also increase or decrease the height of the row by selecting Format command in the Cells group on the Home Tab. If you select the option 'Row Height...' then the Row Height dialog box will appear. Enter the new value for changing the row height and press 'OK'. You can also select the option 'AutoFit Row Height'.

When you select the option 'AutoFit Row Height' then MS Excel will automatically increase or decrease the Row Height as per the cell content.

Merging Rows and Columns

For merging rows and columns, follow the following steps.

1. Select the cells that you wish to merge.
2. Click on merge and center command to open drop down menu in the alignment group of Home tab.
3. Click on the option as per your requirement. An illustration is given below.



Adding Spellcheck in Spreadsheet

To add spellcheck in spreadsheet, follow the following steps.

1. Click the **Review** tab on the ribbon.
2. Click **Spelling**.



NOTES

3. If the program finds spelling mistakes, a dialog box appears with the first misspelled word found by the spelling checker.
4. After you decide how to resolve the misspelling (ignoring it, adding it to the program's dictionary, or changing it), the program moves to the next misspelled word.

Alternatively, you can press F7 key to apply spellchecker.

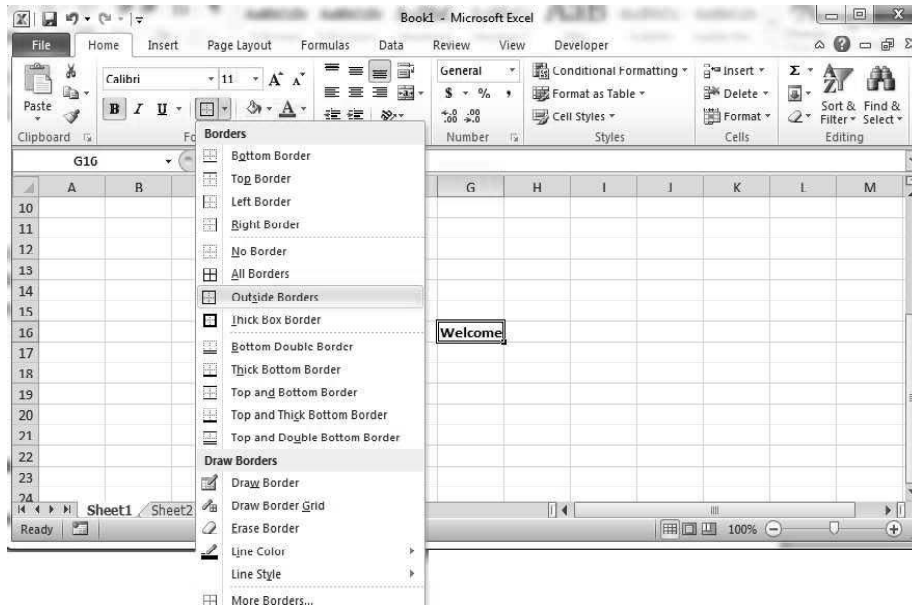
Adding Borders and Colors to Cells

To set border

Steps to set **border** are as follows:

1. Select the cell.
2. Go to **Home** tab.
3. Click on **Borders** option in **Font** group.
4. Select Border style that you want to use.

The screenshot to select **Borders** is as shown:



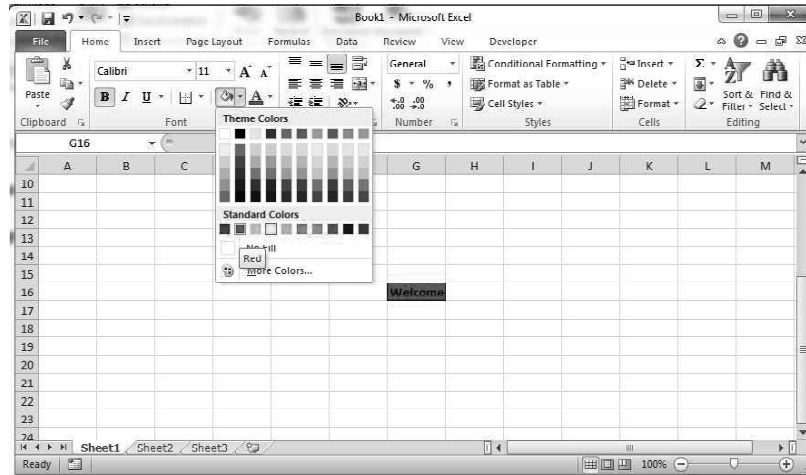
NOTES

To set Fill cell colour

Steps to **Fill cell colour** are as follows:

1. Select the cell.
2. Go to the **Home** tab.
3. Click on **Fill colour** option in **Font** group.
4. Select color that you want to fill in selected cell.

The screenshot to select **Colours** is as shown:

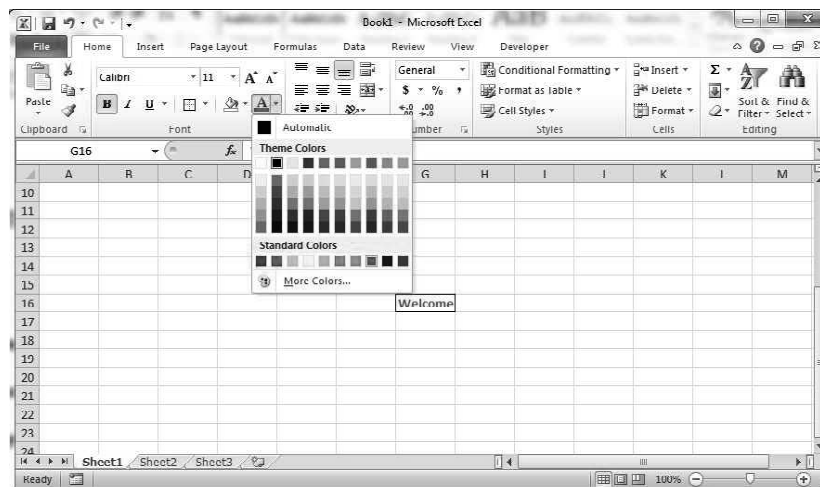


To set the Font colour

Steps to set the **Font colour** are as follows:

1. Select the cell.
2. Click on **Home** tab.
3. Click on **Font colour** option in **Font** group.
4. Select color that you want to set in selected cell. Text in cell gets displayed with selected color.

The screenshot to select **Font colour** is as shown:

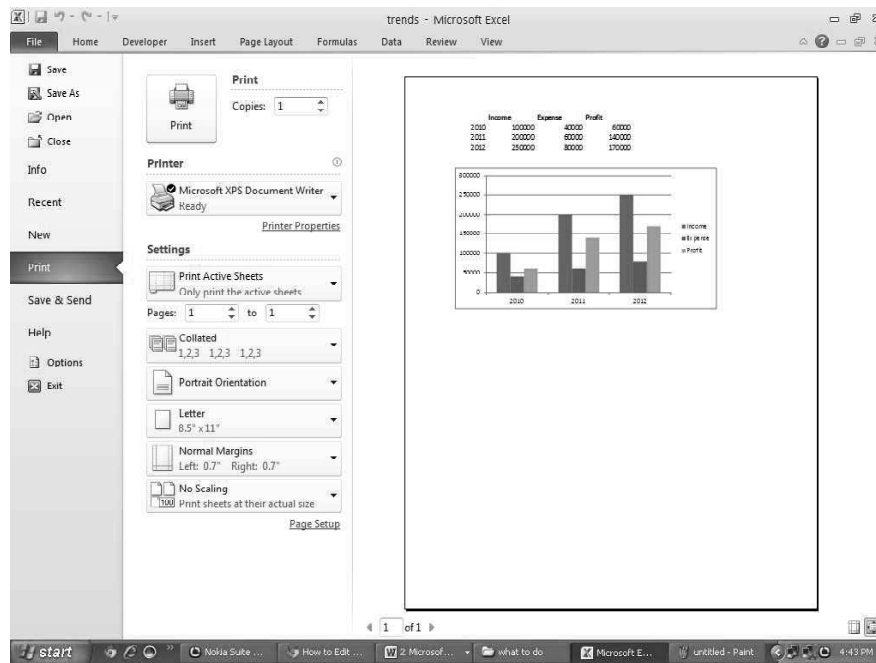


Printing a Worksheet/Workbook

Windows, MS Office and
Outlook 2010

You can print all pages of a worksheet or only selected page(s) by using **Print** option of **File** menu.

Snapshot of **Print** options is as shown:



Before printing, you can set print styles as per your requirement. While printing a document, there are various settings available. These are as follows:

1. **Copies:** You can set number of copies to be printed as well as select the pages to be printed.
2. **Printer:** Select printer, if there are more than one printer attached. Otherwise, Excel will automatically recognize the attached printer.
3. **Settings:** In settings, you will find the following options:
 - A. **Print** option, which gives you three choices— **print active sheet**, **entire workbook** or **print selection**. You can select whichever you need and see the preview of the printing layout on the right side.
 - B. **Pages:** Here, you can set the range of pages to be printed..
 - C. **Collated** refers to printing the pages one by one and **Uncollated** means printing the required copies of first page, then those of second page and so on.
 - D. **Portrait** and **Landscape Orientation** options leave you to decide whether you want to print the pages lengthwise or widthwise, respectively.
 - E. You can set the **Page Type** for the print to be taken. In this snapshot, **Letter** size is selected.
 - F. You can set **Margin** as well. By default, it is set as **Left: 0.7" Right: 0.7"**.

NOTES

G You will find some scaling options as well, which are helpful if you want to compress or enlarge the size of the content.

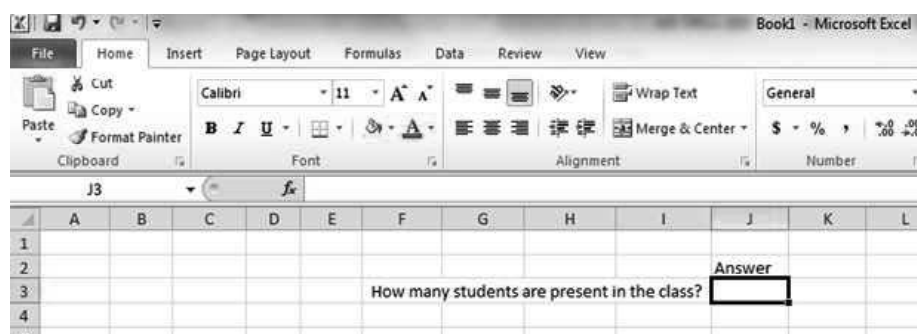
At last, you click on **Print** button on top, to take the final print out(s).

NOTES

Validating and Analyzing Worksheet Data

Validation is important aspect of MS Excel; it provides the facilities to enter only those values as input in the cell which are appropriate. For example: If in a worksheet, numbers between 1 to 10 are required to be entered, then, you must not enter number greater than 10 or less than 1. Also, only integer values should be present in the cell. Validating provides constraints that eliminate the chances of entering wrong data and hence increases the reliability of the worksheet.

The screenshot displaying an example of validation is as shown:



As shown in the screenshot, in the questions like “How many students are present in the class?” The number must be less than or equal to total number of students in class and greater than or equal to zero. In validating and analyzing, there are certain terms that require understanding. These are explained in the succeeding sections.

Conditional Formatting

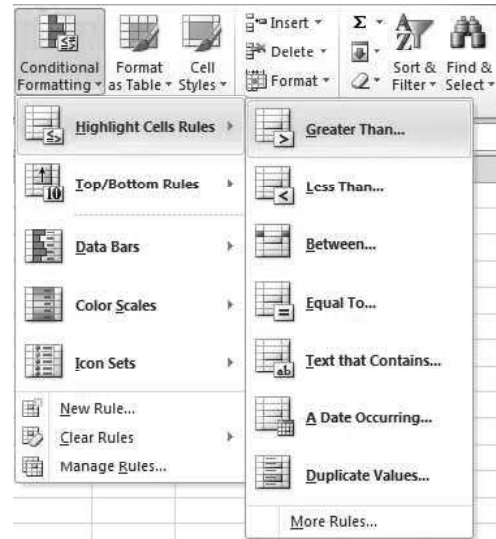
Conditional formatting is a tool of formatting a cell of the worksheet that contains some condition. In the conditional formatting there are different types of rules, bars of option, color scale, etc. All these tools can be used to make the worksheet formatted according to the condition provided by the user. It helps the user to analyze the data in worksheet. You can use the conditional formatting to highlight or change the color of the cell. There are many mathematical relations present in the MS Excel 2010 such as greater than less than, between, equal to, etc.

Steps to apply conditional formatting are as follows:

1. Click on **Home** tab.
2. Click on **Conditional Formatting**.

The screenshot of selecting the **Conditional formatting** is as shown:

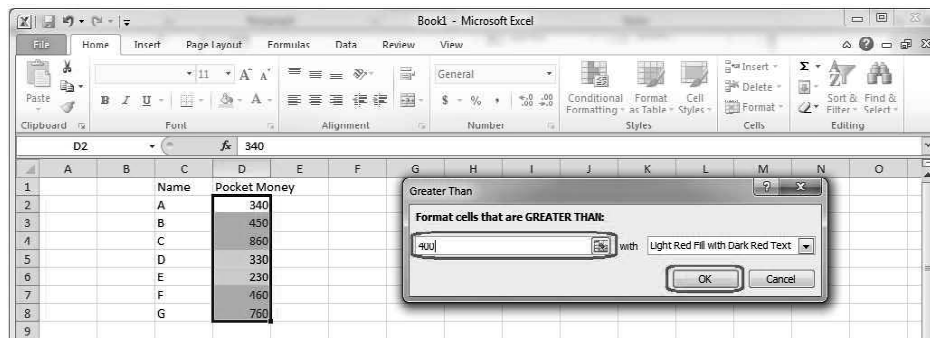
Windows, MS Office and Outlook 2010



NOTES

3. Select **Highlight Cells Rules** and then, select **Greater Than**.
4. Fill the value in **Greater Than** Text area.
5. Click **OK** button.

The screenshot of output displaying the cell, outlined with red color, which contains the value greater than **400** is as shown:



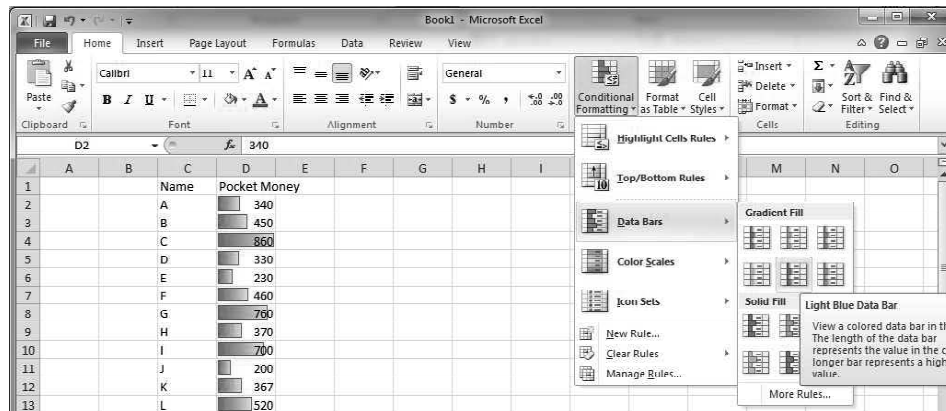
Other options like Data bars are used to see the difference between the values in the cell of the worksheet.

Steps to format the worksheet with a bar are as follows:

1. Click on **Home** tab.
2. Click on **Conditional Formatting**.
3. Select **Data Bars**.
4. Click on the different Color bar.

The screenshot of **Conditional Formatting** is as shown:

NOTES



In the preceding screenshot, the length of bars in each cell is according to the value of the number. The longest bar appears in the cell containing the largest number, which in this case is 860.

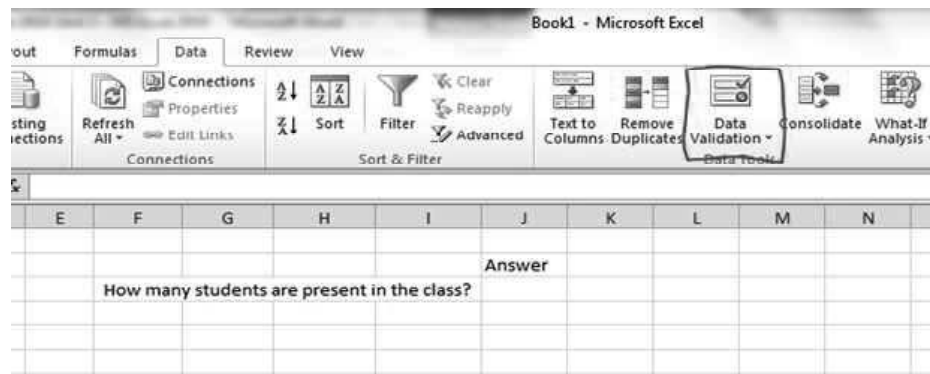
Setting up Validation

Validation is the process of checking the values for desired type, that is, it provides the facilities to enter only those values as input in the cell which are appropriate. It is very useful in official documents or as database for the program. Validation is very important where accurate data must be entered in the cell of the worksheet.

Steps to validate the worksheet are as follows:

1. Click on **Data** tab.
2. Click on **Data Validation** in **Data Tools**.

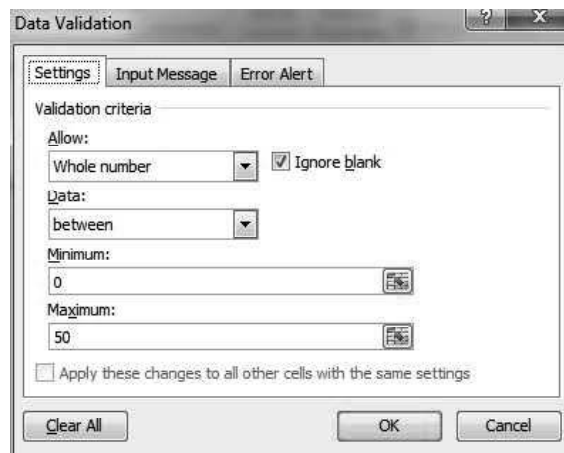
The screenshot of **setting the validation** is as shown:



3. Click on **Settings tab** of the **Data Validation**.
4. Select data type wants to allow in cell. In this case select **Whole number**.
5. Select the **Ignore blank** check box to ignore the blank cell.
6. Select the **Data** either greater than, less than, between, equal to, not equal to, greater than or equal to, less than or equal to. In this case select **between**.

7. Give the **Minimum** and **Maximum** number which is allowed in worksheet.
8. Click on **Input Message** Tab.
9. Select the check box to show the message when cell is selected.
10. Write in **Title** area to set the title of message.
11. Write the **Input message** in input area to show the message when user selects the cell.
12. Click on **Error Alert** tab.
13. Select the check box to show error after inserting the invalid output.
14. Select the **Style** of error message from options stop, information and warning.
15. Write the **Title** for the Error Alert.
16. Write the error message in **Error Message** area which gets displayed when error occurs.
17. Select **OK** to save the setting or **Clear All** to remove the setting or **Cancel** to keep previous setting on.

The screenshot of **Data Validation** dialog box is as shown:



Checking the Validation:

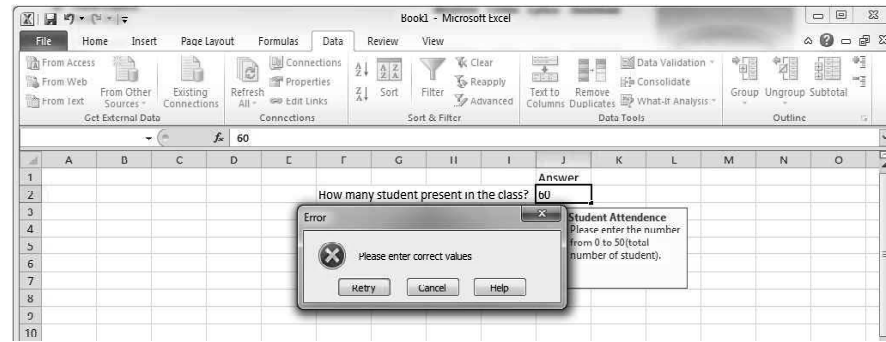
After modifying the setting of validation, we must check whether the setting is working properly or not. So, follow these steps to check the validation:

1. Select the Cell on which validation is required to be done.
2. Write the value in the range 0 to 50. If the value entered is correct, then no error message appears.
3. Enter a value either greater than or less than the range. We can also give alphabetic word to check the validation. Let's enter 60 in cell.

The snapshot of checking the validation is shown as:

NOTES

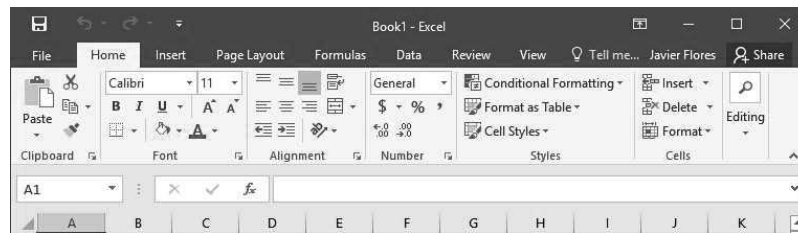
NOTES



If for the values do not lie in the range 0 to 50, error message gets displayed, which means that the validation is successful.

New Features in Excel 2016/ Difference between Excel 2010 and 2016

Excepting superficial colour and labelling, Microsoft Excel 2016 is not much different from Excel 2013 or 2010. All the menu items and often-used features are the same. The screenshot of ribbon interface is given below.



However, new features of Excel 2016 include:

- **Pulling in data from almost anywhere:** A former add-on called Power Query is now built into Excel. Power Query can pull data from various sources for analysis.
- **More charting options:** Excel also has six new chart types.
 - o Box and whisker charts
 - o Pareto charts
 - o Sunburst charts
 - o Histograms
 - o Treemap charts
 - o Waterfall charts
- **Predicting the future:** The new time series forecasting functions can predict future values from historical data. The more data, the better the forecast.
- **Ink Equations:** It allows you to 'write' equations on a tablet device using a stylus or your finger and then converts them to text. If you are not using a touchscreen device, you can do the same using a mouse.
- **Features common with Word 2016:** Tell Me, Smart Look Up and Insight options.

Check Your Progress

1. Windows 10 operating system.
2. When is the .dotx file format used?
3. When is the ignore option used in MS Word?
4. State the uses of header and footer.
5. What is file management?
6. List the steps to set border.
7. What is conditional formatting?

NOTES

5.5 MICROSOFT ACCESS 2010

Microsoft Office Access 2010 or MS Access 2010 is a most popular database management system tool for windows. It combines the object-oriented relational database management system engine with graphical user interface and software-development tools from Microsoft.

Microsoft Access 2010 has some design modifications over MS Access 2007. It is a relational database application meant for allowing users or programmers to manage, sort and filter large amounts of data. It is a powerful tool for creating and running database solutions and producing meaningful reports.

It is designed to help the organizations, such as a company, school, institutions to collect, organize, manage and access useful data in such a way that it can be easily retrieved and made available to users in a convenient way.

Also, it allows users to create custom databases in systematic structure that store the information. This application provides a graphic interface for creating custom forms, tables and data queries. It stores data in its own organized format based on the Jet Database Engine and it is the application of MS Office which is suitable for small as well as large scale business. The software can be downloaded from the MS Office Website. Following are some terminologies related to MS Access:

- **Tables:** Table allows the user to store a collection of data about a specific subject, such as student record.
- **Queries:** Queries allow the user to view, change and analyse data in different ways. This feature helps the user to fetch data from the created tables.
- **Forms:** Forms help users to modify or update the existing data by adding or deleting the new information.
- **Reports:** After storing the data in a database, user need to have a way to view it; this is where reports come into the picture.

5.5.1 Features of MS Access 2010

The features of MS Access 2010 are as follows:

- For sharing a database on the Web, five templates are available in MS Access 2010: Contacts, Assets, Projects, Events, and Charitable Contributions. You can also modify any template, before or after you publish.

NOTES

- An existing database can be converted to a Web. All desktop database features are not supported on the Web, so you may have to adjust some features of your application. The new Web Compatibility Checker can be run to identify and fix any compatibility issues.
- Access 2010 offers a new macro builder that features IntelliSense and a clean, straightforward interface. When you add an action, more options appear in the macro builder. In addition to traditional macros, you use the new macro builder to create data macros which is a new feature.
- Data macros help support aggregates in Web databases and also provide a way to implement ‘triggers’ in any MS Access 2010 database.
- The Expression Builder now features IntelliSense, so you can see options as you type. It also displays help for the currently selected expression value in the Expression Builder window.
- A field can be created that displays the results of a calculation and called calculation field. The calculation must refer to other fields in the same table. You can use the Expression Builder to create the calculation.
- MS Access 2010 includes a suite of professionally designed database templates for tracking contacts, tasks, events, students and assets, among other types of data. You can use them right away or enhance and refine them to track information exactly the way that you want.
- Each template is a complete tracking application that contains predefined tables, forms, reports, queries, macros and relationships. The templates are designed to be immediately useful out-of-the-box so that you can get up and running quickly.
- Functionality to an existing database can be easily added by using an application part. This is also a new feature in MS Access 2010, an application part is a template that comprises part of a database.
- In MS Access 2010, layouts have been enhanced to allow for more flexible placement of controls on forms and reports. You can split or merge cells horizontally or vertically, enabling you to easily rearrange fields, columns, or row. You must use Layout view when designing Web databases, but Design view still remains available for desktop database design work.
- The new user interface — introduced in MS Office Access 2007 and enhanced in MS Access 2010 — was designed to make it easy for you to find commands and features that before were often buried in complex menus and toolbars.
- The ribbon is a collection of tabs that contain groups of commands that are organized by feature and functionality. The ribbon replaces the layers of menus and toolbars found in earlier versions of MS Access.
- In MS Access 2010, Backstage view contains commands that you apply to an entire database, such as compact and repair, or open a new database. Commands are arranged on tabs on the left side of the screen, and each tab contains a group of related commands or links.

- MS Access 2010 provides an intuitive environment for creating database objects. Create tab is used to quickly create new forms, reports, tables, queries, and other database objects.
- MS Access 2010 adds new conditional formatting features, enabling you to achieve some of the same formatting styles as are available in MS Excel 2010. Conditional formatting is not available in Web databases.
- In MS Access 2010, you can export data to a PDF (Portable Document Format) or XPS (XML Paper Specification) file format for printing, posting and e-mail distribution, provided that you first install the Publish as PDF or XPS add-in. Exporting a form, report, or datasheet to a .pdf or .xps file lets you capture information in an easy-to-distribute form that retains all your formatting characteristics, but does not require others to install MS Access on their computers to print or review your output.

NOTES

5.5.2 Creating a Table

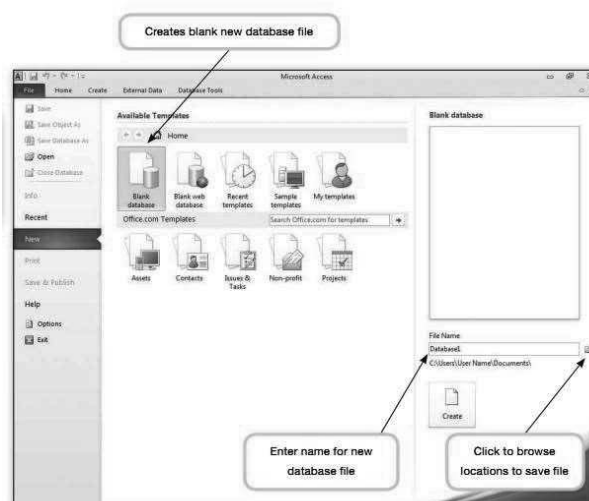
The term ‘database’ is used to describe a collection of related ‘data’ (information) stored on computers. An electronic database should allow you to store, sort and retrieve the data.

The Database Management System or DBMS is the software which acts as an interface between the underlying database and the user to define, update, retrieve and administer the database. Relational database is the collection of the logically related data items which are organized in the form of tables.

Certain steps are used to **create the database** in MS Access 2010, which are as follows:

1. Open **Microsoft Access 2010** from the **Start** menu.
2. Now, click on the **File** tab on the menu bar and then, go to **New option available on left hand side of the window**. Choose **Blank database** and assign name to the Database on the right hand side of the window.

The screenshot displaying the creation of a **Blank database** is as shown below:



3. After choosing the **Blank database** option and writing the name in the **File_Name** box, click on **Create** option to create the database.

Creating Table with Wizard

NOTES

A 'Table' is a grid containing columns and rows in it. It contains cells, which can contain any content inside them. At the moment you create the database, automatically Table 1 is created.

For creating the new table, follow the steps as given below:

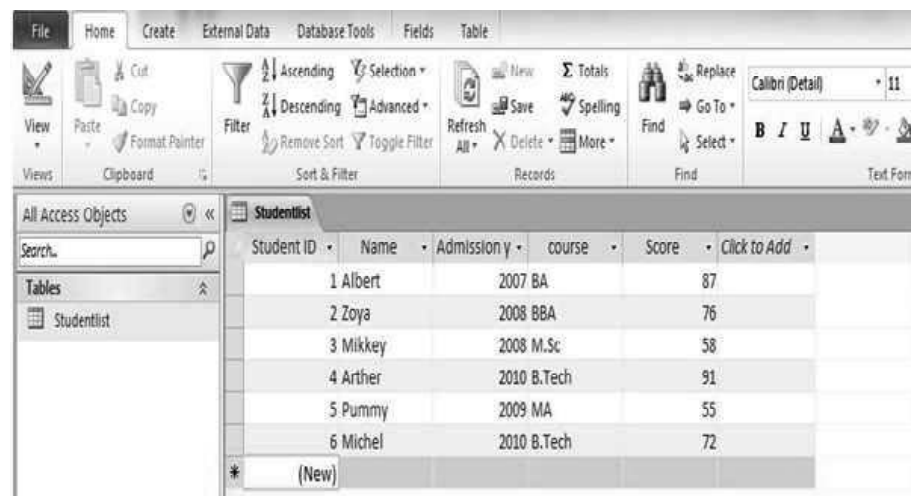
1. Click on the **Create** tab on the Ribbon view.
2. As you click on **Create** tab, **Tables** group containing the various options offers you alternative methods to create a new table.

The screenshot containing the **Tables** group inside the **Create** tab is as shown below:



3. Now, click on the Table option in the Tables group. As you click on it, a blank datasheet will be created for you, with an ID field and an Add New Field column.
4. As you enter data into the Add New Field column, another Add New Field appears and the previous Add New Field column is re-labeled as **Field1**.

The screenshot displaying the Student record in table in **Date sheet view** is as shown below:

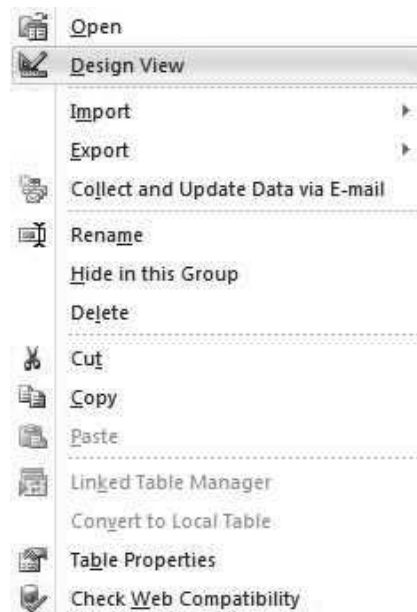


Creating Table in Design View

Steps for creating table in design view area as follows:

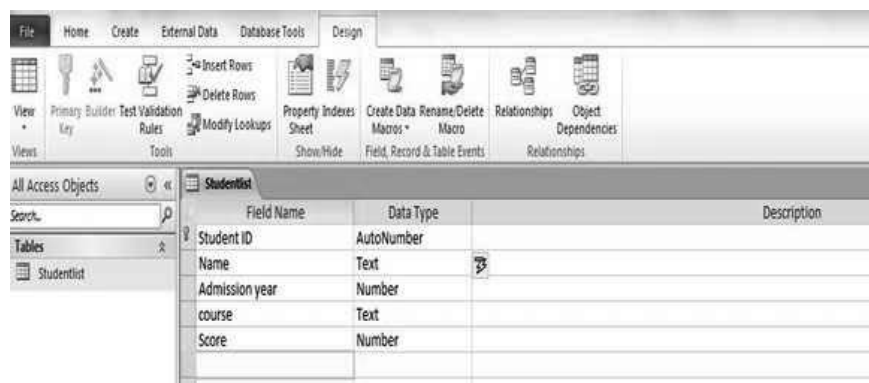
1. Click on the **Create** tab on the Ribbon and select the **Table Design** from **Tables** group.
2. For create a table design view of the existing student list Table; **right Click** on the Student- list and select the **Design view**.

The screenshot of displaying the **Design view** is as shown below:



3. As you click on the **Design view** option a table having the different fields or attributes as **Field Name**, **Data Type** and **Description** appears on the computer screen.

The screenshot displaying the **Table created in Design View** is as shown below:



Creating Table by Entering Data

The table can also be created by entering the data in the table. In this, the structure of the table is not designed already. Rather the data is entered and the structure of the table is designed based on the data that you enter in the column.

NOTES

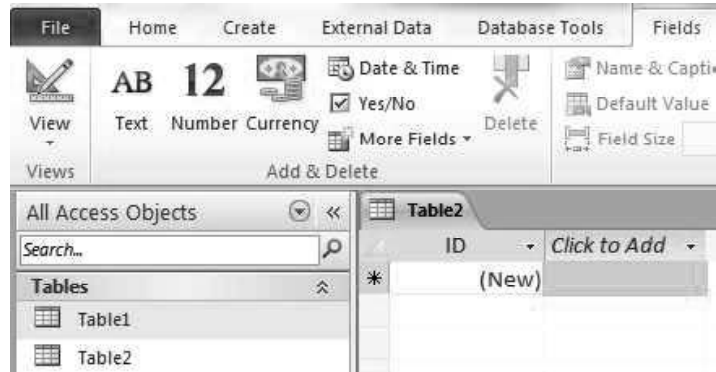
Once the cursor is moved to the next column, the column is saved. This removes the need for defining the table structure as a separate task.

Steps for creating the table by entering data are as follows:

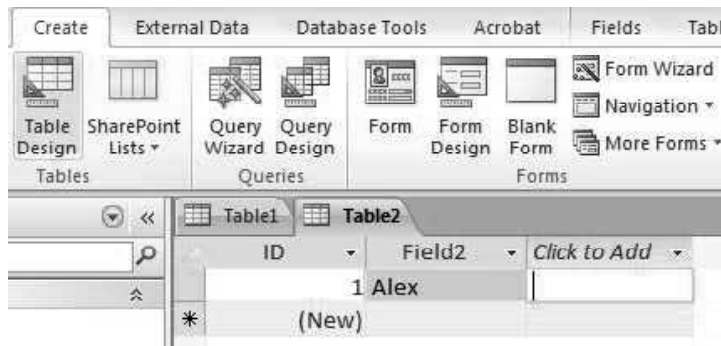
NOTES

1. In the ribbon, click on the **Create** tab and then click on **Table** button.

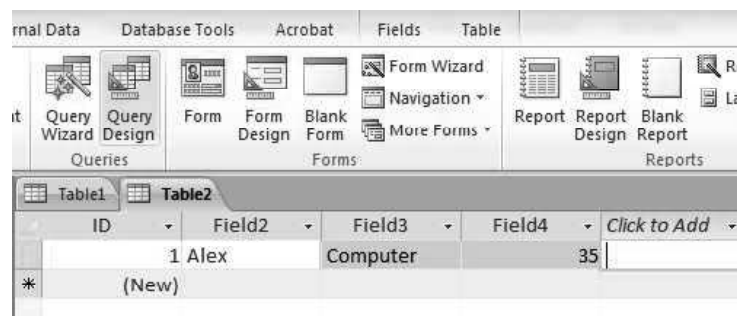
This creates a new table as shown in the screenshot below:



2. Now, start entering the data in the field. Enter the data in the field and when done, press the right arrow key on the keyboard. This sets the **ID** field to 1 and the type of data for the field would be set according to the data that was entered. The screenshot displaying the type of data would be set to **Text** for the **Field1** is shown below:



3. Keep entering the data till the time you want more columns. Once all the columns have been filled, click on the new row. Table structure would be saved according to the data entered in the first row. The screenshot displaying a table with three columns of type **Text**, **Text** and **Number** respectively, is as shown below:



5.5.3 Form Design in MS Access 2010

There are two simple ways of creating a form in MS Access, either by using **AutoForm** or **Form Wizard**. AutoForm is a very quick and an easy way to produce a form as it performs the required task just at the click of a button.

Steps to create **AutoForm in MS Access** are as follows:

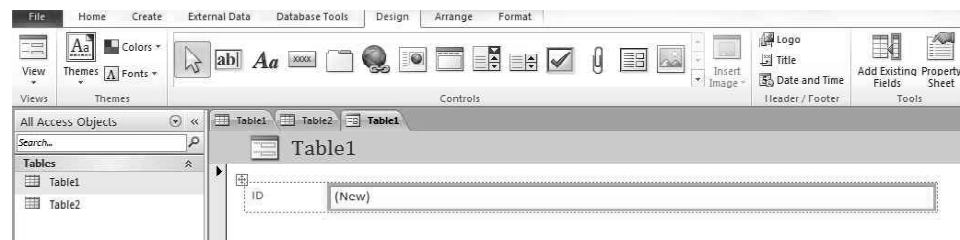
1. Click on the **Create** tab available on the **Ribbon View**.
2. As you click on the **Create** tab, **Forms** group containing the various options or controls for creating the form appears in front of you.

The screenshot displaying the **Forms** group is as shown below:



3. Now, click on the **Form** option of the **Forms** group.
4. As you click on the **Form** option, a new form

The screenshot displaying a new form created named **Table1** is as shown below:



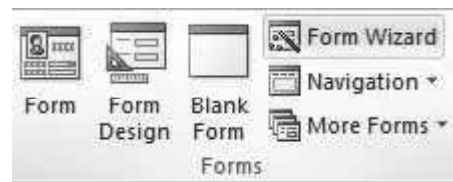
Creating Form by Wizard

The Form Wizard is equally easy to use and offers you various additional options.

Steps to create the form by Form Wizard are as follows:

1. Click on the **Create** tab available on the **Ribbon View**.
2. As you click on the **Create** tab, **Forms** group containing the various options or controls for creating the form appears in front of you.
3. From the **Forms** group select the **Form Wizard** option and click on it.

The screenshot displaying the **Form Wizard** option is as shown below:

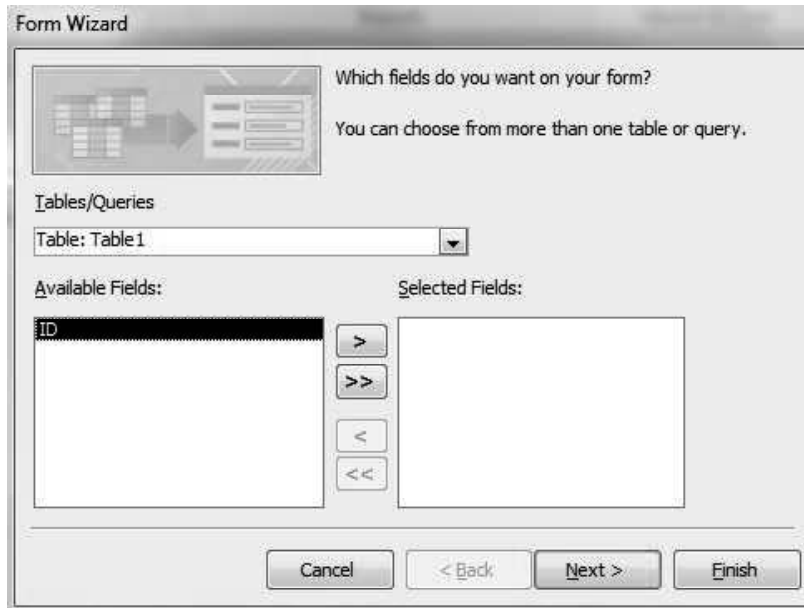


NOTES

- As you click on **Form Wizard** option, a **Form Wizard** dialog box appears on the screen.

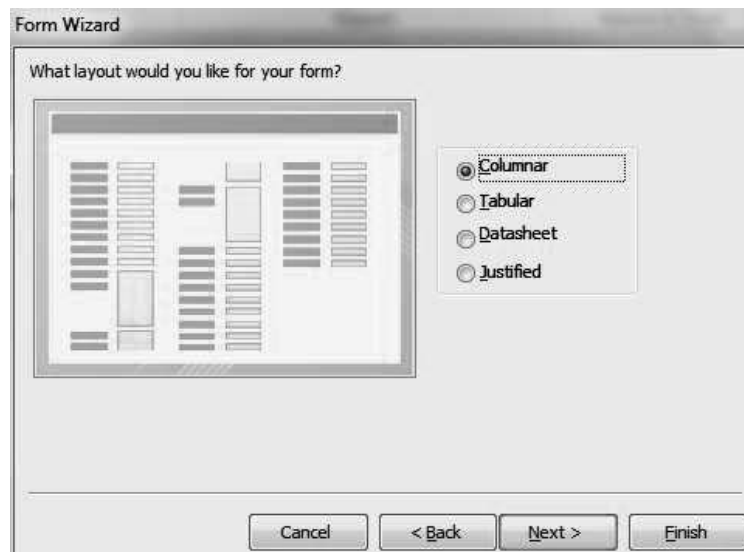
Screenshot displaying the **Form Wizard** dialog box is as shown below:

NOTES



- In the Form Wizard dialog box, all the fields available in the table are shown in the **Available Fields:** ListBox.
- Click on the double arrow >> in the Available Fields: ListBox to move the fields available in the Available Fields: ListBox to the Selected Fields: ListBox.
- Now, click on Next button present in the bottom of the Form Wizard dialog box.
- As you click on the Next button, a new dialog box containing the different layouts for the form appears.

The screenshot of the dialog box containing the different layouts for the form is as shown below:



9. In the **Form Wizard** dialog box, select the layout **Columnar** and then click on **Next** and then, **Finish** button.
10. As you click on the **Finish** button, a new form will be created using Wizard.

The screenshot of the **Form created using Wizard** and having ID Field is as shown below:



NOTES

5.5.4 Query Design and Running in MS Access 2010

Queries are the commands that are used to retrieve and update the data based on the conditions depending upon the business requirements. Queries also help to organize the view of the data. Two ways to create or design the query are as follows:

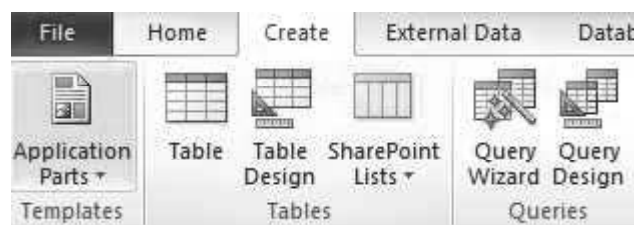
- Query with wizard
- Query with design view

Query in Design View

For creating or designing query in design view, the following steps are used:

1. First, click on **Create** tab of menu bar of Microsoft Access 2010.
2. As you click on the **Create** tab, **Queries** group containing the two options as **Query Wizard** and **Query Design** appears.

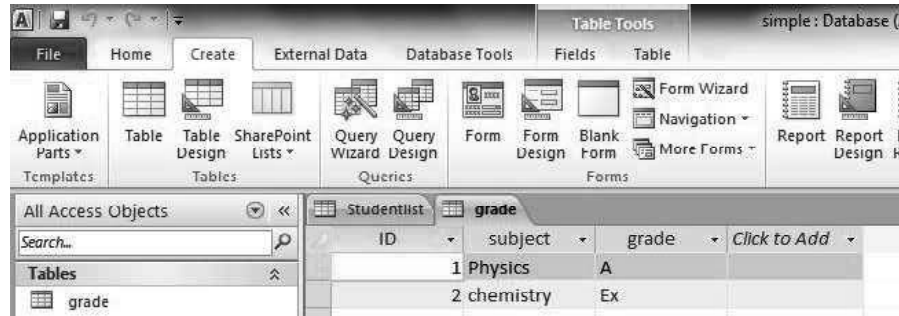
The screenshot displaying the **Queries** group is as shown below:



3. Now, suppose if there are two tables: Studentlist and grade. Now, click on the **Query Design** option of the **Queries** group.

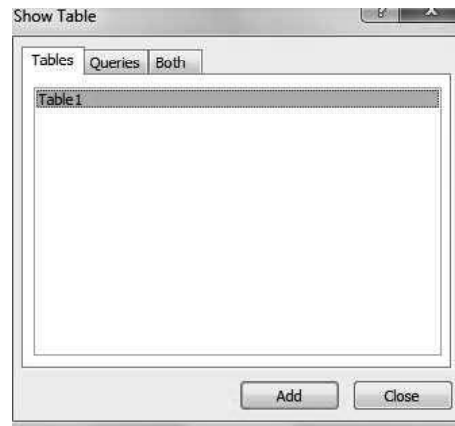
The screenshot displaying the two tables named Studentlist and grade is as shown below:

NOTES



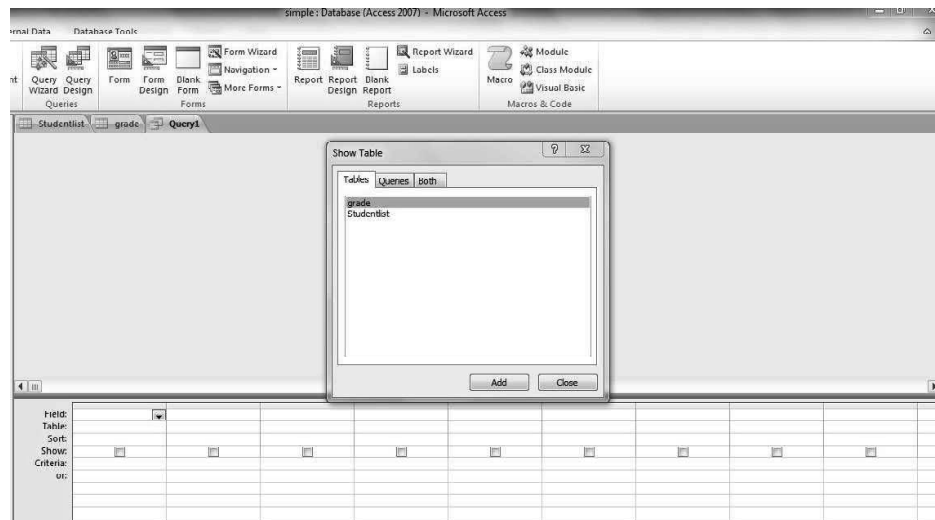
4. Now as you click on the **QUERY DESIGN**, a **Show Table** dialog box appears on the screen.

The screenshot displaying the **Show Table** is as shown below:



5. In the **Show table** dialog box, there are two tables as Studentlist and grade, click on the **Add** button available at the bottom of the dialog box.

The screenshot displaying the **Add** button is as shown below:



6. As you click on the **Add** button, both the tables will be added for designing the query.

7. Select the fields for solving the query.

The screenshot displaying the fields selected for the query is as shown below:

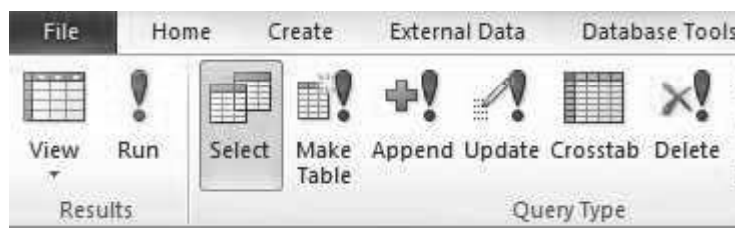
Windows, MS Office and Outlook 2010



NOTES

8. At last, click on **Run** in the **Design** Tab.

The screenshot displaying the **Run** command in the **Design** tab is as shown below:



9. As you click on **Run**, final query solution appears.

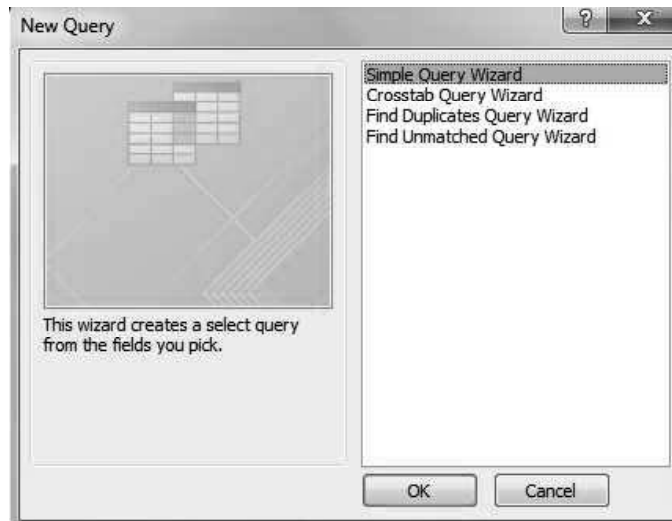
ID	grade.subject	grade	Field0	Student ID	Studentlist	Admission year	course	Score	Field1
1	Physics	A	Physics	1	Albert	2007	BA	87	Albert
2	chemistry	Ex	chemistry	1	Albert	2007	BA	87	Albert
3	computer	A	computer	1	Albert	2007	BA	87	Albert
4	biology	B	biology	1	Albert	2007	BA	87	Albert
5	english	B	english	1	Albert	2007	BA	87	Albert
1	Physics	A	Physics	2	Zoya	2008	BBA	76	Zoya
2	chemistry	Ex	chemistry	2	Zoya	2008	BBA	76	Zoya
3	computer	A	computer	2	Zoya	2008	BBA	76	Zoya
4	biology	B	biology	2	Zoya	2008	BBA	76	Zoya
5	english	B	english	2	Zoya	2008	BBA	76	Zoya
1	Physics	A	Physics	3	Mikkey	2008	M.Sc	58	Mikkey
2	chemistry	Ex	chemistry	3	Mikkey	2008	M.Sc	58	Mikkey
3	computer	A	computer	3	Mikkey	2008	M.Sc	58	Mikkey

Query Design with Wizard

Steps used to design query with Wizard are as follows:

1. Go to the **Create** tab and click on the **Query Wizard** option of the **Queries** group.
2. As you click on the **Query Wizard** option, a **New Query** dialog box appears on the screen.

The screenshot displaying the **New Query** dialog box is as shown below:



3. Select the **Simple Query Wizard** option in the **New Query** dialog box.
4. Click on the **OK** button available at the bottom of the **New Query** dialog box.
5. As you click on the **OK** button after selecting the **Simple Query Wizard** option, a new dialog box named **Simple Query Wizard** appears on the screen.

The screenshot displaying the **Simple Query Wizard** dialog box is as shown below:



NOTES

6. In the **Simple Query Wizard** dialog box, select the query for selected fields.

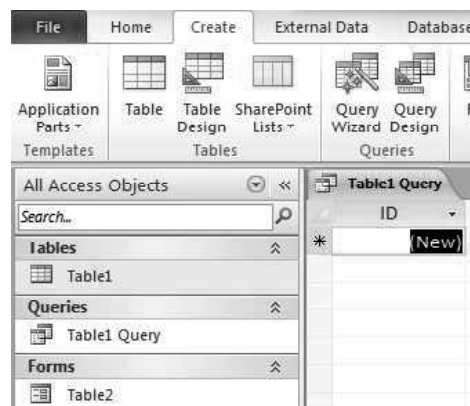
The screenshot displaying the **Selected Fields** as ID is as shown below:



NOTES

7. Now, click on **Next** button and then, **Finish** button to end the process.

The screenshot displaying the new window opened by designing the query with **Query wizard** is as shown below:



5.5.5 Report Design in MS Access 2010

Reports are designed to create an organized output of data from your database. With a Report, you can group and summarize information. You cannot edit the data in a Report but if you make the modifications in the Table, Query, or Form you will see the results when you open the Report again.

There are two ways of designing the Report:

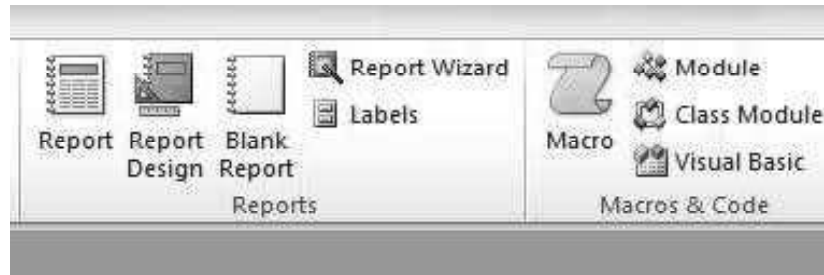
- Design with wizard and design with design view.
- Begin by creating a report for the student list table using design view.

Report Design with Wizard

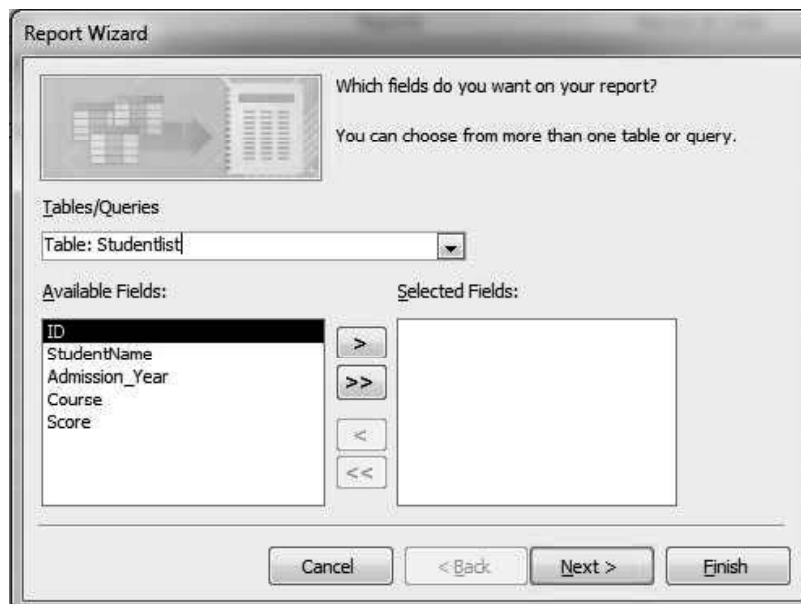
Wizard helps in designing the reports by providing options step by step for designing the reports. Report design with wizard can be done using the following steps:

NOTES

1. Go to **Create** tab and click on **Report Wizard**. The screenshot of **Report Wizard** button is as shown below:

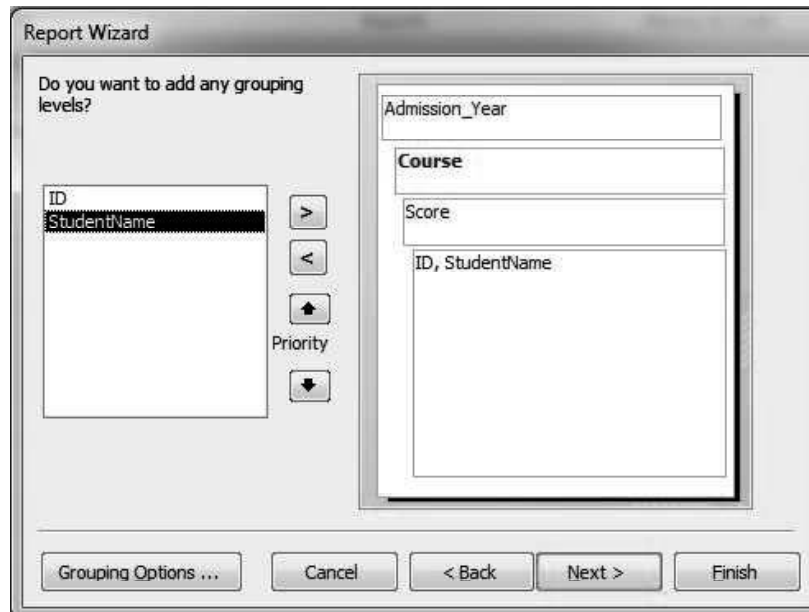


2. The window that appears leads through the steps for creation of report. The first window asks for choice of table and related columns. This lets you choose the table from which the report should be prepared for and what columns should be displayed. Choose the relevant table and columns from here. After this, click on **Next**. The screenshot displaying the first window of **report wizard** is as shown below:

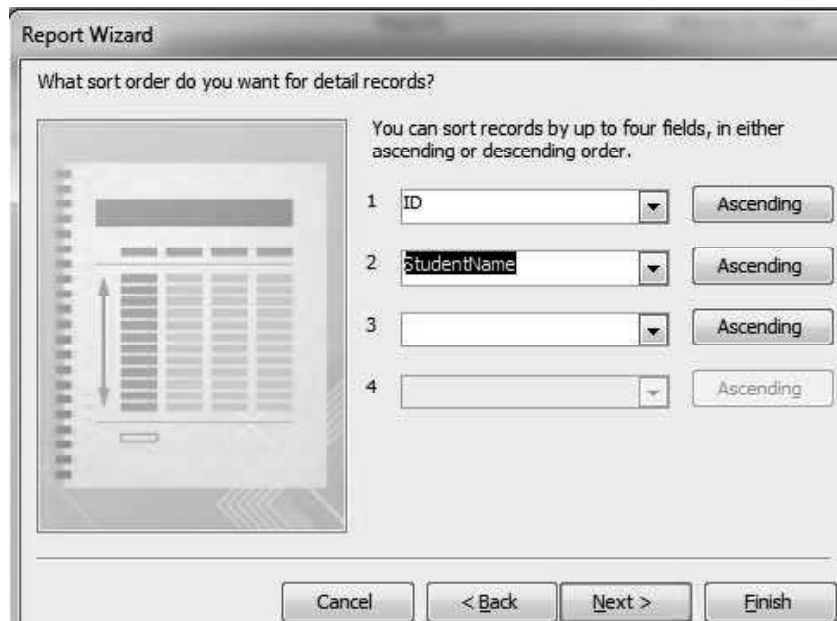


3. Now, the next window of wizard asks to set any grouping levels if required. The report would be shown according to the grouping that has been set. Select the appropriate settings and click **Next**. The screenshot displaying the grouping options is as shown below:

NOTES



4. Next window asks to set any sorting order if required. The data can be sorted according to any column. Multiple sorting filters can be applied. If there is same data in a column that is used for sorting, the sorting would be done according to the next column. The screenshot showing the sorting option window is as shown below:



5. Now choose the layout options from the window. This window allows you to choose whether the report layout should be portrait or landscape

and many other options. The screenshot of the layout window is as shown below:

NOTES



6. Now choose a name from the window that is shown. Select the appropriate choice and click the **Finish** button. The screenshot of the last window is as shown below:



The screenshot of the generated report is as shown below:

Windows, MS Office and Outlook 2010

Studentlist1

Admission_Year
2007

Course
BA

Score
90

ID	StudentName
6	

Admission_Year	2007
Course	BA

Score	90
ID	StudentName
1	Albert

NOTES

Report Design with Design View

The reports can be created using the design view of report creation. In this format of report can be specified manually. The steps to design the report with design view are as below:

1. Click on the table in the Objects list. There is no need to open it.
2. Move to the Create tab and click on Report. MS Access shows you the results in **Layout View**. The dashed lines show the page breaks, but it is much easier to see the layout in Print Preview as shown in the screenshot below:

The screenshot shows the MS Access Report Design View. The report title is 'Studentlist'. The table contains the following data:

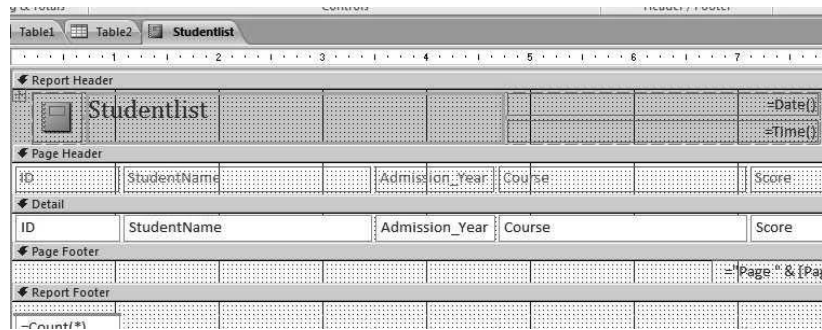
ID	StudentName	Admission_Year	Course
1	Albert	2007	BA
2	Zoya	2008	B. Com
3	Vincent	2008	MCA
4	Richard	2009	BCA
5	Sanjay	2010	LLB
6			

NOTES

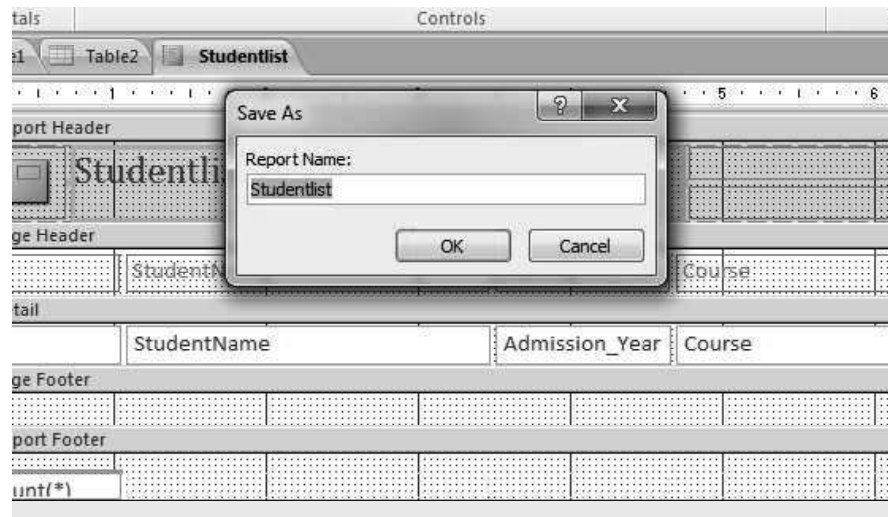
3. Right click on the tab that shows the table name. The screenshot displaying the table tab with the **Save** option is as shown below:



4. Click on **Design View** button. This opens the window where you can format the report. Change the names of the columns according to the need. The screenshot displaying the **Design View** is as shown below:



5. When the design is completed, save the report. For this right click the tab and click on **Save**. Give an appropriate name for the report and click **OK**. The screenshot displaying the **Save As** dialog box is as shown below:



5.5.6 Using MS Access 2010 for Business Decisions

Microsoft office is business cost effective tool that integrates the most popular office software. MS Access Database can be integrated effectively and efficiently with programs, such as MS Excel - for charts, pivot tables and graphs, MS Word for mail merges and labels. MS Access provides better results for business decisions. Organizations can start tracking the information immediately with a rich library of pre-built solutions. Forms and reports are already built for convenience and can be easily customized to meet business needs of organizations. Contacts, issue tracking, project tracking and asset tracking are only few of the solutions included in MS Office Access 2010. Multiple reports can be easily created with different views of same information. Creating a report in MS Office Access 2010 is based on 'What You See Is What You Get' (WYSIWYG) experience. A report can be easily modified with real time visual feedback and save various views for different audiences. The new grouping pane and filtering and sorting capabilities help the organization to display the information so you can make more informed business decisions.

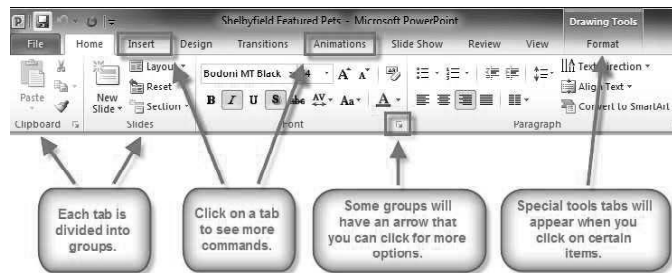
With automatic data type detection, table creation in MS Office Access 2010 is as easy as working with a Microsoft Office Excel table. MS Office Access 2010 will automatically recognize the typed information whether it is a date, currency or another common data type. MS Office Access 2010 also enables new field types, such as attachments and multivalued fields. Any document, image can be easily attached or spreadsheet to any record in application for various business decisions. More than one value can be selected with the help of multivalued field. With MS Office Access 2010, organizations can create forms using Microsoft Office InfoPath or HTML to gather information for the database. This form can be send to teammates via e-mail and use their responses to populate and update MS Access table to eliminate the need to retype any information. Organizations can share information in MS Access database with the rest of team using Windows SharePoint Services. With the advanced features of both applications, teammates can access and edit the data and view real time reports directly through a Web interface. MS Office Access 2010 can be used as a rich client interface to analyse and create reports from Windows SharePoint Services lists. You can even take the list offline and then synchronize any changes when you reconnect to the network, making it easy for you to work with data at any time. The data become more transparent by moving it to Windows SharePoint Services. This way, the backup of the data can be taken on the server on the routine basis, recover deleted data from the recycle bin, track revision history and set access permissions, so the data of an organization can be managed easily for better decision-making. Using MS Office Access 2010, tables can be linked to the database from other sources, such as other MS Access databases, MS Excel spreadsheets, Windows SharePoint Services sites, ODBC data sources, Microsoft SQL Server databases and other data sources. These linked tables can be used to easily create reports so that an organization can base decisions on a more comprehensive set of information.

NOTES

5.6 MICROSOFT POWERPOINT 2010

NOTES

Ribbon view is a substitute that provides the facilities of accessing the commands for applications. It is organized using a horizontal bar. It provides an easy way to access the commands, as all the commands are organized using tabs and groups. With the help of Ribbon, performing various actions to the Office application is easy.



The ribbon contains three components:

- **Tabs:** They appear across the top of the Ribbon and contain groups of related commands. Home, Insert, Page Layout are example of ribbon tabs.
- **Groups:** They organize related commands; each group name appears below the group on the Ribbon. For example group of commands related to fonts or or group of commands related to alignment etc.
- **Commands:** Commands appear within each group as shown in the screenshot.

VIEWS IN SLIDES

PowerPoint views can be accessed from two locations.

- Views can be access quickly from the bottom bar just to the left of the zoom settings



Quick Access to Presentation Views

- Views can also be accessed from the Presentation Views section in the View ribbon

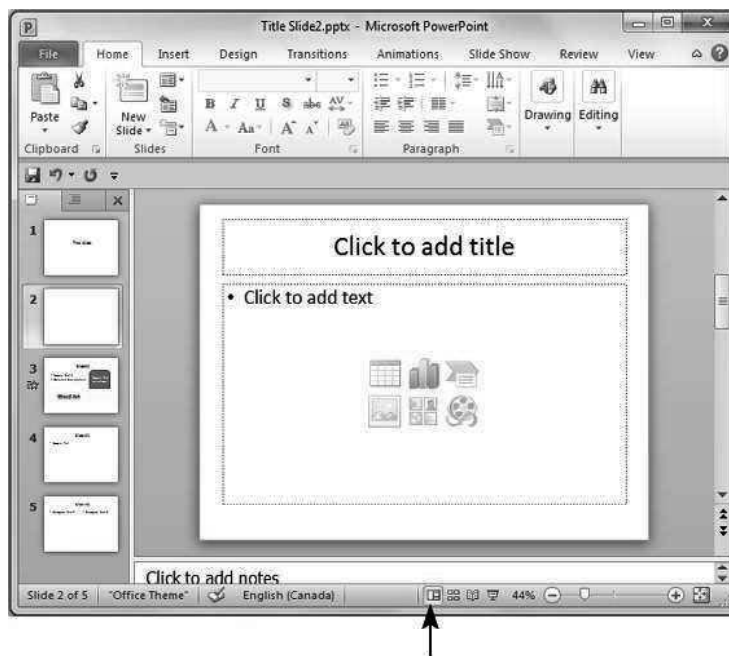


NOTES

A short description of the different views and their features.

Normal View:

This is the default view in PowerPoint and this is primarily used to create and edit slides. User can create/ delete/ edit/ rearrange slides, add/ remove/ modify content and manipulate sections from this view.

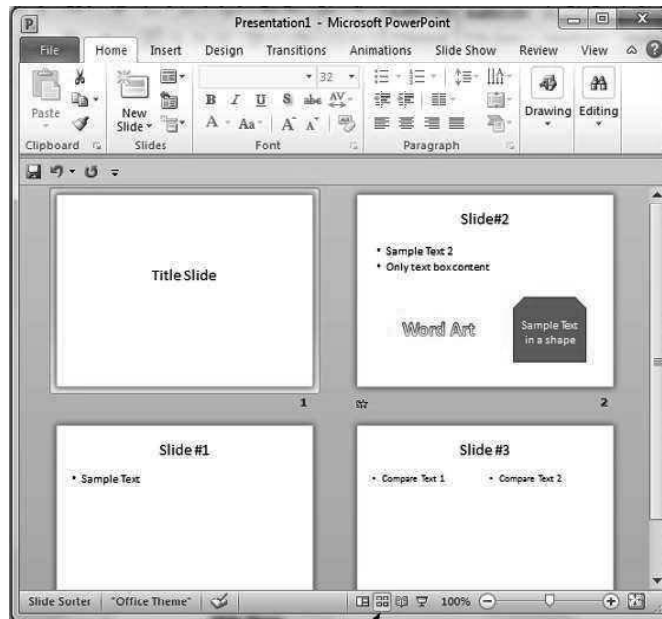


Normal View icon

Slide Sorter View:

This view is primarily used to sort slides and rearrange them. This view is also ideal to add or remove sections as it presents the slides in a more compact manner making it easier to rearrange them.

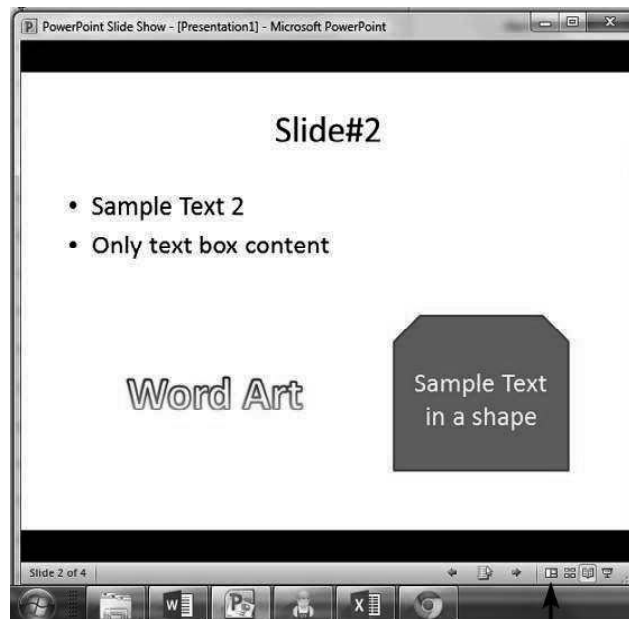
NOTES



Slide Sorter View icon

Reading View:

This view is new to PowerPoint 2010 and it was created mainly to review the slide show without losing access to rest of the Windows applications. Typically when User run the slide show, the presentation takes up the entire screen so other applications cannot be accessed from the taskbar. In the reading view the taskbar is still available while viewing the slide show which is convenient. User cannot make any modifications from this view.

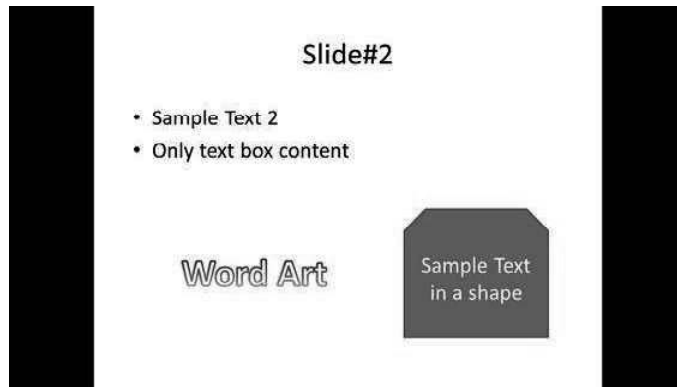


Windows Taskbar visible

PowerPoint bottom bar visible

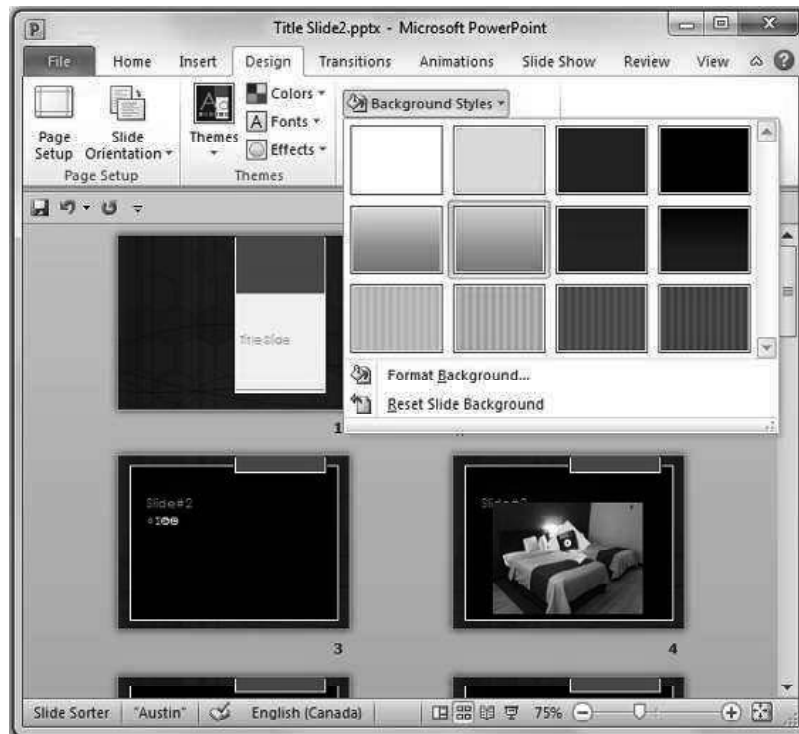
Slide Show:

This is the traditional slide show view available in all the earlier versions of PowerPoint. This view is used to run the slide show during presentation.



NOTES

Step (1): In the **Design** ribbon, under the **Background** group click on the **Background Styles** command.

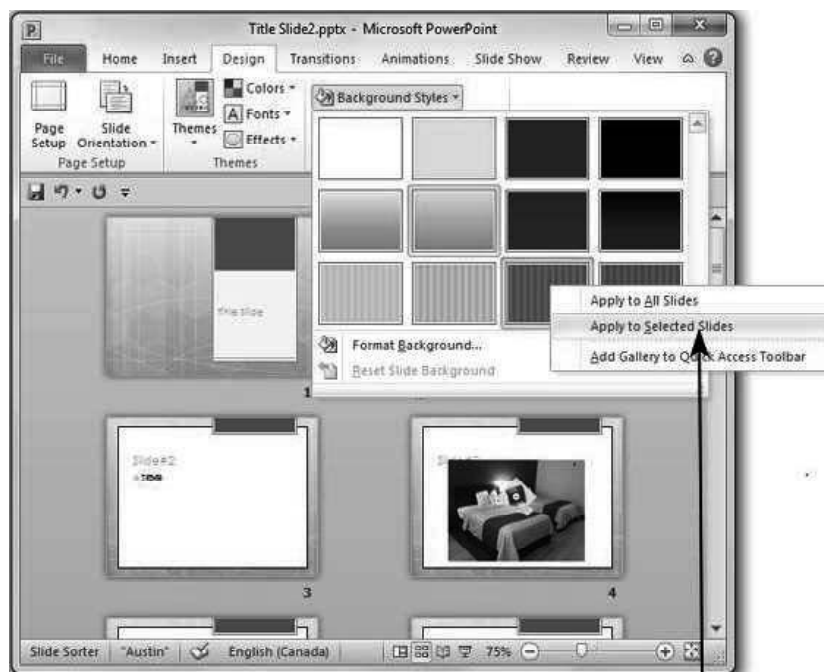


NOTES

Step (2): Select one of the background styles that suits your requirements



Step (3): To edit the background for a specific slide, right click on the desired background slide and select “Apply to Selected Slides”



Apply to Selected Slides option

Step (4): Selected slide(s) now have the new background.

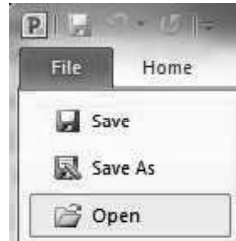
5.6.1 Creating Presentation and Slides

Opening a presentation in PowerPoint

Steps for opening a presentation are as follows:

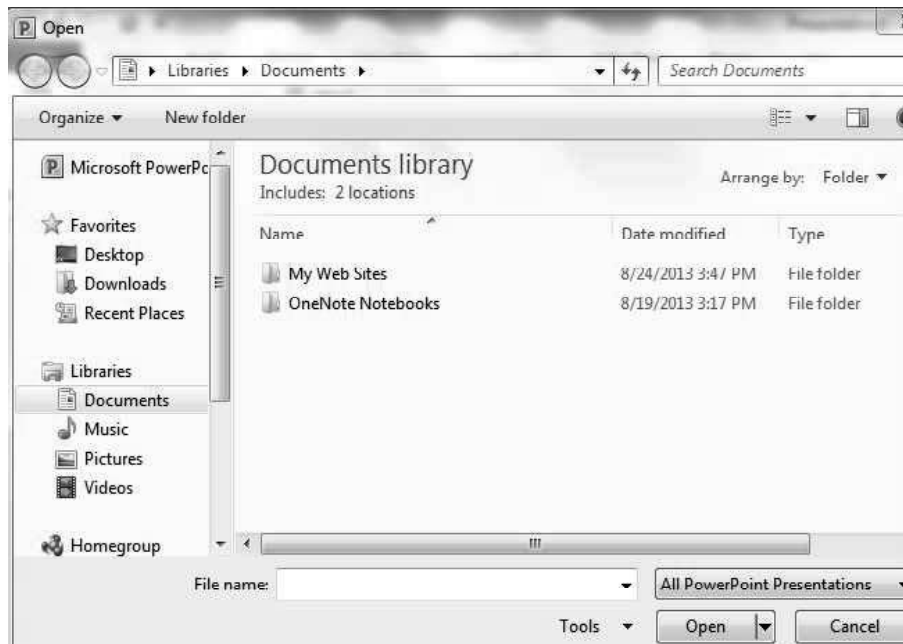
1. Go to the **File** tab which is present on the left hand side of the ribbon.
2. Go to the **Open** option of the **File** tab.

The screenshot of an **Open** option is as shown:



3. As you click on the **Open** option, an **Open** dialog box appears. Choose the appropriate drive and folder from where you want to open the file.

The screenshot of an **Open** dialog box is as shown:



4. Choose the file which you want to open by writing its name in the **File name** text box.
5. Click on the **Open** option to open the file.

Saving a document

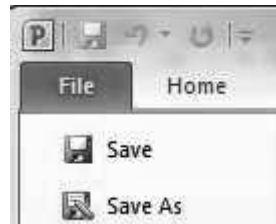
In PowerPoint 2010, by default, files are saved in .pptx format. In the previous version, files are saved in .ppt format. There are different ways of saving the files in your computer depending upon the requirements.

NOTES

Steps for saving the presentation in PowerPoint are as follows:

1. Go to the **File** tab which is present on the left hand side of the ribbon.
2. Go to the **Save As** option of the **File** tab on the left most side of the screen.

The screenshot of **Save As** option in **File** tab is as shown:



3. As you click on the **Save As** option, a **Save As** dialog box appears. From the dialog box, you can choose the required drive or folder where you want to save the file.

The screenshot of **Save As** dialog box is as shown:



4. In the **Save As** dialog box, go to the **File Name** text box to write the name of the file.
5. Choose the format in which you want to save the File. By default file format is .pptx.
6. Click on the **Save** option available on the bottom right side of the dialog box to save the file.

Creating a new slide

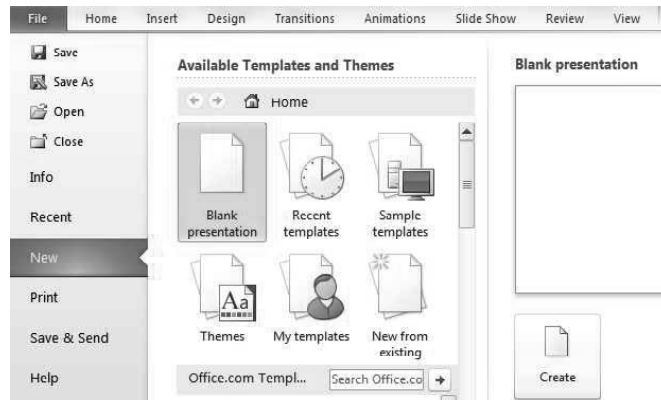
Steps to create a new slide are as follows:

1. Click on the **File** menu.
2. Select the **New** option of the **File** tab and click on it.

NOTES

3. A new window of **Available Templates and Themes** appears.
4. Click on the **Blank presentation** and then click on the **Create** button on the right hand side of the window.

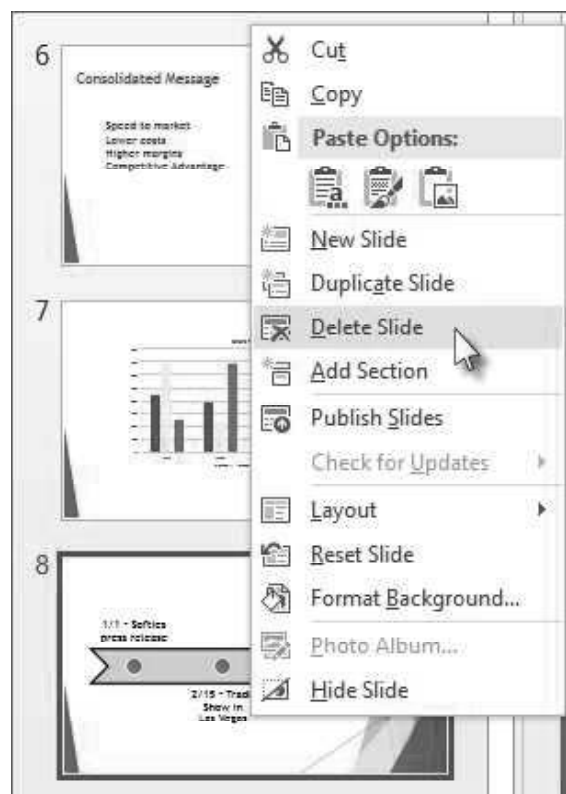
The screenshot displaying the **Blank presentation** is as shown:



NOTES

Changing layout of the slide and deleting a slide

You can change the layout of the slide and delete the slide by simply right click on thumbnail of the **slide**. The illustration is shown below.



Changing the order of the slides

You can change the order by simply right click the thumbnail of the slide that you want to move, and then drag it to the new location.

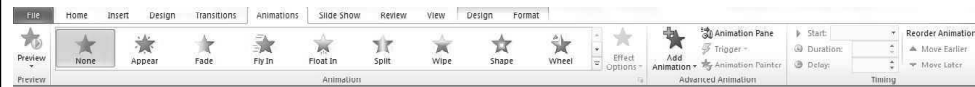
5.6.2 Applying Animation and Transition Effects

NOTES

Applying Animation

Animation is one of the most common features of MS PowerPoint. It is used to make an attractive presentation and for creating the interest in audience. MS PowerPoint 2010 provides different varieties of animations that can be applied to text, picture or other graphics in multiple ways.

The screenshot of **Animations** tab is as shown:

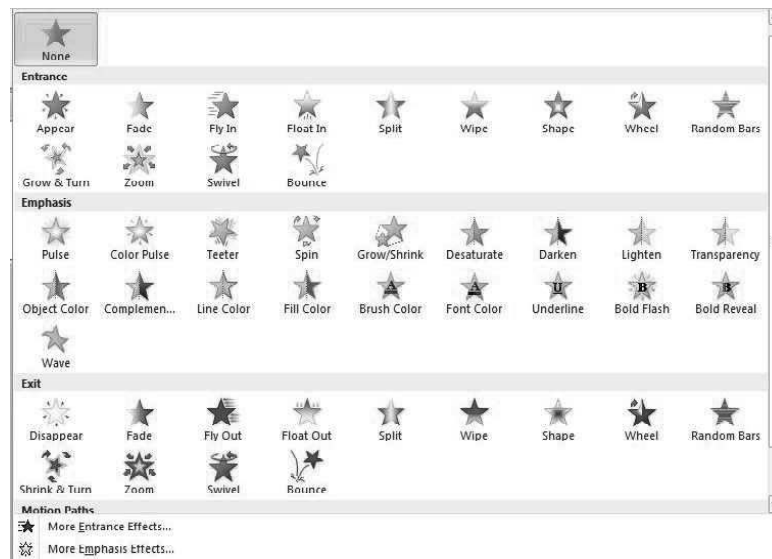


The first part of the **Animations** group provides the various options for applying animations to a character, text, picture and graphics within each slide.

Steps to **choose an animation** for applying it to objects of a slide are as follows:

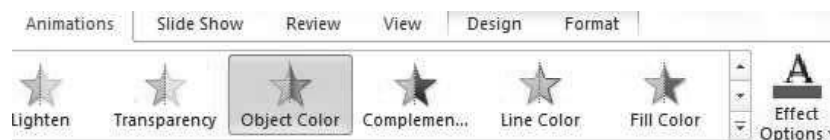
1. To get all of the options from **Animations** group, click on the Down Row, Up Row and Many arrows to the right of the **Animations** options pane.
2. Select the Animation option to apply on the text or other graphics.

The screenshot of **Animation** options is as shown:



3. Click on **Effect Options** icon to apply an animation effect to the selected object.

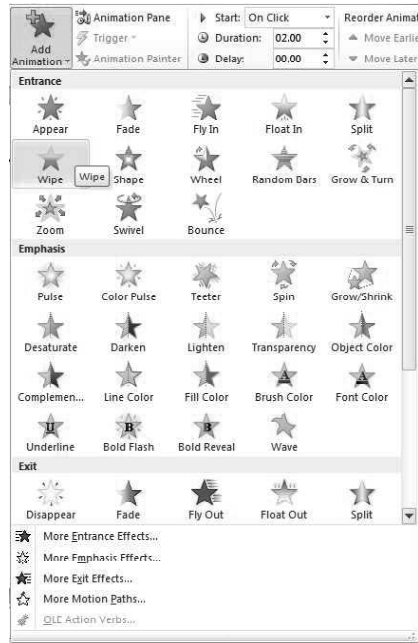
The screenshot of **Effect Options** icon of an **Animations** tab is as shown:



4. Click on the **Add Animation** icon present next to **Effect Options**. As you click on **Add Animation**, a pictorial view of all of the animation effects to animate text or graphics appears.

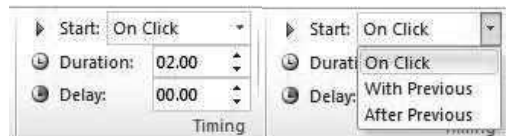
- Choose the animation effect for adding it to the required object. The selected animation effect will be highlighted in golden color.

The screenshot of **Add Animation** and selecting a specific animation effect is as shown:



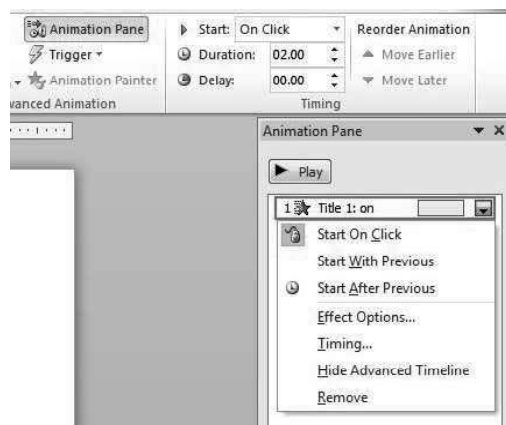
- Timing** group of **Animations** tab allows user to adjust the sequence and timing of selected animations. User can also adjust settings if an animation should be performed when the mouse is clicked or automatically.

The screenshot of **Timing** group is as shown:



- Click on the **Animation Pane**, it displays all the animations applied to each slide.

The screenshot of **Animation Pane** is as shown:



NOTES

8. After selecting all the animations as per the requirement, you can see the preview by clicking on the **Preview** icon available below the **File** menu.

The screenshot of **Preview** option is as shown:



NOTES

Transition Effects

Slide Transition is also an animation effect which is to define effects while proceeding from one slide to next slide during the slide show. It also includes adding sound effects. You can apply the different kinds of transition effects in slides of a presentation.

Steps to apply the transition effect are as follows:

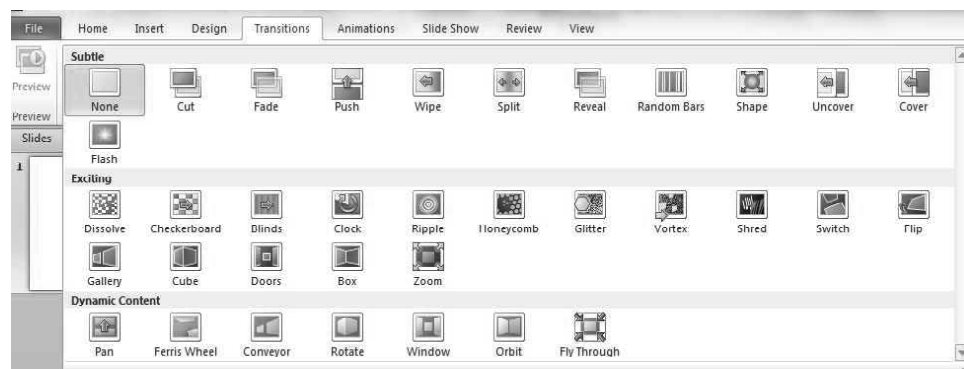
1. Click on the **Transitions** tab which is available on the Ribbon view.

The screenshot of **Transitions** tab is as shown:



2. Click on the 'Up' and 'Down' arrow of **Transition to This Slide** group to choose a special effect that gets applied during the transition between the previous slide and next slide.

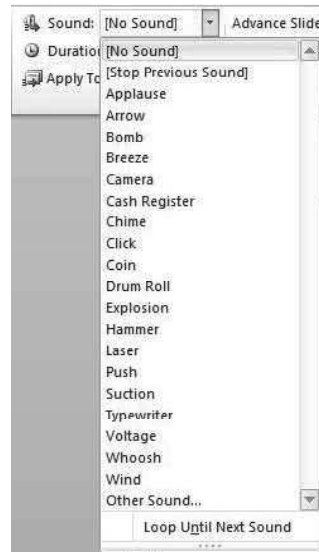
The screenshot displaying all the **Transition Effects** is as shown:



3. As you click on any of the transition effects, a preview of the theme plays in front of you. To set the transition to slide, left click on the choice which you want.
4. Click on the **Apply To All** option of the **Transition** tab to apply the same transition effect on all slides.
5. Click on the list box of the **Sound** option of **Transition** tab.

6. Select and click on the sound which you want to play during the transition between the current and the previous slide.

The screenshot displaying different types of **Sound** options is as shown:



NOTES

7. After applying all the transitions you can also see the preview of a slide by clicking on the **Preview** option.

5.6.3 Adding Sound and Video in Slides

Microsoft Office PowerPoint 2010 allows its users to easily use audio-visual content in the presentation. It supports a large variety of video files such as, Windows Media File, MP4 Video, MPEG-2 TS Video and Window Media Video File, along with audio files such as ADTS Audio, AIFF audio file (aiff), AU audio file, MP3 and MP4 audio files.

Audio and video content can be added to Microsoft Office PowerPoint 2010, using the **Media Clips** Option under the **Insert Tab**. Thus, it is possible to insert sound, picture or media files in PowerPoint presentations or slides to make them more attractive and presentable.

Steps for adding audio or video clips in PowerPoint 2010 are as follows:

1. First go to the **Insert** menu of the menu bar and click on it.
2. When you click on it, you will discover the **Media** group containing **Video** and **Audio** options.
3. Now, click on the **Audio** or **Video** icon, as per your requirement.

Screenshot displaying the **Audio** and **Video** icon, available under the **Media** group of the **Insert** tab is as shown:



• Adding audio or sound Files in Microsoft Office PowerPoint 2010:

Steps for adding Audio clip to the PowerPoint file is as shown:

1. Click on the drop down arrow of the **Audio** icon under the **Media** group.
2. As you click on the drop down arrow, a list of options available under this icon appears.

Screenshot displaying the **Audio** icon, visible after clicking the drop down arrow, is as shown:



3. You can choose any of the three options to add audio from various sources, like from a previously saved file on your computer or from Clip Organizer.
4. Suppose you choose **Audio from File** option, then on clicking **Audio from File**, **Insert Audio** dialog box appears, showcasing a set of folders having the music or song files present on your system.
5. Now, navigate to the appropriate folder to use or place the desired audio file into the PowerPoint File
6. Click on **Libraries** and then double click on the **Music** library file or option. When you click on **Music**, the **Sample Music** folder appears. Next, double click on it.
7. After clicking on **Sample Music**, a window containing all the music files available in your system appears. Here, you can choose any media or sound file, you want to insert in your presentation.

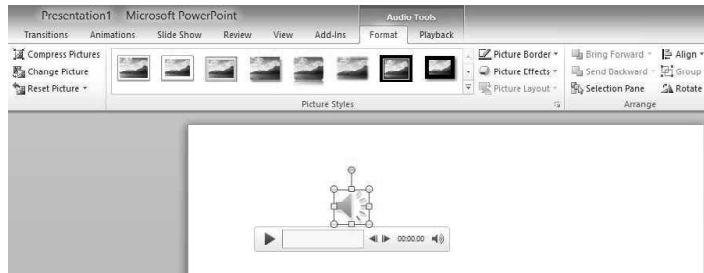
Screenshot of the **Insert Audio** dialog box and files showcased under the **Music Folder**, after clicking on **Audio from File**, is as shown:



NOTES

8. After choosing the desired file, click on the **Insert** button which will add the audio file to the slide.

Screenshot of the slide after the **Audio File** has been added is as shown:



NOTES

• Adding Video Files in Microsoft Office PowerPoint 2010:

1. Click on the drop down arrow of the **Video** icon of the **Media** group.
2. When you click on the drop down arrow, a list of available options appears.

Screenshot displaying the **Audio** icon after clicking the drop down arrow is as shown:



4. Now, you can choose one of the three options to add Video from various sources. Click on **Video from File** or to add from a website, click on **Video from Web Site**. To add from Clip Organizer, click on **Video from Clip Organizer**. Here, for example, let's click on **Video from File**. On clicking **Video from File**, **Insert Video** dialog box appears displaying the set of folders having all the music or song files present on your system.

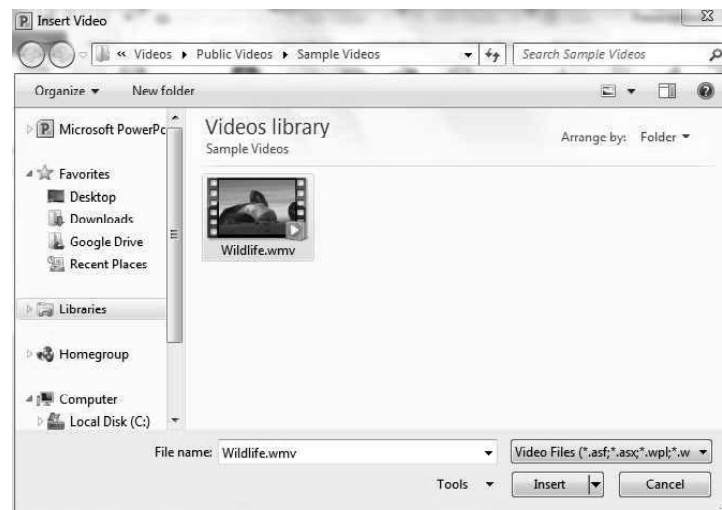
Screenshot displaying the **Insert Video** dialog box containing the **Pictures** and **Videos** library is as shown:



NOTES

5. Now, navigate to the appropriate folder to place the desired video file in the PowerPoint file.
6. Click on **Libraries** and then, double click on the **Video** or **Picture** library file.
7. When you click on the **Video** library, the **Sample Video** folder appears. Double click on it.
8. After clicking the **Sample Video**, a pop-up window containing all the video files available in your system appears, from where you can choose any media or video file to insert in your presentation.

Screenshot of the video file present in the **Sample Video** folder, visible after clicking on **Video library**, is as shown:



9. After choosing the desired file, click on **Insert** button. This will add the video file to the slide.

5.6.4 Working with Tables and Pictures

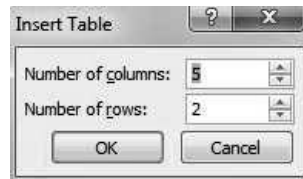
A Table is a representation in the form of rows and columns for managing and presenting the data. For using tables in the slides, there are various processes like create table, insert table, format table, etc.

- **Create and Format a table**

Steps to Create and Format a table are as follows:

1. Select the slide where you want to add a table.
2. Select **Table** option from the **Insert** tab in the **Tables** group.
3. Click on the **Insert Table** command from the drop-down menu. An **Insert Table** dialog box appears where you can add the required number of rows and columns.

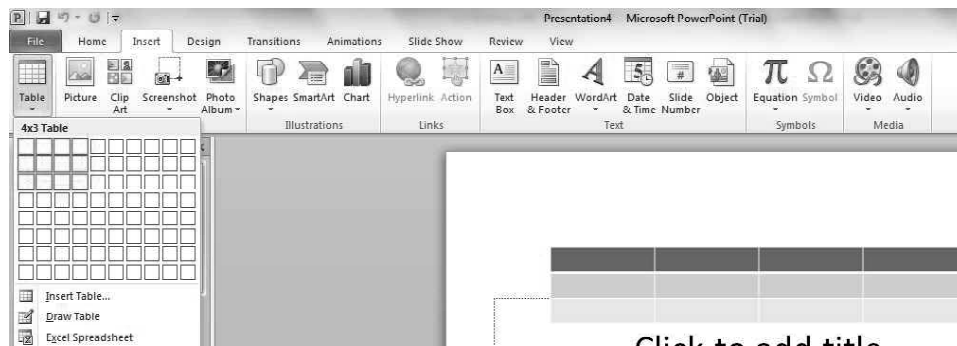
The screenshot of an **Insert Table** dialog box is as shown:



NOTES

4. You can also insert the table by selecting the required number of rows and columns on the Table grid of the drop-down menu.

The screenshot of inserting a table using table grid is as shown:



5. For adding text to the table cells, click on a cell and enter the text.
6. For adding a row at the end of a table, click on the last cell of last row and then press **TAB**.

• For Adding a Column or a Row:

Steps to add a column or a row are as follows:

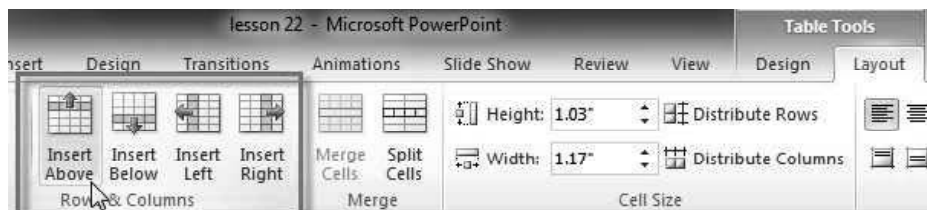
1. Click on the cell (insertion point) adjacent to which you wish to add row or column.

The screenshot displaying the insertion point is as shown:

Classics	\$18,580.00	\$69,225.00
The insertion point	78,970.00	\$82,262.00

2. Use commands available in **Rows & Columns** group under **Layout** tab of **Table Tools** to insert rows or columns (above, below, left and right) of the selected cell.

The screenshot of **Layout** tab having **Rows & Columns** group is as shown:



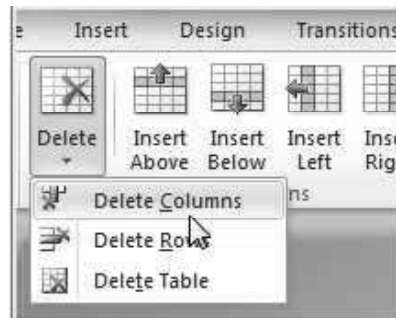
NOTES

• Deleting a Row or Column

Steps to Delete a Row or Column are as follows:

1. Click on any cell of a desired row or column you want to delete.
2. Click on the **Layout** tab of the **Table Tools** and then, click on the **Delete** icon available in the **Rows & Columns** group.
3. As you click on the **Delete** icon, a drop-down menu containing three options such as **Delete Columns**, **Delete Rows** and **Delete Table** appears. You can select any one of them according to the requirement.

The screenshot of **Delete** icon available in **Rows & Columns** group is as shown:



Creating Photo Album

While creating the Photo Album in PowerPoint, you can use templates which are built in or create on your own. This feature allows you to add the pictures or photos of your choice in a presentation. Also, by exploring this feature of the MS PowerPoint, you can choose, rearrange and add text to the pictures. You can also add pictures from your drives to the Photo Album. Several effects like layouts, slide transitions and themes can be applied to the Photo Album. You can also add captions to the photos and frames around them.

For Creating Photo Album

Steps to create a Photo Album are as follows:

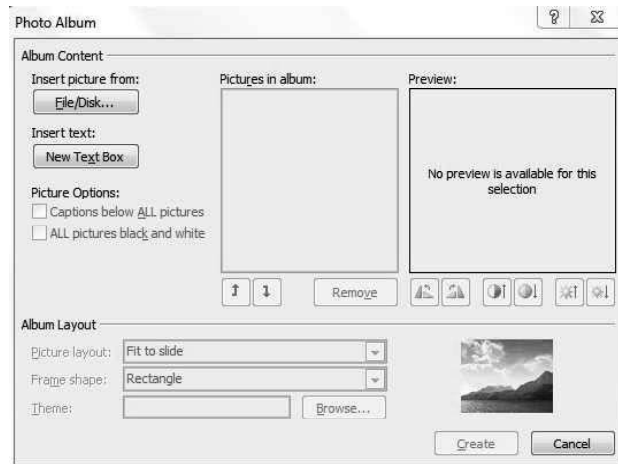
1. Select and click on the **Insert** tab available in the ribbon.
2. Click on the **Photo Album** option in the **Images** group and then select **New Photo Album** from the drop-down menu.

The screenshot of the **Photo Album** is as shown:



3. As you click on the **New Photo Album**, a **Photo Album** dialog box appears.

The screenshot of **Photo Album** dialog box is as shown:



NOTES

4. Click on the **File/Disk...** button of the **Photo Album** dialog box.
5. As you click on the **File/Disk...** button, an **Insert New Pictures** dialog box appears on the screen.

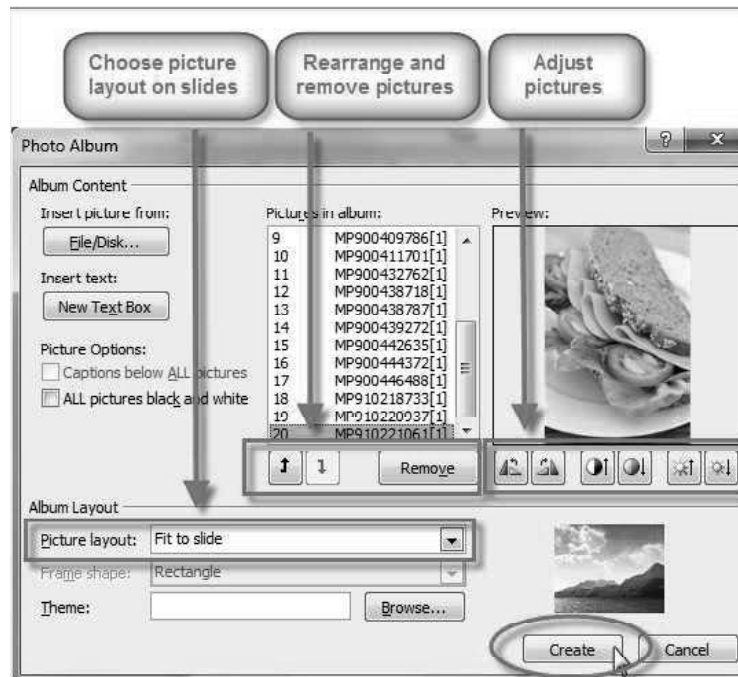
The screenshot of **Insert New Pictures** dialog box is as shown:



6. Now, you can choose the desired images or photos from the corresponding drive by writing the name of the file in the **File name** textbox and then click on the **Insert** button.
7. A preview of the selected image appears inside the **Photo Album** dialog box. There are several options available for changing the layout, rotating the image, adjusting it as needed.

- Now, click on the **Create** button to insert images into the **Photo Album**. The screenshot of the **Picture layout** and **Theme** in the **Photo Album** dialog box is as shown:

NOTES



- A presentation with one image in each slide will be created. A title page is also automatically added in the presentation.

5.6.5 Working with SmartArt

A SmartArt can be easily inserted in MS PowerPoint 2010. SmartArt icon can be found in **Illustrations** group. The various graphics available in this option are All (basic block list, alternating hexagon, etc.), matrix, relationship, hierarchy, etc.

Steps to use the **SmartArt** icon are as follows:

- Go to the **Insert** tab and then go to the **Illustrations** group.
- Click on the **SmartArt** option of the **Illustrations** group.

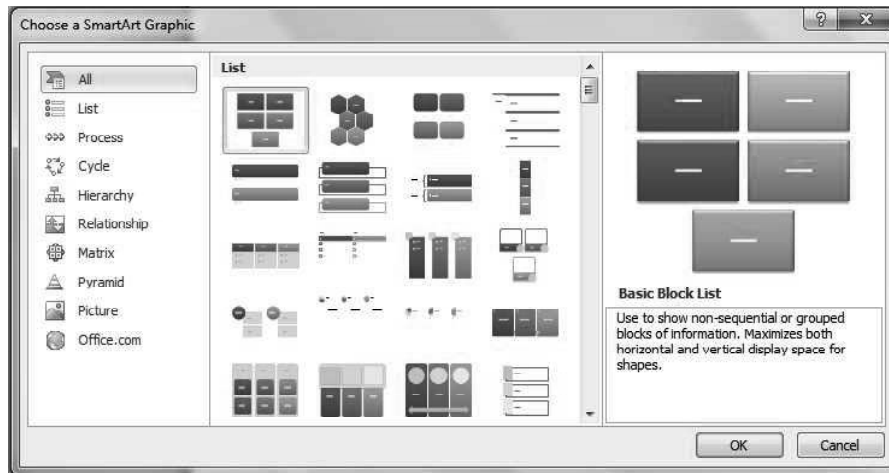
The screenshot of **SmartArt** option in **Illustrations** group is as shown:



- As you click on the **SmartArt**, a dialog box **Choose a SmartArt Graphic** appears having different categories such as **List**, **Relationship** and so on. You can choose any one of them according to the requirement.

The screenshot of **Choose a SmartArt Graphic** dialog box is as shown:

Windows, MS Office and Outlook 2010



NOTES

4. Click **OK** button of the dialog box.
5. As you add **SmartArt** graphic by clicking the **OK** button, the **SmartArt Tools** appear on the PowerPoint slide. To change particular individual shapes, click on the **Format** tab of the **SmartArt Tools** and do the changes as required. Also, a text can be added from Ribbon.
6. You can also change the layout by selecting any layout from the **Layouts** group of **Design** tab in **SmartArt Tools**.

The screenshot of **Layouts** group is as shown:



7. You can also change style, color and effects of the **SmartArt** Graphic by using the **SmartArt Styles** group.

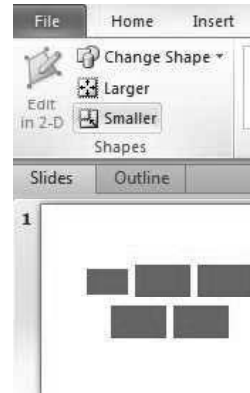
The screenshot of **SmartArt Styles** group is as shown:



8. You can change the size of the SmartArt shapes by clicking the **Smaller** and **Larger** icons from the **Shapes** group of the **Format** tab.

The screenshot of **Shapes** group is as shown:

NOTES



SmartArt Types and Their Uses

SmartArt is a method, which helps to improve the artistic view of a presentation. **SmartArt** is nothing other than the visual representation of textual data presented in a document. You can create SmartArt graphics in PowerPoint, Excel, Outlook, and Word.

You can try out different layouts, and then choose one which suits your needs. You can choose among nine SmartArt graphics.

SmartArt Graphics are available under the **SmartArt** icon or button of **Illustration** group under the **Insert** tab.

Steps for using **SmartArt** graphics in your presentation are as follows:

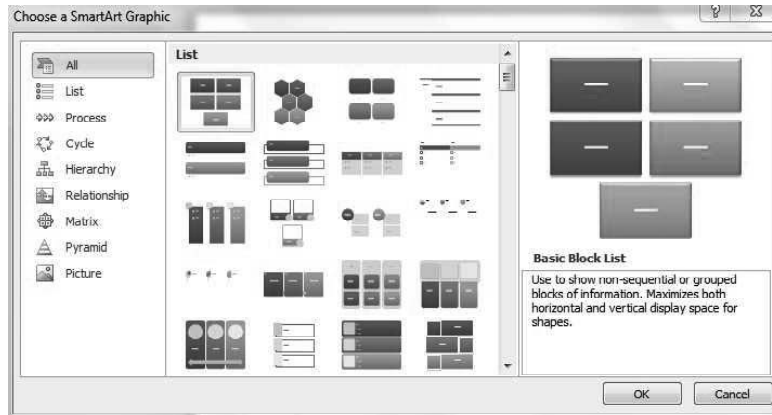
1. Go to the **Insert** tab of the menu bar and click on it.
2. When you click on it, the **Illustration** group will open displaying three icons or buttons, namely **Shapes**, **SmartArt** and **Chart**,

Screenshot displaying the **SmartArt** icon in the **Illustrations** group is as shown:



3. Now, as you click on the **SmartArt** icon, a new window or dialog box with the caption **Choose a SmartArt Graphic** will appear. This window contains several different categories or layouts of graphics such as **List**, **Relationship**, etc. You can choose any one of them, as you see fit.

Screenshot displaying the **Choose a SmartArt Graphic** dialog box is as shown:



NOTES

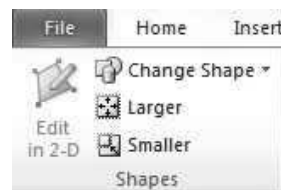
4. Choose the desired type and layout and click on the **OK** button.
5. When you click on the **OK** button, the selected **SmartArt Graphic layout** will be added to your presentation file.
6. Also, the **Design** tab of the menu bar will automatically appear, displaying the **Layouts** group and **SmartArt Styles** group. The layout of the graphic can be changed anytime by selecting a different layout from the **Layouts** group under **Design** tab in **SmartArt Tools**.
7. You can also change individual shapes by clicking on **Format** tab of the **SmartArt Tools** and making the required changes. Text can also be added from Ribbon.

The screenshot displaying the **Layouts** group and **SmartArt Styles** group of the **Design** tab is as shown:



8. You can change the size of the **SmartArt** shapes by clicking on the **Smaller** and **Larger** icons, available under the **Shapes** group of the **Format** tab.

The screenshot of **Shapes** group is as shown:



• Uses of different SmartArt Graphics or Layouts:

SmartArt Graphics represent information and ideas in visual form. It is chiefly designed for expressing concepts or ideas in a much more attractive and presentable way, as opposed to presenting the content solely in text format. In Microsoft

NOTES

PowerPoint 2010, the **eight categories of SmartArt Graphics** enable the user to access or insert all the layout types. These graphics play a key role in providing a new shape and style to your content. Further, they enhance the versatility of your document.

Some of the frequently used **SmartArt Graphics**, their purpose and the advantages they have are as follows:

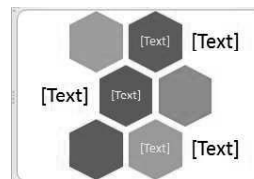
1. List

This type of layout can be used for data written in the form of a list. It transforms the bulleted list or text into a **SmartArt Shape** or Graphic. The shapes can be colored, resized and animated in order to highlight and emphasize the importance of your main points.

The layouts that belong to this group, do not present sequential information and thus, they do not have arrows for indicating the flow.

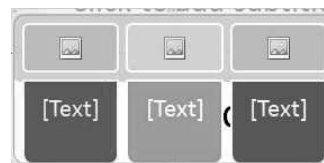
For example: Basic Block List, Alternating Hexagons, Picture Caption List, Table Hierarchy, Stacked List, Pie Process, etc. are examples of **Layouts** under the **List** category.

Screenshot displaying the pattern under **Alternating Hexagons** type of **List** is as shown:



This type of layout is used for representing content and ideas, as an interconnected series. Using this shape, **Text** appears in two levels. This implies that text is expressed inside the hexagons as well as outside the hexagons. Also, a couple of **List** layouts allow you to express your data using pictures. So, you can add pictures along with the text to highlight important information. **For Example:** **Horizontal Picture List** of the **List** group shows data in a non-sequential form and includes top shapes for inserting in pictures.

Screenshot displaying the **Horizontal Picture List** with two shapes is as shown:



2. Process

Process layout is used for expressing data having a directional or sequential flow. It is used to illustrate the sequential steps needed to complete a task or to depict phases involved in development of a product.

Process layouts can be used as horizontal steps, vertical steps, circular bending

processes, etc.

For example: Step Up Process, Ascent process, Chevron List, Segmented Process, Arrow Ribbon, Equation are some of the types of Layouts under the **Process** category.

Basic TimeLine Layout can be used to represent the sequence or flow of steps in case of any work, event, task or design procedure.

Step up Process is used to express text and present the information in an ascending or increasing manner.

Screenshots displaying the pattern or shape of the **Basic TimeLine** and **Step up** type of **Process** are as shown:



NOTES

3. Cycle

This type of **Smart Art** graphic layout is used to represent a repetitive or circular process such as Season cycles, Product development cycles, various life cycles, Website publishing cycle and many more. In these layouts, there are no defined or fixed beginning and ending points.

For example: Basic cycle, Text cycle, Block cycle, Segmented cycle, Basic Pie, Radial cycle, Cycle Matrix, Gear, etc. are some of the examples or types of Layouts under the **Cycle** category.

Basic Cycle and **Segmented Cycle** patterns or shapes are used for denoting the progression of processes, tasks or events, in circular flow.

Radial Cycle is used to show how the outer pieces of information or data connect to the central data and what the surrounding shapes are contributing to the central hub.

Screenshots displaying the **Segmented Cycle** and **Radial Cycle SmartArt Graphic** are as shown:



4. Hierarchy

Hierarchy type is used to express information in a hierarchical manner. It basically denotes and emphasizes the relationships in an organization. Hierarchy can also

NOTES

be used to represent tree structures like decision trees, binary trees, B-trees. It is mostly used in case of organization charts, related to companies or employment.

Hierarchy Type is used to show the various levels of organizations and depicting the sequence of steps to be followed, for the purpose of reporting and handling the employment system.

For Example: Organization Chart, Hierarchy, Labelled Hierarchy, Horizontal Hierarchy, Hierarchy List and Lined list are some of the types of Layouts under the **Hierarchy** category.

Screenshots displaying the **Organization Chart** and **Hierarchy List** are as shown:



Organization Chart is mainly used for expressing the relationship among the various levels of the organization or company. Also, **Hierarchy List** shown in the second screenshot is used for denoting and representing the relationship and flow of information in a collective manner, as groups.

5. Relationship

This layout is used to express or denote non-progressive and non-hierarchical relationships. **Relationship** layout can be used to show conceptual relationships such as overlapping between two or more sets of things. Venn diagrams, showing the overlapping areas at the centre of intersection, are a good example of the **Relationship SmartArt Graphics**.

Thus, this type of layout is used for illustrating connections and depicting how parts or pieces of information are related to the whole.

For example: Circle relationship, Arrow Ribbon, Basic Venn, Basic Target, Radial cluster, and Diverging Radial are some of the common layouts under the **Relationship** category.

Screenshots displaying **Arrow Ribbon** and **Basic Target Relationship** are as shown:



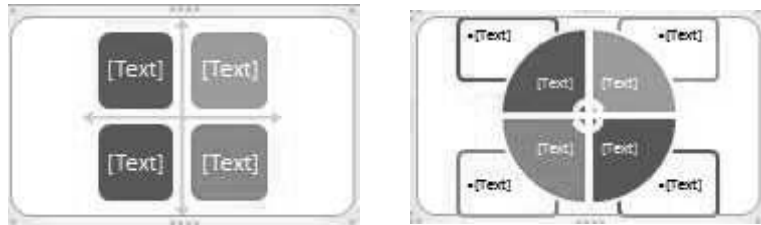
Arrow Ribbon (displayed in the first screenshot) is used for denoting concepts either as related or contradicting. **Basic Target** layout (shown in the second screenshot) is used to show gradations or containment, while the **Basic Radial** layout is used to collectively highlight processes in order to merge various concepts into a single central idea.

6. Matrix

Matrix layout is used to represent the relationship between pieces of information or data, and express this information as a collective unit called matrix. It mainly classifies or expresses information in two dimensional forms.

For example: Basic Matrix, Titled matrix, Grid Matrix and Cycle Matrix are some of the popular layouts under the **Matrix** category.

Screenshots displaying the **Grid** matrix and **Cycle** Matrix are as shown:



In the first screenshot of **Grid** layout, the four lines showing the axes of level 1 data appear in quadrants. The concepts are placed along the two axes.

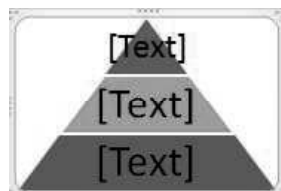
In **Cycle** matrix, first four lines of level 1 are shown in the form of a Pie and a rectangular shape is used for showing level 2 data.

7. Pyramid

It is used to show hierarchical or proportional relationships, with largest component being placed at one end and the smallest component at the other end. Thus, all the patterns that belong to this category have an upwardly pointing shape. This layout is basically used for expressing information or relationships in two approaches, either top to down or bottom to up.

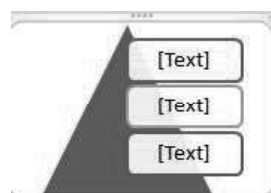
For example: Basic Pyramid, Inverted Pyramid, Pyramid List and Segmented Pyramid are some of the popular layouts under the **Pyramid** category.

Screenshot displaying the **Basic Pyramid** is as shown:



Another type or layout of the **Pyramid** category known as **Pyramid List**, can be used to write text in rectangular boxes provided on top of the pyramid background.

Screenshot displaying the shape of the **Pyramid List** is as shown:



NOTES

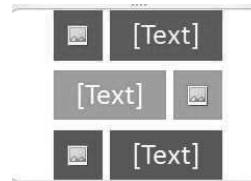
8. Picture

Picture layout is used for representing text and information using pictures. It gives your presentation an authentic touch and makes it more interactive, as well as attractive.

NOTES

For example: Accented Picture, Alternating Picture Blocks, Bubble Picture List, Captioned Pictures, Snapshot Picture List, Vertical Picture List and Spiral Picture are some of the widely used types of Pyramid category layouts.

Screenshot displaying the **Alternating Picture Blocks** is as shown:



Screenshot displaying **Vertical Picture Accent List** layout of **Picture** category is as shown:

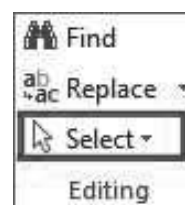


This layout is used for denoting information in non-sequential form by grouping blocks. It also contains circles attached with text blocks, for attaching and showing the pictures.

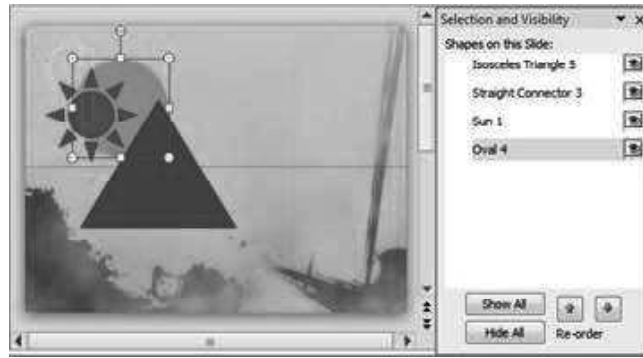
5.6.6 Working with Objects

Selecting objects

On the **Home** tab, in the **Editing** group, click **Select** and do one of the following:





- To select all of the objects and placeholders on the slide, click **Select All**.
- To select objects that are hidden, stacked, or behind text, click **Select Objects**, and then draw a box over the objects.
- To open the selection pane, where you can select, multi-select, show, hide, or change the order of objects, click **Selection Pane**, and then click the options that you want.



NOTES

Grouping, ungrouping and regrouping objects


Grouping shapes, pictures or object

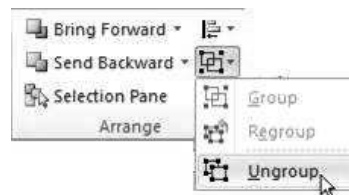
1. Press and hold CTRL while you select the shapes, pictures, or objects that you want to group.
2. Do one of the following:
 - o To group shapes and other objects: under **Drawing Tools**, on the **Format** tab, in the **Arrange** group, Click  , and then click **Group** .
 - o To group pictures: under **Picture Tools**, on the **Format** tab, in the **Arrange** group, Click  , and then click **Group**.

If you do not see the **Drawing Tools** or **Picture Tools** tabs, make sure that you selected a shape, picture, or other object. You might have to double-click the object to open the **Format** tab.

Ungrouping shapes, pictures or object

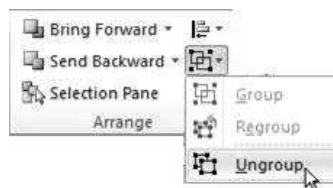
To ungroup a group of shapes, pictures, or other objects (for example, if you want to move a group but leave one shape behind or make extensive changes to one shape without changing the other shapes), do the following:

1. Select the group that you want to ungroup.
2. Do one of the following:
 - o To ungroup shapes or other objects, under **Drawing Tools**, on the **Format** tab, in the **Arrange** group, click **Group**  , and then Click **Ungroup**.





- o To ungroup pictures, under **Picture Tools**, on the **Format** tab, in the **Arrange** group, Click  , and then click **Ungroup**.

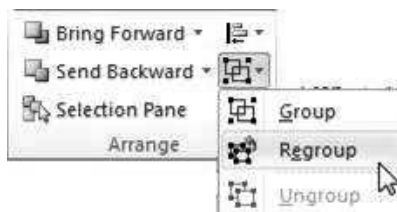
NOTES



If you do not see the **Drawing Tools**, **Picture Tools**, or **Format** tabs, make sure that you selected a group of shapes, pictures, or other objects. You might have to double-click the object to open the **Format** tab.

Regrouping shapes, pictures or object

1. Select any one of the shapes or objects that were previously in a group.
2. Do one of the following:
 - o To regroup shapes and objects: under **Drawing Tools**, on the **Format** tab, in the **Arrange** group, click , and then click **Regroup**.
 - o To regroup pictures: under **Picture Tools**, on the **Format** tab, in the **Arrange** group, click , and then click **Regroup**.

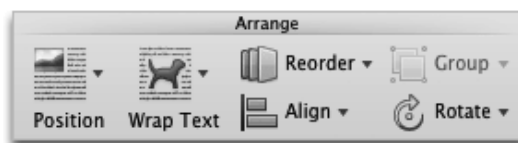


If you do not see the **Drawing Tools**, **Picture Tools**, or **Format** tabs, make sure that you selected a group of shapes, pictures, or other objects. You might have to double-click the object to open the **Format** tab.

Aligning object on a slide

Follow the steps given below to align object on a slide.

1. Hold down Shift, click the objects that you want to align, and then click the **Format** tab.
2. Click **Arrange > Align > Align to Slide**.

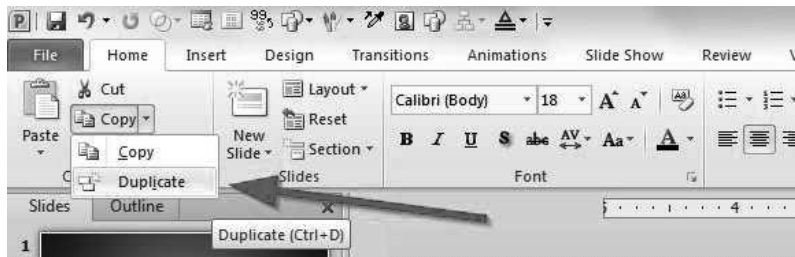


3. Click **Arrange > Align**, and then click the alignment that you want.

Cut, Copy, Paste and Duplicating an object

You can cut, copy, paste and duplicate an object in a slide using the illustration given below. In case of duplicating an object, Click to select the object that you want to **duplicate**. Hover over the object with your mouse and press the CTRL

key until you see the plus symbol next to the arrow. While still pressing the CTRL key, click and drag the object to the new location.



NOTES

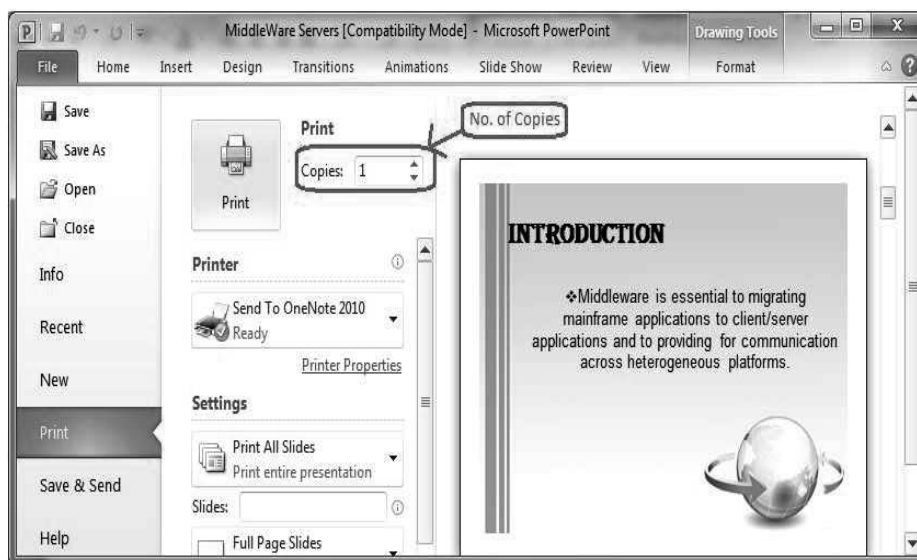
Printing Slides

You can modify the common print settings by using the drop-down menus. The print previews are automatically adjusted to allow you to see how the pages will appear before actually printing them.

Steps to print the slides are as follows:

1. Select the **File** tab in the Ribbon.
2. Select and click on the **Print** option. Then, fill the number of copies required in the **Copies** drop-down box.
3. As you click on the **Print** option, **Print Preview** appears on the right side of the PowerPoint window.
4. Click on the **Print All Slides** drop-down box in the **Settings** group to select the options like **Print All Slides**, **Print Current Slide**, and **Custom Range** of slides.
5. After checking the **Print Preview**, click on the **Print** button.

The screenshot of printing the presentation is as shown:



• Print Layout

There are many options to print the slides in different formats. These options are available in print layout tab. The options in Print Layout are as follows:

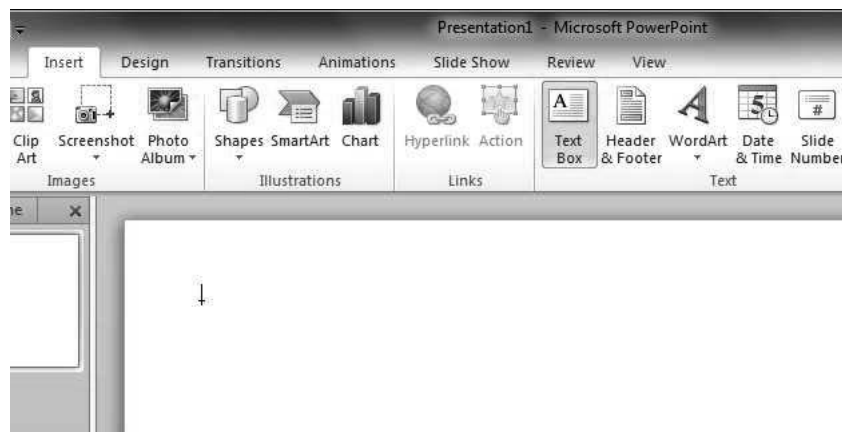
NOTES

- **Full Page Slides:** It prints one slide on a full page.
- **Handouts:** There are various options available in the **Handouts** group. You can choose any one of them like **1 Slide, 2 Slides**. Also, you can print slides in horizontal and vertical manner by using **4 Slides Horizontal, 4 Slides Vertical**, etc.
- **Notes Page:** It prints one slide per page including notes that are added in the **Notes Pane** while making the presentation.
- **Outlines View:** It provides the facility to print the slides in an outline format.

Putting Text on Slides

To add text anywhere on slides, follow the steps given below.

1. Click on Text box command in the Text group that is available in the Insert menu.
2. Click and hold down the left mouse button while you drag the mouse. Release the mouse button when the size of box is as per your text size.
3. Click inside the text box and start typing.



Using a Shape as a Text Box / Typing Text in Shapes

Microsoft Office PowerPoint 2010 allows its users to use **Shapes** as **Text Boxes**, which implies that you can add a **Shape** to your slide and then you can use it as a **Text Box** to write text inside it, making presentations more beautiful and informative.

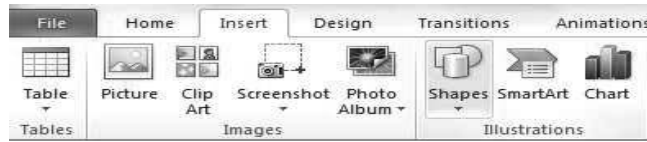
• Turning a Shape into a Text Box:

Shapes can be added to Microsoft Office PowerPoint 2010, using the **Shapes** option under the **Illustrations** group of the **Insert Tab**. You can insert any shape of your choice and use it to write text.

Steps to add a rectangular shape and use it as a **Text Box** are as follows:

1. First go to the **Insert** menu of the menu bar and click on it.
2. When you click on the **Insert** tab, you will find **Illustrations** group containing three options namely, **Shapes, SmartArt** and **Chart**.

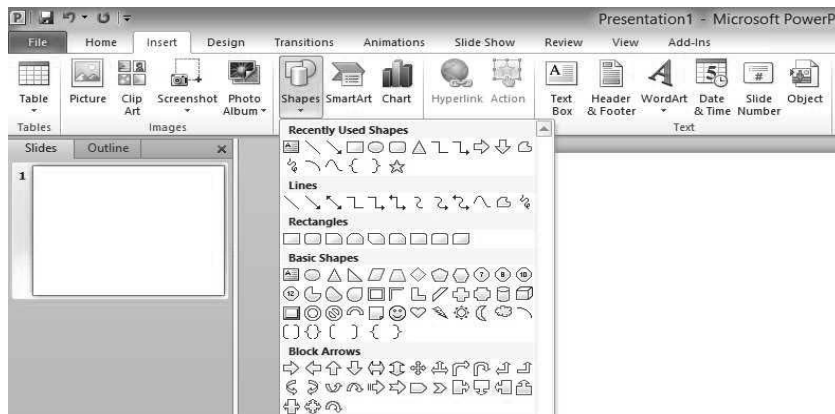
Screenshot displaying **Shapes** under the **Illustrations** group of the **Insert** tab is as shown:



NOTES

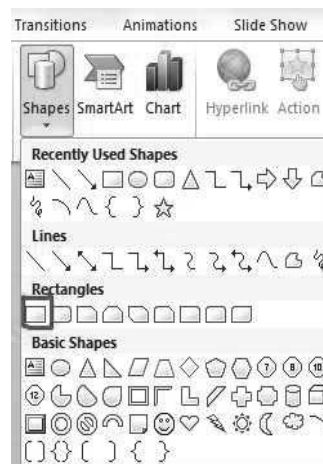
3. Now, click on the drop down arrow of the **Shapes** icon, to select the ready-made list of available shapes like Lines, Rectangles, Basic Shapes, Block Arrows, Equation Shapes, Flow charts etc.

Screenshot displaying several options for selecting and inserting shapes from **Shapes** option is as shown:



4. Now, suppose you want to insert the rectangle shape. Click on the first shape under the Rectangles section.

Screenshot displaying the selection of rectangle shape is as shown:



5. Now, after selecting the desired rectangular shape, drag the cursor and click on the area where you want to create your rectangle. Move the cursor away to create the rectangle.

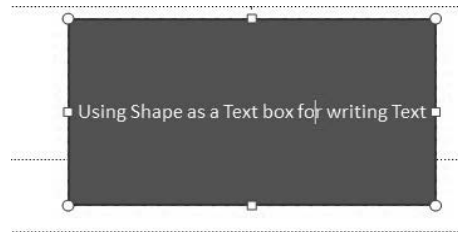
Screenshot displaying the rectangle inserted into the PowerPoint file is as shown:



NOTES

- Now, as the rectangle is created and inserted, just click on the Rectangle and wait for it to show a blinking cursor in the middle. This cursor will allow you to enter text.
- Enter the text in the **Rectangle** and press **Enter**. Thus, we see how, Microsoft PowerPoint 2010 provides us the facility of turning any shape into a text box.

Screenshot of the **Rectangle Shape** seen in the PowerPoint slide, after entering the text is as shown:



• Turning a Text Box into a Shape:

It is also possible to convert the text box into a shape and apply various **Quick Styles** options to it, in order to resize, colour and modify according to your choice. Steps to turn a text box into a shape are as follows:

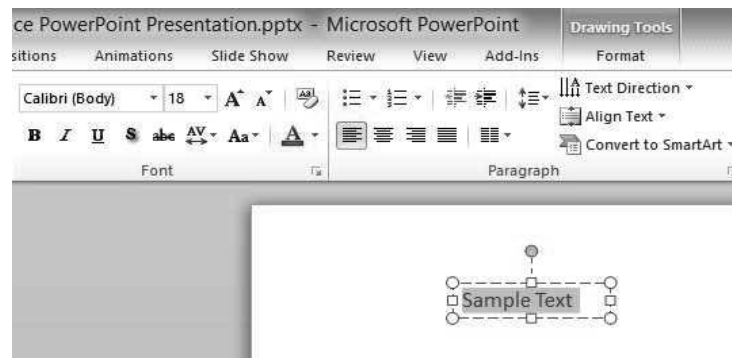
- Go to the **Insert** tab of the ribbon view and click on the **Text Box** icon of the **Text** group.

Screenshot displaying the **Text Box** icon, available in the **Text** group of the **Insert** tab is as shown:



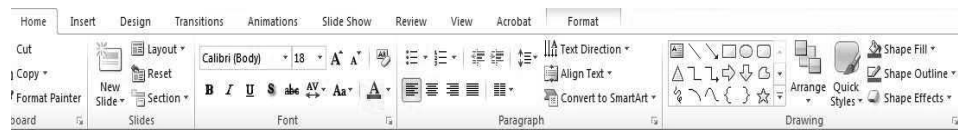
- When you click on the **Text Box**, the shape of the cursor on the screen changes.
- Now, click on the desired location and drag the cursor to insert a text box.

Screenshot of a **Text Box** with sample text is as shown:



4. Now, go to the **Home** tab, and use the shape styling options of the **Drawing** group to use this text box as a shape.

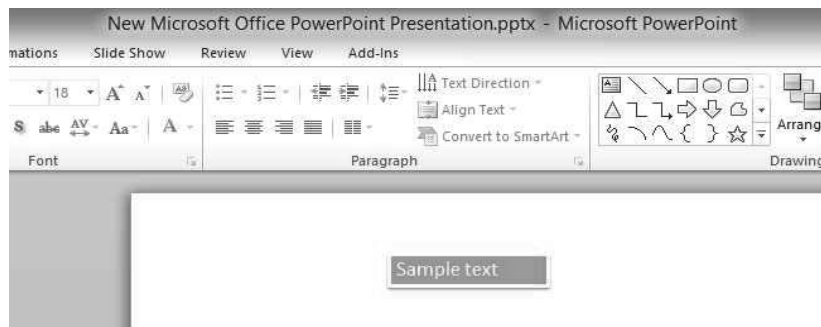
Screenshot of the **Drawing** group of the **Home** tab is as shown:



NOTES

5. Choose from the **Quick Styles**, **Shape fill**, **Shape Outline** or **Shape Effects** to add a shape effect.

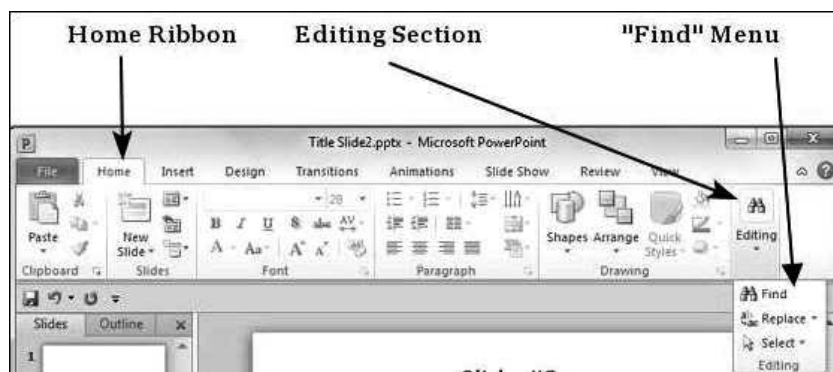
Screenshot of a text box, after adding **Quick Style** into it, is as shown:



Finding and Replacing Text

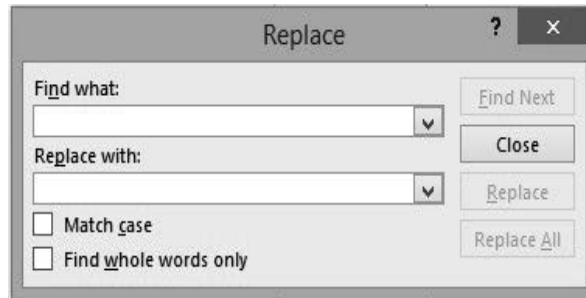
To find and replace the text in slides, follow the steps given below.

1. Click on Home tab.
2. Go to the editing group where you will get Find and replace button. Click on it.



3. A new window will open as shown below. Type the text which you want to find and its replacement also. After that click on Replace button. Using Replace All button, you can replace every occurrence of text in the slide.

NOTES



New Features in PowerPoint 2016/ Difference between PowerPoint 2010 and 2016

PowerPoint 2016 is not much different from PowerPoint 2013 or 2010. Here also, all the menu items and often-used features are the same. The screenshot of ribbon interface is given below.



Although there aren't any dramatic changes, the latest version, PowerPoint 2016, brings some new enhancements to this top presentation programme, such as:

- **PowerPoint Designer:** With the new PowerPoint Designer, as soon as you insert an image, the *Design Ideas* panel opens on the right, offering several layout options for the graphical design of the current slide.
- **Morph Transition:** PowerPoint 2016 adds a brand new transition type called *Morph*. This transition can help you make smooth animations, transitions, and object movements across the slides in your presentation.
- **Quick Shape Format:** Draw a shape on a slide and you'll see an increase to the number of "preset" styles available, another enhancement in PowerPoint 2016.
- **Six New Chart Types:** Microsoft Office 2016 has six new charts to show off your work in PowerPoint, Excel, and Word. New chart types include Waterfall, Treemap, Histogram (including Pareto), Box and Whisker, and Sunburst.
- **Screen recording:** you can now include screen recordings in your PowerPoint presentations with just a few clicks.
- **Simple Collaboration and Sharing:** Collaboration is more and more integrated into Office 2016. Choose **Share** on the ribbon to share your presentation with others on SharePoint, OneDrive, or another online location. With real time co-authoring, you can now not only see when others are working on the same shared presentation but also see an indicator that shows you exactly what slides they're working on, and where in the slides they are making changes.

- **Improved Version History:** Now you can go to **File > History** to see a complete list of changes that have been made to your presentation and even access earlier versions. **Note:** This feature is only supported for files stored on OneDrive for Business or SharePoint.
- **Tell Me:** One of the other PowerPoint 2016 new features is the “Tell Me” help tool. Look at the right side of the PowerPoint 2016 Ribbon menu after the last tab, to see a light bulb that says “Tell me what you want to do.” Just enter the words and phrases that describe the steps you want to do next and quickly get the features or actions you want to use.
- **Smart Look Up:** The Smart Lookup feature, which is available in most of the Office 2016 programs, gives you access to quick searches from inside of PowerPoint 2016.

NOTES

5.7 MICROSOFT OUTLOOK 2010

When you click on MS Outlook 2010 for the first time, you will see an e-mail configuration window. A startup wizard guides the process of creating a new e-mail account. You will have to configure an e-mail account before sending and receiving e-mail through Microsoft Outlook 2010.

Steps to **configure e-mail** account when you start MS Outlook 2010 for the first time are as follows:

1. Click on the **Start** menu option in Windows.
2. Click on the **All Programs** options for opening list of your programs.
3. Click on **Microsoft Office 2010**.
4. Choose the option **Microsoft Outlook 2010**. It displays an **Add New Account dialog** box to configure an e-mail account.
5. Enter your name, e-mail ID and password and retype the password in following window.

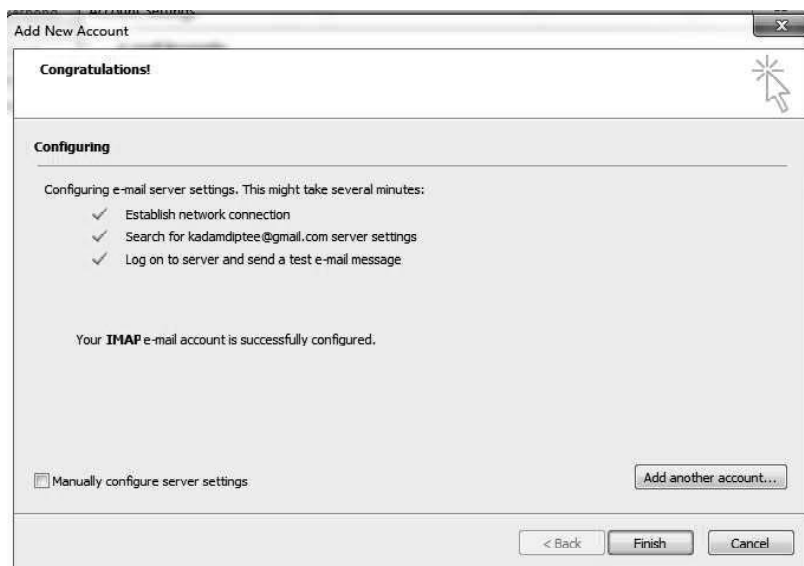
The screenshot of **configuring e-mail account** for the first time is as shown:

The screenshot shows the 'Add New Account' dialog box. It has a title bar with 'Add New Account' and a close button. The main content area is titled 'Auto Account Setup' and includes the instruction 'Click Next to connect to the mail server and automatically configure your account settings.' Below this, there are three radio button options: 'E-mail Account' (selected), 'Text Messaging (SMS)', and 'Manually configure server settings or additional server types'. The 'E-mail Account' section includes four text input fields: 'Your Name' (with example 'Ellen Adams'), 'E-mail Address' (with example 'ellen@contoso.com'), 'Password', and 'Retype Password' (with the instruction 'Type the password your Internet service provider has given you.'). At the bottom, there are three buttons: '< Back', 'Next >', and 'Cancel'.

NOTES

6. Configuring the server settings for an e-mail account can take several minutes. A progress sign indicates that configuration of e-mail account is under process.

The screenshot of **Add New Account** dialog box is as shown:



7. You can also configure more accounts by clicking on **Add another account...** button.
8. Click **Finish** button.

5.7.1 MS Outlook and E-mail Accounts

MS Outlook provides mailing tool which helps in sending or receiving mails. MS Outlook allows access to e-mail account on your computer even without internet, that is, you can view and manage your e-mails when you are offline. A profile in Outlook is created by adding information such as username, e-mail id, password, etc. Most of the people create one e-mail account to send or receive personal e-mails; however, you can also create more than one e-mail account. One can access more than one e-mail account on the same computer. The e-mail accounts are secure in separate profiles with different names. It is the window registry that retrieves the profile information when Outlook starts.

Steps for adding e-mail accounts are as follows:

1. Click on the **File** tab and select **Info**.
2. Select **Account Settings** and click on the **Account Settings** option.

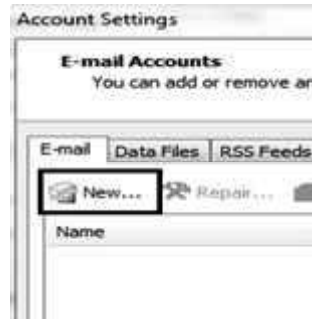
The screenshot displaying all options for **setting the account** is as shown:

NOTES



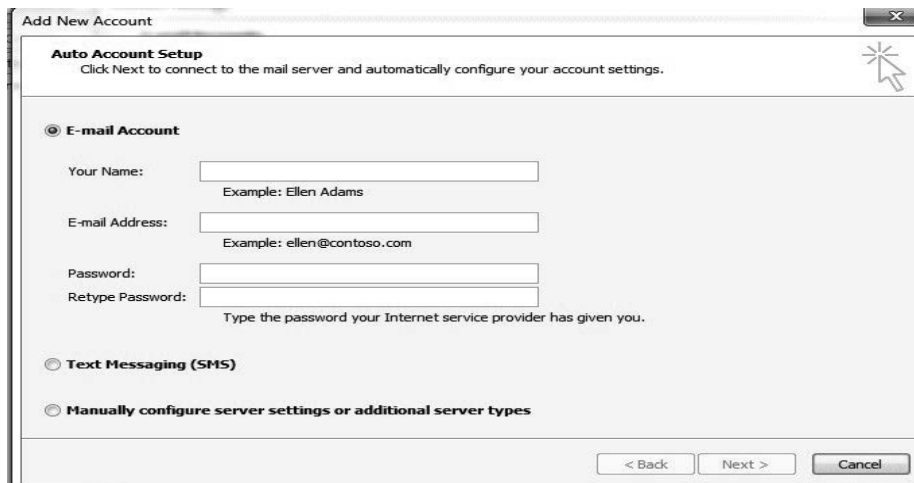
3. Click on **New...** option.

The screenshot to display **New...** is as shown:



4. Enter your name, e-mail id and password and retype the password in following window.

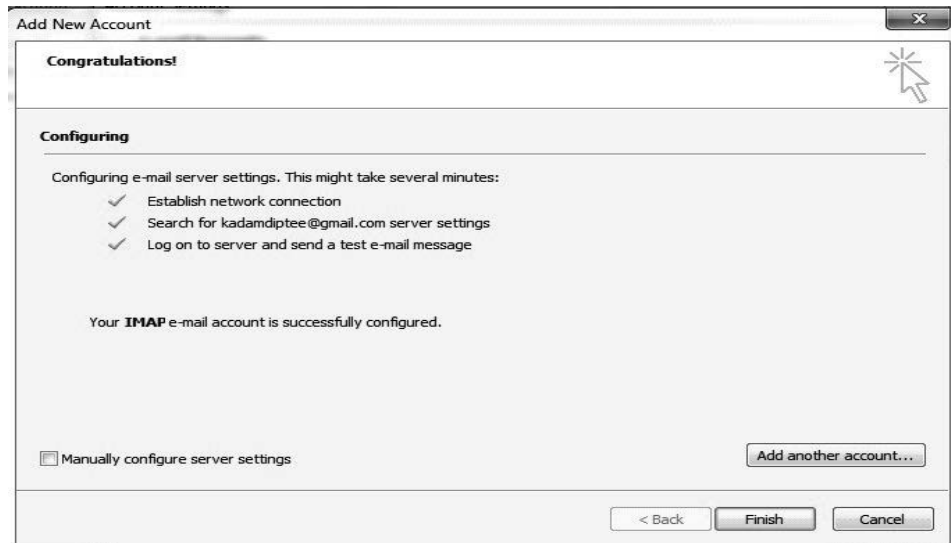
The screenshot of **Add New Account** dialog box is as shown:



5. Click **Next** button. Configuring server setting for an e-mail account can take several minutes. A progress sign indicates that configuration of e-mail account is processing.

The screenshot for configuring the e-mail server setting is as shown:

NOTES



6. Click on **Finish** button.

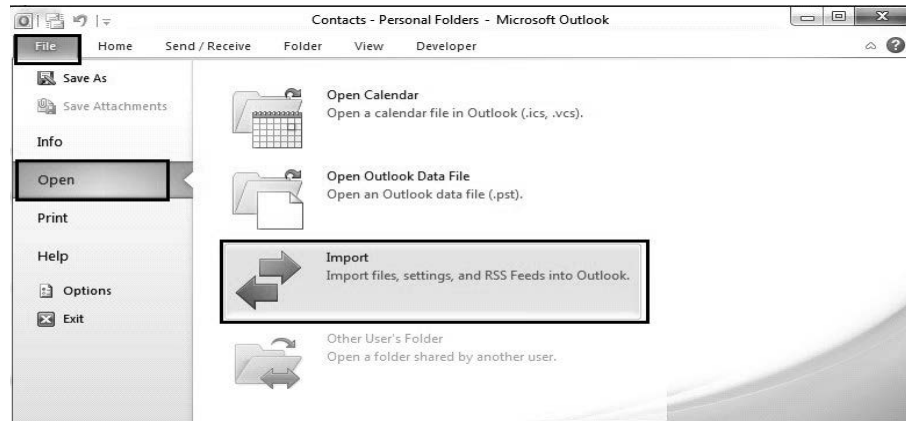
5.7.2 Importing and Exporting Data

MS Outlook provides Import and Export wizard to import data in Microsoft Outlook and to export Outlook data to a text file. You can import and export contact, appointments, tasks, notes, e-mail and e-mail folder.

Steps to **import** data are as follows:

1. Click the **File** tab, select **Open** and then select **Import**.

The screenshot for selecting **Import** option is as shown:



2. Select **Import from another program or file** from **Choose an action to perform** list of **Import and Export Wizard** dialog box. Afterwards, click on the **Next** option.

The screenshot of **choose an action to perform** from the list is as shown:

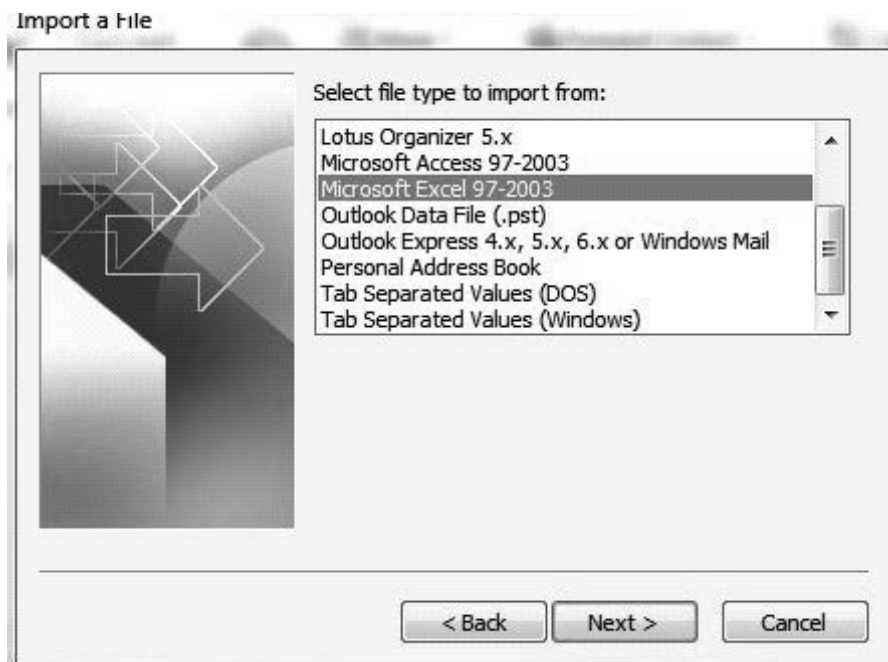
Windows, MS Office and Outlook 2010



NOTES

3. Select the file type to import from **Select file type to import from** list.

The screenshot of **Select file type to import from** in **Import a File** dialog box is as shown:



4. Click on the **Next** button.

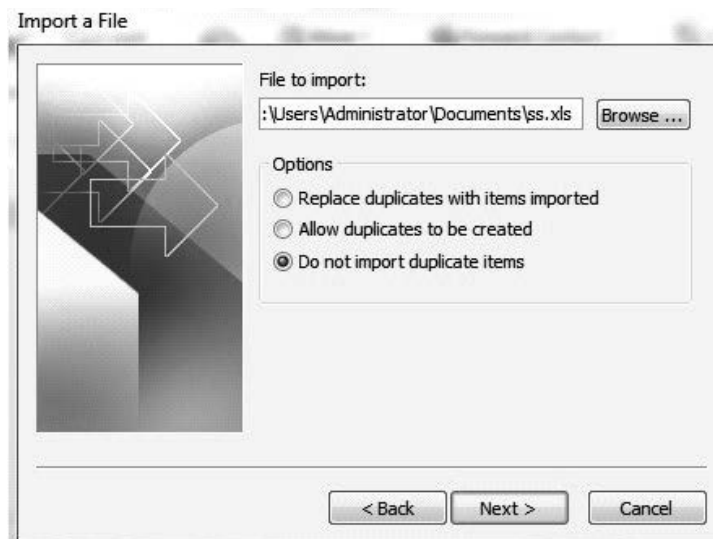
5. Type file name with path in the **File to import** box.

Various importing options are as follows:

NOTES

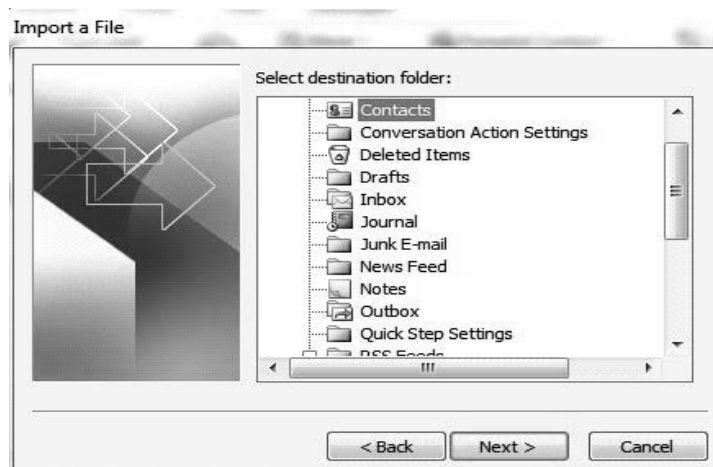
- If MS Outlook data already exists and you want to replace this data with new data, then select **Replace duplicates with items imported** option.
- If MS Outlook data already exists and you want to create imported data as a second data, then select **Allow duplicates to be created** option.
- If MS Outlook data already exists and you do not want to import same data, then select **Do not import duplicate items** option.

The screenshot displaying three options for importing the file is as shown:



6. Select the folder that you want to import from the **Select destination folder** list and click on the **Next** button.

The screenshot for selecting the destination folder is as shown:



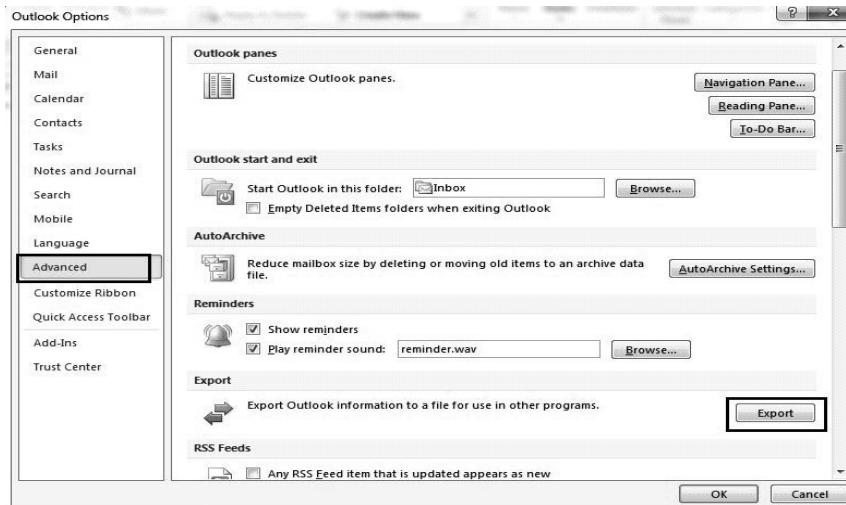
7. Click **Next** button in the **Import a File** dialog box.
8. Click **Finish**.

In addition, you can export contacts to a file such as Excel or database application.

Steps to **Export data** are as follows:

1. Click on the **File** tab.
2. Click on the **Option**.
3. Click on the **Advanced** and click on **Export** button in **Export section**.

The screenshot for selecting **Export option** is as shown:



NOTES

4. Click on **Export to a file** in the **Choose an action to perform** list.

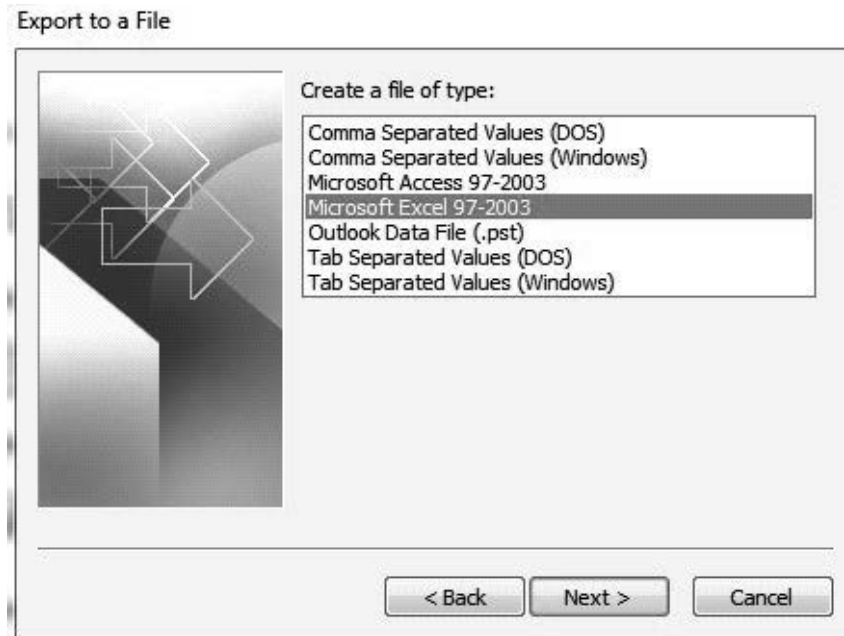
The screenshot of **Import and Export Wizard** is as shown:



5. Click **Next** button.
6. Select a format for the file from **Create a file of type** list and click on **Next** button.

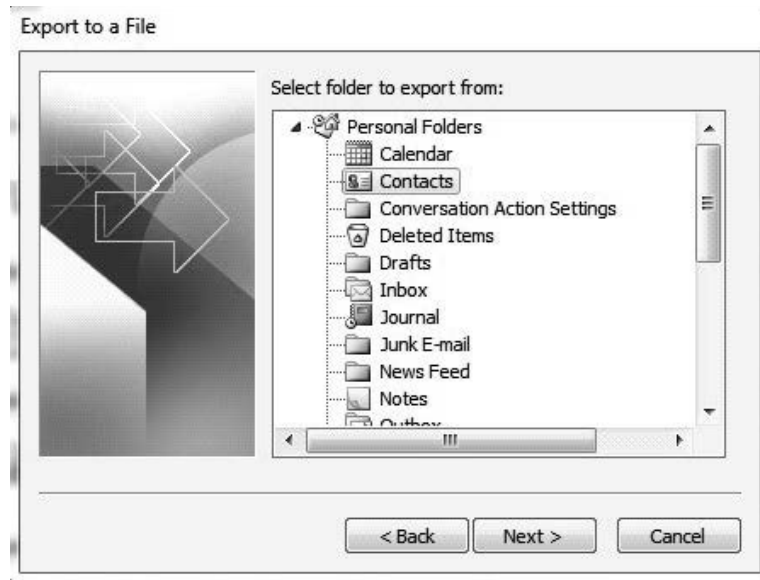
The screenshot of **Export to a file** with option **Microsoft Excel 97-2003** in **Create a file of type** is as shown:

NOTES



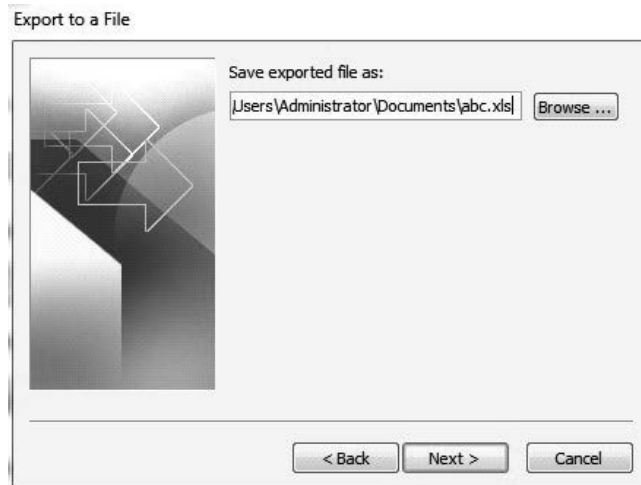
7. Click the folder from which you want to export in the **Select folder to export from** list and click **Next** button.

The screenshot of **Export to a File** with **Select folder to export from** dialog box is as shown:



8. Type the file name with path in '**Save exported file as**' box and click **Next** button.

The screenshot of **Save exported file as** for saving the file at appropriate location is as shown:

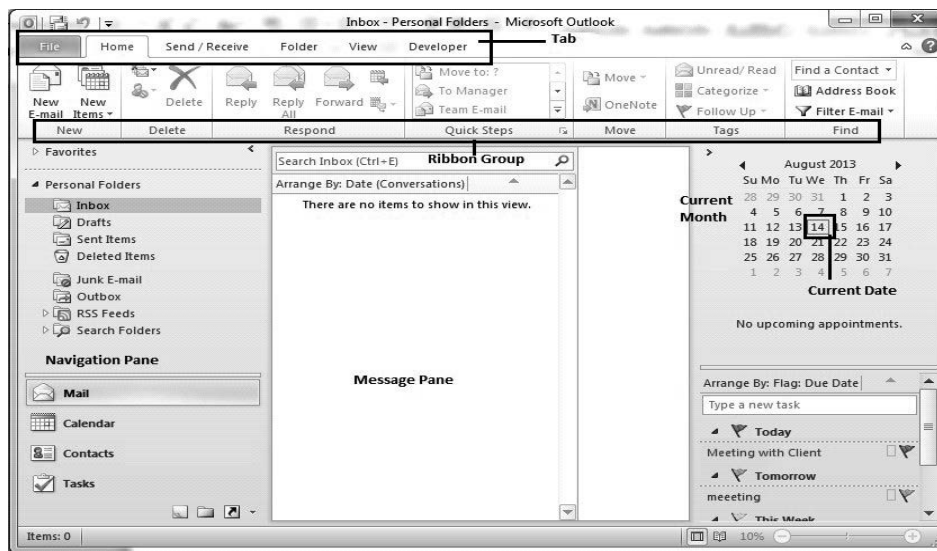


9. Click **Next** button in the **Export to a File** dialog box.
10. Click **Finish** button.

Navigating the Outlook

Navigation Pane is available on the left side of the MS Outlook window. It is used to switch different areas of Outlook by clicking on each tool. There are five main tools such as Mail, Calendar, Contacts, Tasks and Notes. When you are working with a specific tool, Navigation Pane displays folders for each tool. For example, if you click on the Inbox folder in Mail tool, it shows mail in message pane. The bottom of Navigation Pane contains Notes, Folder List and Shortcuts.

The screenshot of **Outlook** window is as shown:



5.7.3 Managing E-Mail in Outlook

E-mail is one of the primary functions of Outlook. The e-mails are usually received or sent in bulk, therefore, they need to be organized and managed. In **MS Outlook**, there are various tools available for keeping and organizing the e-mails

NOTES

efficiently. Let's first learn working with mails.

Working with E-mails

NOTES

The e-mails are the electronic messages which are sent or received digitally from one person to another. In e-mail page of the Outlook, there are four default folders which are **Inbox, Sent Items, Drafts and Deleted Items**. Inbox contains received e-mails, Sent Items contains e-mail that were sent, Drafts contains saved e-mails and deleted Items contains deleted e-mails. Categorize and Rules option also help in storing e-mails in a more organized manner. Let's learn about Folders, Categorize and Rules.

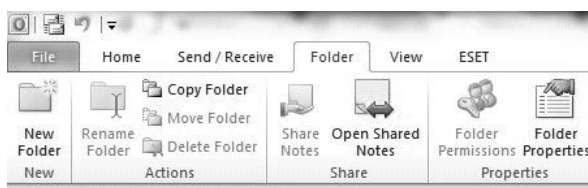
• Folders

Folders help in storing e-mails of same type at one location or folder. You can create folders on you own within the default folders. You can also perform other operations on already created folders such as copy a folder, move a folder from one location to other folder and delete a folder.

Steps to create a new folder are as follows:

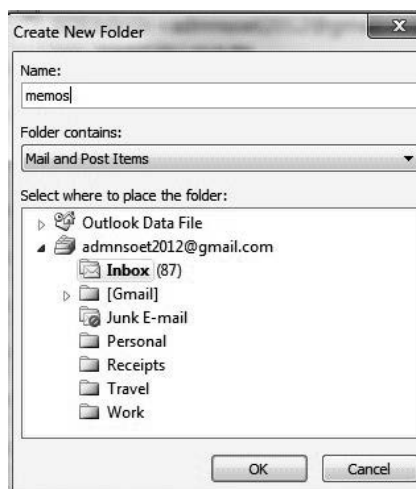
1. Select **Folder** tab from the Ribbon for accessing folder commands.

The screenshot of **Folder** tab is as shown:



2. Click on **New Folder** option of the **Folder** tab. Then a **Create New Folder** dialog box appears on the screen.

The screenshot of **Create New Folder** dialog box is as shown:



3. Write the name of new folder, say memos, in the **Name** field.
4. Select the location where folder has to be created.
5. Click **OK** button. A new folder is created at the selected location.

The screenshot of created new folder **memos** is as shown:

Windows, MS Office and
Outlook 2010



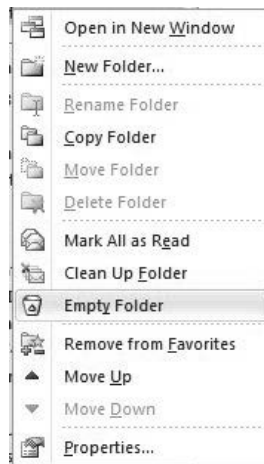
NOTES

The e-mails that are deleted automatically move to the Deleted Items folder.

Steps to **delete e-mail** from an e-mail account are as follows:

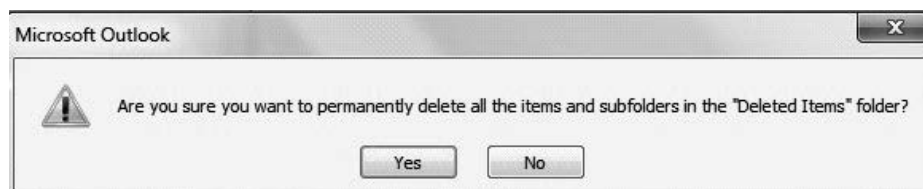
1. Go to the **Deleted Items** folder and right click on it.
2. In the drop down list, select the option **Empty Folder**.

The screenshot of list appeared after right click on **deleted folder** is as shown:



3. After selecting **Empty Folder** option, a message box appears on the screen. Then, click on **Yes** button. It will delete all the e-mails permanently.

The screenshot of **Microsoft Outlook** message box is as shown:



• Using Categories

Categories help just like folders in categorizing mails. However, multiple categories can be applied to a number of e-mails. We can also rename categories, create new category and change the category color. Categories are also used for finding and managing the e-mails.

To customize categories:

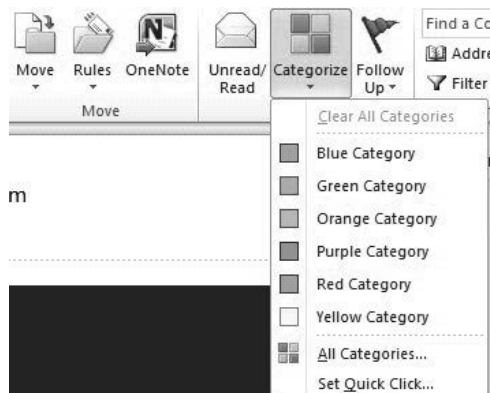
There are six categories by default. You can customize categories before using them to organize your e-mails.

NOTES

Steps to categorize a mail are as follows:

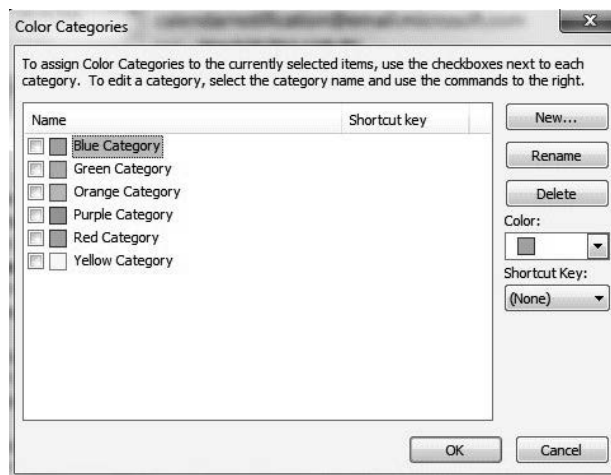
1. Click on the **Home** tab and select the **Categorize** option from the **Ribbon** view.

The screenshot of **Categorize** option is as shown:



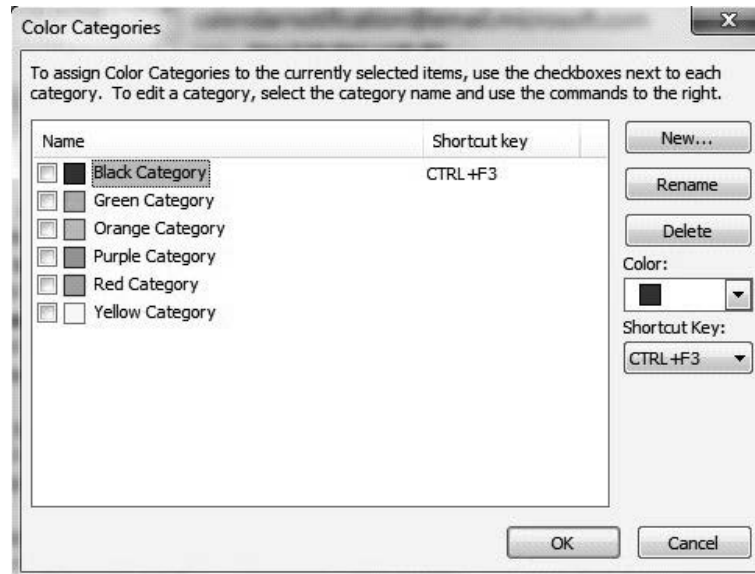
2. From the drop down list of categories, select the **All Categories....** option. A **Color Categories** dialog box appears on the screen.

The screenshot of **Color Categories** dialog box is as shown:



3. You can create a new category, rename a category or delete any category using options in **Color Categories** dialog box. For example, to rename a category, click on the check box near the category name, click on the **Rename** button and give the new name to the category.
4. You can also change the color and set a shortcut for the category. To set the color, first select the category. Then, select new color from drop down list named **Color** option. To set the shortcut, select a key from the drop down list **Shortcut Key**.
5. To apply the changes made by you in the **Color Category** dialog box and click **OK** button.

The screenshot of modified category (with color changed and shortcut key assigned) is as shown:



NOTES

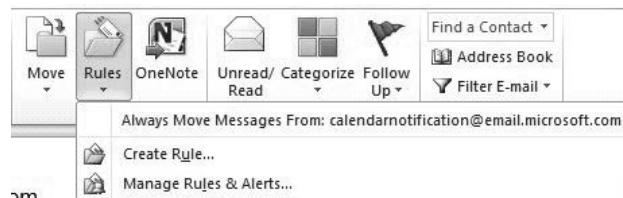
• Using Rules

Rules are the conditions specified by the user to organize mails. These are created for a specific sender or a recipient. With the help of Rules, e-mails in inbox or sent items are automatically moved or deleted depending upon the conditions set.

Steps to **create a new rule** are as follows:

1. Click on **Home** tab. Then click on **Rules** option of Home tab.
2. As you click on **Rules** option a List appears. Select **Manage Rules & Alerts...** option from the list.

The screenshot of **Rules** option is as shown:



3. A dialog box **Rules and Alerts** appears on screen. Click on **New Rule...** button which lies inside the **E-mail Rules** tab of dialog box.

The screenshot of **Rules and Alerts** dialog box is as shown:



4. After clicking on **New Rule** button, a **Rules Wizard** dialog box appears on the screen.

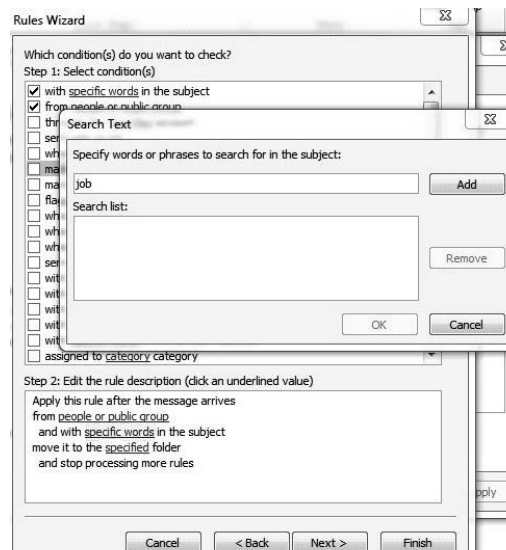
The screenshot of **Rules wizard** is as shown:

NOTES



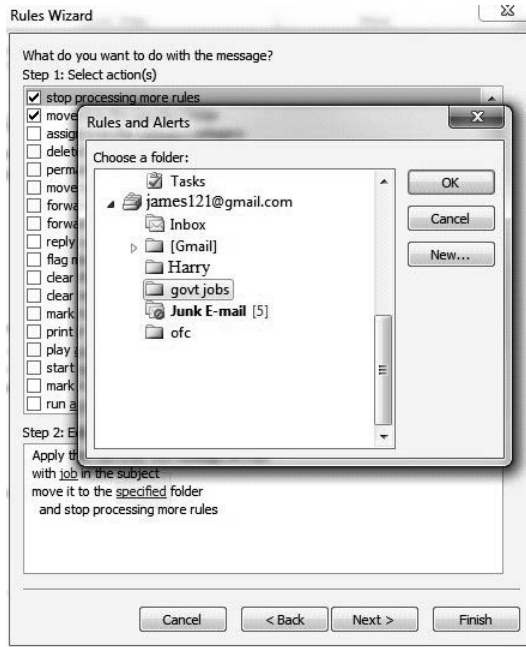
5. Select any template from **Select a template** and then, click on **Next** button.
6. Check the checkboxes according to your need. For example, check the checkboxes having text **with specific words in the subject** of **step 1: Select condition(s)**.
7. Click on link **Specific Words** of **step 2** as highlighted in the screenshot. A **Search Text** dialog box appears on screen. Type the word **job** in the textbox and then, click on **Add** button.

The screenshot for adding word in **Search Text** is as shown:



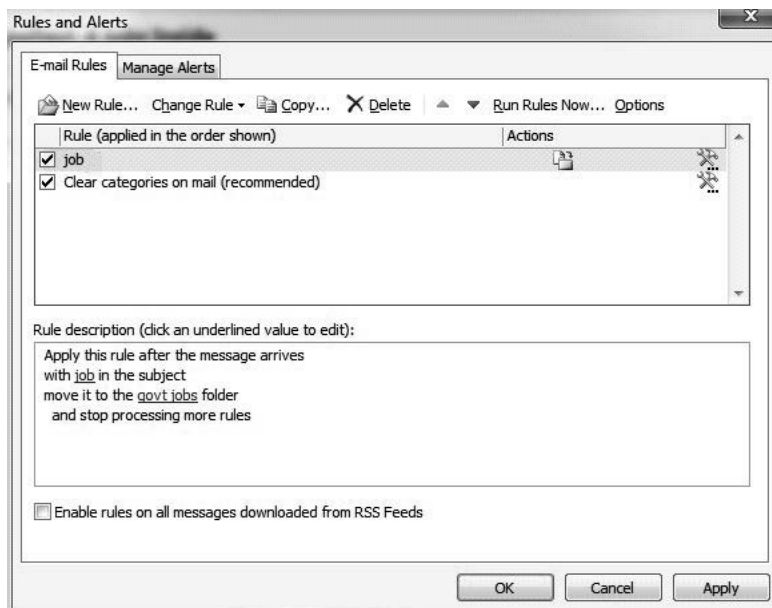
- Click on link **Specified** of **step 2** as highlighted in the following figure. Another **Rules and Alerts** dialog box appears on screen. Select the folder where you want to save your e-mail which contains **job** word in subject. Then, click **OK** button.

The screenshot to select the destination folder for moving e-mails having **job** as a word in subject is as shown:



- After selecting the destination folder, click on **Finish** button.
- Click **OK** button of **Rules and Alerts** dialog box.

The screenshot of new added rule is as shown:



NOTES

NOTES

Tips on Using your Mailbox

Microsoft Outlook is used for managing your mailbox. However, managing your mailbox with Outlook is a time consuming process. You can easily control your mailbox using some simple tricks. You can also use keyboard shortcuts for managing your emails. For example, CTRL+D is used for deleting current mail and CTRL+F is used to forward current mail.

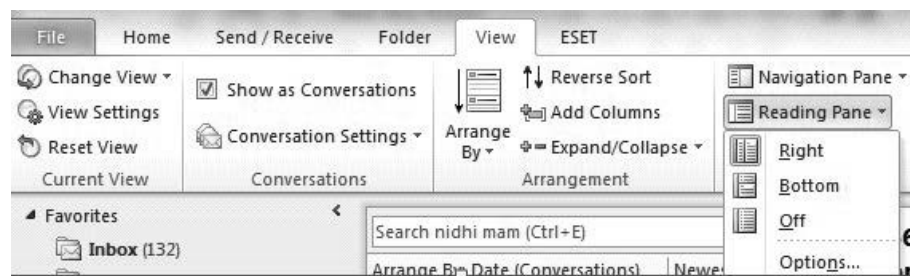
• The Reading Pane and Preview

The reading pane in Outlook 2010 is used for previewing the currently selected email without opening it. The email message text is shown in the reading pane and you can open an attachment or follow any hyperlink using the reading pane.

Steps for displaying the reading Pane in the ribbon are as follows:

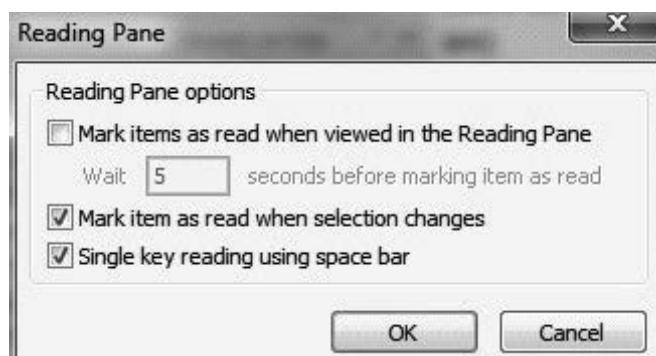
1. Go to the **View** tab and click on the **Reading Pane** under the **Layout** group.

Screenshot displaying various options of **Reading Pane** is as shown:



2. Click on **Options**. When you click on **Options**, **Reading Pane** dialog box will open.

Screenshot displaying **Reading Pane** dialog box is as shown:



3. In the dialog box there are three **Reading Pane** options, enabling you to change the setting of **Reading Pane**.

• Show as conversation

Microsoft Outlook 2010 provides a new feature called conversation view, which allows messages to be viewed in the form of a conversation. The first mail, its reply and forwarded mails are shown as a conversation in a single mail.

Steps for showing emails as a conversation are as follows:

1. Go to **View** tab and click on the **Show as Conversations** checkbox present inside group **Conversations**.

Screenshot displaying **Show as Conversations** options is as shown:



NOTES

2. When you click on the **Show as Conversation** checkbox, a **Microsoft Outlook** dialog box will open.

Screenshot for **Microsoft Outlook** dialog box is as shown:



3. To see the messages inside the current folder in the conversation view click on **This folder**. Alternatively, you can choose **All folders** to see messages of all folders in conversation view. Afterwards, if you don't want the messages to be shown in conversation view, you can remove the conversation view.
4. For removing conversation view, go to the **View** tab and uncheck the checkbox for **Show as Conversations**.
5. When you uncheck the checkbox, the **Microsoft Outlook** dialog box will open.

Screenshot for **Microsoft Outlook** dialog box is as shown:



6. You can click on **All folders** or **This folder** to choose which messages you do not want to see in conversation view.

• Using coloured categories

Microsoft Outlook provides features for Color category. Color category helps in identifying email messages quickly. You can also assign a color category to the related items such as contacts, notes or appointments. This ensures that these email messages can be organized and tracked easily.

NOTES

Steps for grouping email messages by color are as follow:

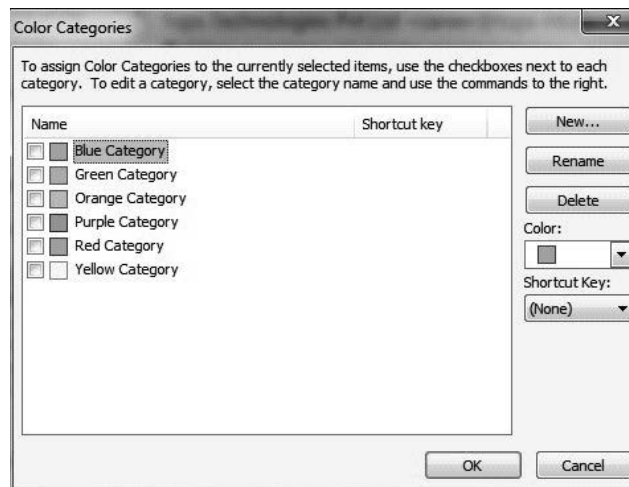
1. Go to **Home** tab present in the ribbon and navigate to the **Tags** group. Click on **Categorize** dropdown box.

Screenshot displaying the **Categorize** dropdown box is as shown:



2. Click on **All Categories** option, the **Color Categories** dialog box will open.

Screenshot displaying **Color Categories** dialog box is as shown:



3. Click on the **New** button, **Add New Category** dialog box will appear. Write name of the category in **Name** box. Choose the color from **Color** dropdown box. If the user wants to assign a shortcut for the category, this can be done with the help of **Shortcut Key** dropdown box.

Screenshot displaying **Add New Category** is as shown:



4. Click **OK** to set the button categories.

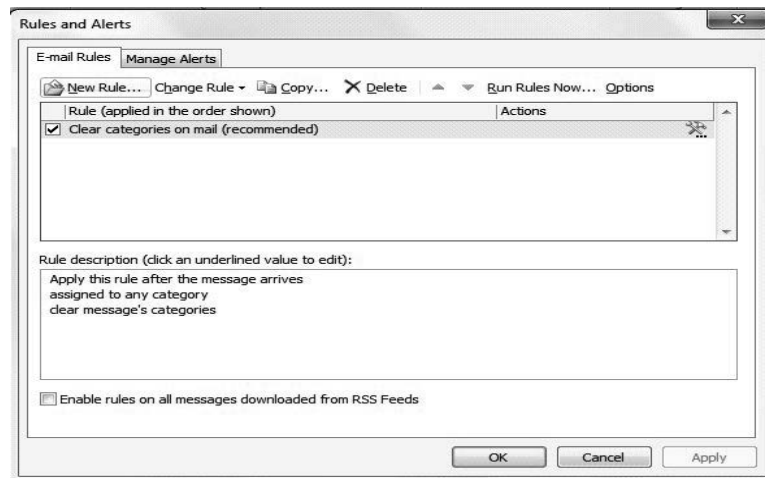
• Rules and Alert

Rules are defined as actions, which Microsoft Outlook automatically performs for incoming or outgoing messages. These actions are performed according to the conditions specified by the user. A rule can be created by taking a message or template into account.

Steps for creating rules are as follow:

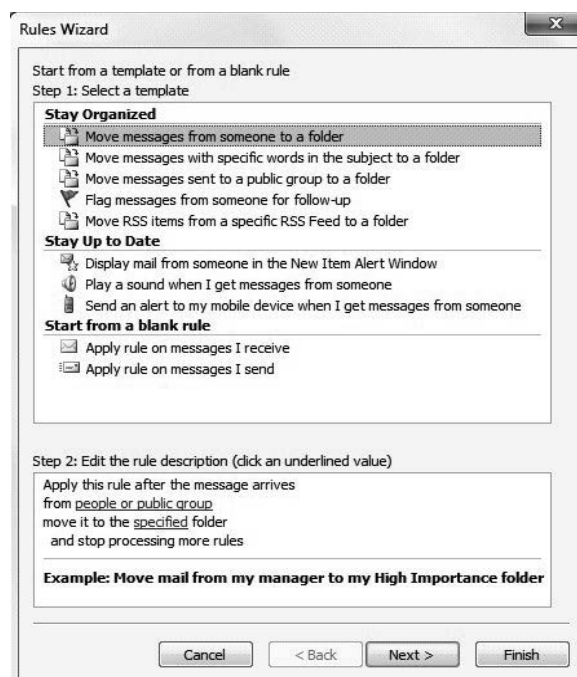
1. Under the **Home** tab, click on **Rules**. In the **Rules** dropdown, click on **Manage Rules and Alerts**. When you click on **Manage Rules and Alert**, the **Rules and Alerts** dialog box will open.

Screenshot for **Rules and Alert** dialog box is as shown:



2. Under **E-mail Rules**, click on **New Rules**. When you click on **New Rule**, **Rules Wizard** will open.

Screenshot for **Rules Wizard** is as shown:

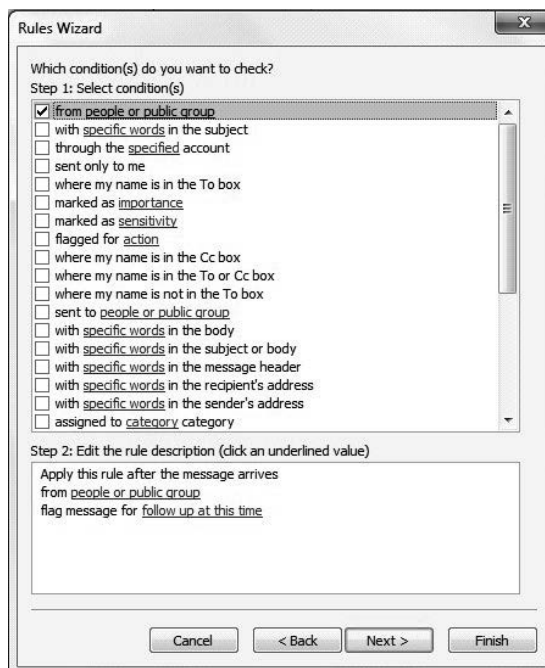


NOTES

NOTES

3. Under the **Rules Wizard**, You can set rules for options—**Stay Organized, Stay Up to Date** and **Start from a blank rule**. From these three options click on one action. Then, click on **Next** button. Another **Rules Wizard** window will open. After that click on the underlined options, to set the value and click on **Next** button.

Screenshot displaying another **Rules Wizard** window is as shown:



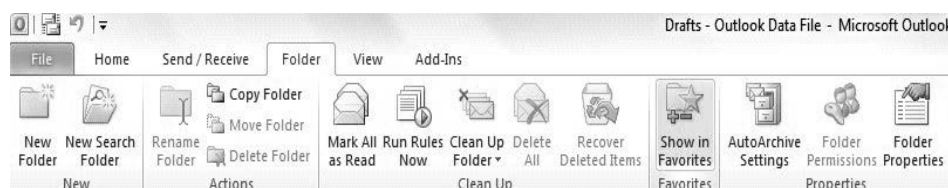
• Favorites in Outlook 2010

The **Favorites** section is present at the top of the Navigation Pane. Shortcuts for inbox, sent items and deleted items are contained in this section. Any folder which you access regularly can be added to this section in order to easily access the mails of that folder.

Steps for adding any folder to **Favorites** are as follows:

1. From the **Navigation Pane**, select the folder which you want to add to the **Favorites** section.
2. Go to **Folder** tab and click on **Show in Favorites** under **Favorites** group.

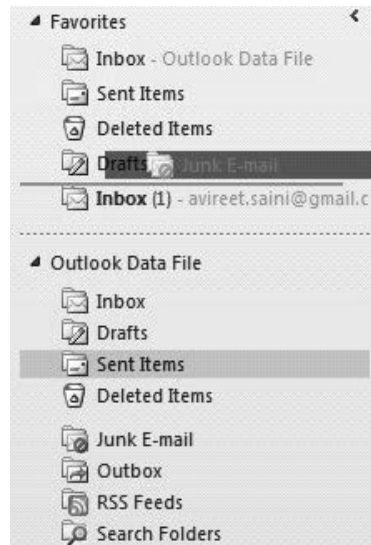
Screenshot for accessing **Show in Favorites** is as shown:



3. The folder you selected will appear in **Favorites**.
4. You can also add a folder to favorites by clicking on any folder and dragging it to **Favorites**.

Screenshot for adding folders in, **Favorites** by dragging is as shown:

Windows, MS Office and
Outlook 2010



NOTES

5. You can also remove any folder by dragging the folder to the Outlook data file.

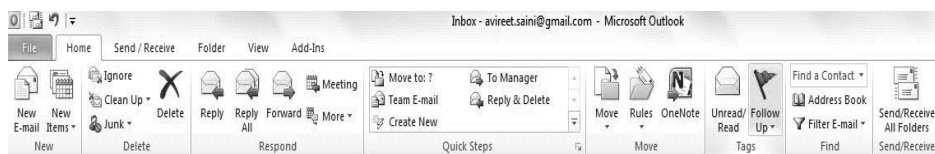
• For Follow Up Folder

Microsoft Outlook supports reminders and flagging of items in folders, other than the default inbox. When you are creating a message, you can add a flag or set a reminder for that mail.

Steps for following up a folder are as follows:

1. Click on any email message which you want to follow up.
2. Under the **Home** tab, click on the **Follow Up** dropdown box present in the **Tags** group.

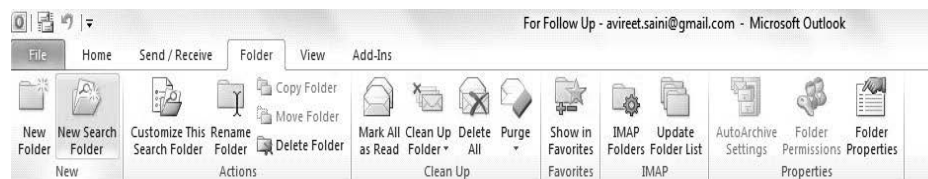
Screenshot for accessing **Follow Up** is as shown:



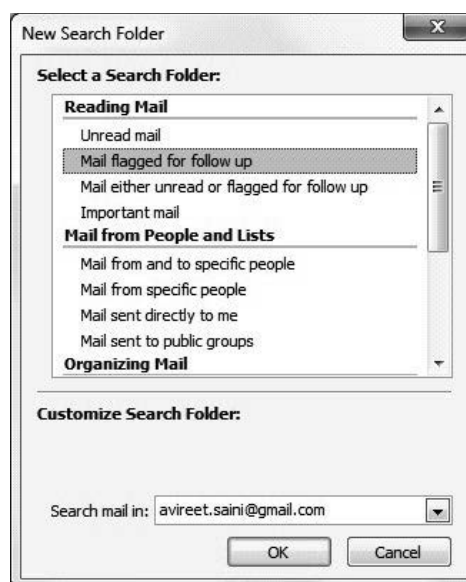
3. When you click on a **Follow Up** option, a flag is added to the selected email message.
4. You can remove it by clicking on **Clear Flag** under **Follow Up**.
5. For following up a folder, go to the **Folder** tab and click on **New Search Folder**.

Screenshot for accessing **New Search Folder** is as shown:

NOTES



6. When you click on **New Search Folder**, **New Search Folder** dialog box will open. Screenshot for **New Search Folder** dialog box is as shown:



7. Click on **Mail flagged for follow up** and click on **OK** button. All the mails that are flagged will be shown to you.

Blocked attachment

Microsoft Outlook helps to protect your computer by blocking certain attachments such as .exe, .asp, .bat, .bas etc. This is necessary because there is a possibility that these file extensions can introduce a virus in your computer. These files are by default blocked in Microsoft Outlook. You can access the blocked attachment by using the following options:

- **Shared server should be used:** The user can ask the sender of the email message to save the file on a shared server or FTP website, from where you can access that file easily. The sender will send the link of the file that he has stored on a shared server and you just have to follow that link in order to access the blocked file.
- **File compression utility should be used:** File compression utility is used to win zip or compress a file and store it with a different extension. This file extension is not recognized by Outlook. Hence this attachment will not be blocked. You can use any file compression utility.

- **Rename the file:** Request the sender of the email message to rename the file with a different file extension that Outlook does not recognize as a threat.

5.7.4 Managing E-mail Accounts

For managing e-mail accounts more efficiently, Outlook provides some additional features such as flags. Flags help in drawing attention to mails that are important. It works more efficiently than folders, rules and categories.

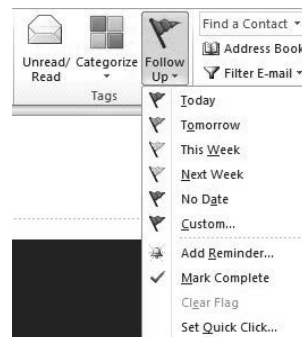
Flags:

Flags in Outlook are used as a reminder. It prompts the user to finish a task until it is finished. Flags help us in responding urgent messages quickly and efficiently.

Steps to **apply a flag** on e-mail are as follows:

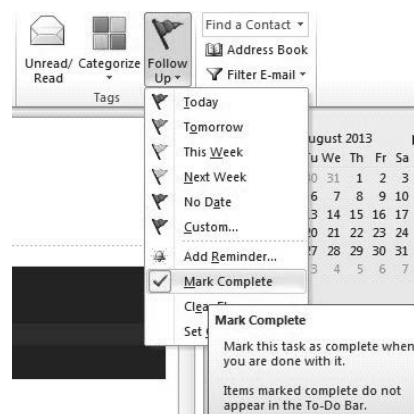
1. Select e-mails that you want to flag and click the **Follow Up** option on the Ribbon.

The screenshot of **Follow Up** option is as shown:



2. Set the corresponding time for the task up to which it has to be completed from the given option of drop-down menu. The selected Flag will be applied to the e-mail.
3. When the task is completed, then click on flag and select the option **Mark Complete**.

The screenshot of **Mark Complete** option is as shown:



NOTES

NOTES

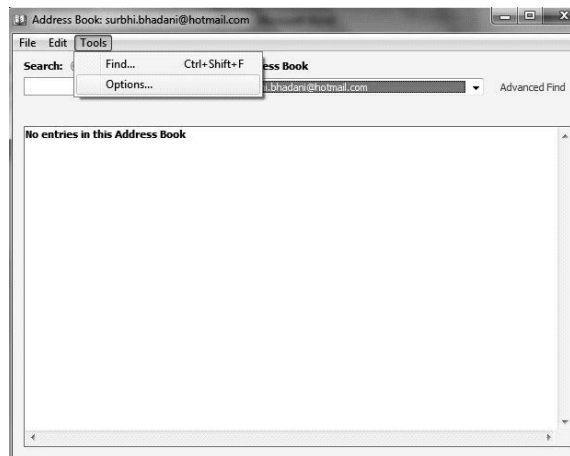
Setting Which Address List Shows First

In Outlook, there are multiple address books, for storing contact information about various departments in the organization. Hence, we need to sort the address book according to the requirements of the organization. The address book is mostly required in order to set the first address list.

Steps needed for deciding which address list should be shown first, are as follows:

1. Go to the **Home** tab present in the ribbon. Next, navigate to the **Find** group and click **Address Book** button.

Screenshot obtained after clicking **Address Book** is as shown:



2. Go to the **Tool** menu and click on the **Options** button.

Screenshot after clicking **Options** button is as shown:



3. Select **Custom** and choose contact **idap.services.wise.edu**.
4. From dropdown box, select the address list which is to be shown first. Here, **idap.services.wise.edu** is selected from the dropdown list.
5. Click **OK** button after selecting.
6. Now, close MS Outlook and reopen it. By default the address book shown first is **idap.services.wise.edu**.

Check Your Progress

8. Define the term Microsoft office access 2010.
9. How are files saved in PowerPoint 2010?
10. State the option to group pictures.
11. How can shapes be added to Power Point 2010?
12. State about the Microsoft outlook 2010.

NOTES

5.8 ANSWERS TO ‘CHECK YOUR PROGRESS’

1. Windows 10 is a personal computer operating system developed and released by Microsoft as part of the Windows NT family of operating systems. It was officially unveiled in September 2014 following a brief demo at Build 2014.
2. .docm file format is used when the document is to be saved in macro-enabled documents.
3. When you click on the Ignore option then MS Word will skip checking the word or phrase in the document without making any change or modification to it.
4. Header and footer are the important parts of any document. They are used for representing information about the document such as the page number, heading of the document etc. Header is present on the top of the document and footer is present at the bottom of the document.
5. File management is the task of maintaining the newly created folders or the existing folders in the computer. It involves various tasks such as saving the folder, maintaining the data and many more.
6. Steps to set border are as follows:
 - Select the cell.
 - Go to Home tab.
 - Click on Borders option in Font group.
 - Select Border style that you want to use.
7. Conditional formatting is a tool of formatting a cell of the worksheet that contains some condition. In the conditional formatting there are different types of rules, bars of option, colour scale, etc.
8. Microsoft office access 2010 or MS Access 2010 is a most popular database management system tool for windows. It combines the object-oriented relational database management system engine with graphical user interface and software-development tools from Microsoft.
9. In PowerPoint 2010, by default, files are saved in .pptx format. In the previous version, files are saved in .ppt format.
10. To group pictures: under Picture Tools, on the Format tab, in the Arrange group, Click, and then click Group.

NOTES

11. Shapes can be added to Microsoft Office PowerPoint 2010, using the Shapes option under the Illustrations group of the Insert Tab. You can insert any shape of your choice and use it to write text.
12. When you click on MS Outlook 2010 for the first time, you will see an e-mail configuration window. A startup wizard guides the process of creating a new e-mail account. You will have to configure an e-mail account before sending and receiving e-mail through Microsoft Outlook 2010.

5.9 SUMMARY

- Windows 10 is a personal computer operating system developed and released by Microsoft as part of the Windows NT family of operating systems.
- The first release of Windows 10 also introduces a virtual desktop system, a window and desktop management feature called Task View, the Microsoft Edge web browser, support for fingerprint and face recognition login, new security feature for enterprise environments, and DirectX 12 and WDDM 2.0 to improve the operating system's graphics capabilities for games.
- To reduce the storage footprint of the operating system, Windows 10 automatically compresses system files. The system can reduce the storage footprint of Windows by approximately 1.5 GB for 32-bit systems and 2.6 GB for 64-bit systems.
- Windows 10 is designed to adapt its user interface based on the type of device being used and available input methods. It offers two separate user interface modes: a user interface optimized for mouse and keyboard, and a "Tablet mode" designed for touchscreens.
- Windows 10 incorporates multi-factor authentication technology based upon standards developed by the FIDO Alliance.
- Microsoft Family Safety is replaced by Microsoft Family, a parental controls system that applies across Windows platforms and Microsoft online services.
- Home, pro, enterprise and education are the four main editions for personal computer that is available with Windows 10.
- In MS Word 2010, files are saved in the default format, which is .docx. This is because .docx format is more secured and damages can be easily recovered.
- The 'Print Preview' feature of MS Word helps you to view on the screen how the printed version of the document would look like before printing a hard copy. In latest versions of the MS Word (MS Word 2010 Onwards), there is 'NO' Print Preview Tab or Option.
- MS Word provides significant feature 'Proofing' for checking the text in the document by using the option 'Spelling & Grammar'. If you have made mistakes in the document while typing the text then you can use the numerous proofing features provided by MS Word to produce error-free and professional documents.

- In Microsoft Excel 2010, various new features have been included which help in creating an effective and attractive worksheet.
- Microsoft Excel is used for storing the data in the form of table. It is used by several organizations to perform complex calculations, statistical analysis and tracking income and expenses.
- Excel sheet contains data in cells. These cells can be converted into table using Format as Table command. There are many built-in table styles available in Excel 2010. When you apply table style on cells Design tab appears.
- Excel file is a workbook that contains one or more worksheets. Formatting Excel sheet improves the presentation of the worksheet and makes it easy to use. By default, each Excel file has three worksheets.
- Microsoft office access 2010 or MS Access 2010 is a most popular database management system tool for windows. It combines the object-oriented relational database management system engine with graphical user interface and software-development tools from Microsoft.
- The Database Management System or DBMS is the software which acts as an interface between the underlying database and the user to define, update, retrieve and administer the database.
- In PowerPoint 2010, by default, files are saved in .pptx format. In the previous version, files are saved in .ppt format.
- Animation is one of the most common features of MS PowerPoint. It is used to make an attractive presentation and for creating the interest in audience. MS PowerPoint 2010 provides different varieties of animations that can be applied to text, picture or other graphics in multiple ways.
- Microsoft Office PowerPoint 2010 allows its users to easily use audio-visual content in the presentation. It supports a large variety of video files such as, Windows Media File, MP4 Video, MPEG-2 TS Video and Window Media Video File, along with audio files such as ADTS Audio, AIFF audio file (aiff), AU audio file, MP3 and MP4 audio files.
- MS Outlook provides mailing tool which helps in sending or receiving mails. MS Outlook allows access to e-mail account on your computer even without internet, that is, you can view and manage your e-mails when you are offline.

NOTES

5.10 KEY TERMS

- **Cortana:** It is an intelligent personal assistant created by Microsoft.
- **Universal Windows Platform:** It provides a common app platform available on every device that runs Windows 10.
- **Ribbon:** It is a set of tabs and commands to perform various sets of functions and operations on documents.

NOTES

- **Formatting:** It means changing the property or the style according to the need of the user.
- **Header:** Header is added on the top of the document which contains the important information of the document such as title, page number and other information.
- **Validation:** This is the process of verifying the data to determine if the data is of correct format and in the desired range.
- **Chart:** A chart in Excel refer to the tool that helps in representing data in the graphical form. It increases the understanding of the audience or a user.
- **Database:** The term used to describe a collection of related 'data' (information) stored on computers.**Database:** The term used to describe a collection of related 'data' (information) stored on computers
- **Queries:** The commands that are used to retrieve and update the data based on the conditions depending upon the business requirements.**Queries:** The commands that are used to retrieve and update the data based on the conditions depending upon the business requirements
- **Reports:** These are designed to create an organized output of data from your database.
- **Review Tab:** It contains Proofing, Language, Comments, and Compare options. It is utilized for spelling and grammar editing.
- **Slide Transitions:** Slide transitions display slides in such a way that one slide disappears and the next one appears in an elegant way.
- **MS Outlook:** It is one of the tools of MS Office which is used to store contacts, send e-mail, schedule meetings, create appointments etc.

5.11 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. When was the Windows 10 released?
2. What are the names of the notes that are added at the end of the page and at the end of the document?
3. Write the steps involved in adding the break in the document.
4. Write the steps to add column to the table.
5. Which command does Excel use for replacing data?
6. Why is the text function used?
7. How MS Access 2010 is useful for business?
8. Write the significance of AutoForm in MS Access 2010.
9. How are tables useful in PowerPoint 2010 for organizing and presenting data?
10. State about the E-mail accounts for Microsoft outlook 2010.

Long-Answer Questions

1. Briefly explain the Window 2010 and its features with the help of examples.
2. Discuss briefly header and footer and write the steps involved in adding header and footer to the document and also explain how to remove the header and footer from the document.
3. Describe briefly formatting and how is it done in MS Word 2010?
4. Discuss the differences between MS Word 2010 and 2016.
5. Briefly explain the Microsoft excel 2010 with the help of appropriate examples.
6. Explain the various ways to create a table in MS Access 2010.
7. How the query is designed using design view and wizard? Discuss the steps involved with the help of screenshots.
8. What is the need of designing reports? Describe the various ways for designing reports in MS Access 2010.
9. Explain the Microsoft PowerPoint as the most popular program for creating the interactive presentations.
10. How can you insert charts in Microsoft PowerPoint 2010? Write the steps for inserting data in charts.
11. Explain the new features added to PowerPoint 2016.
12. Discuss about the MS Outlook 2010. Also, explain the use of MS Outlook 2010.

NOTES

5.12 FURTHER READING

Shelly, Gary B. and Misty E. Vermaat. *Microsoft Office 2010: Introductory*, 1st edition. United States: Cengage Learning 2010. Shelly, Gary B. and Misty E. Vermaat. 2010. *Microsoft Office 2010: Introductory*, 1st edition. United States: Cengage Learning.

Habraken, Joe. *Microsoft Office 2010 In Depth*, 1st edition. New Jersey: Que Publishing 2010. Habraken, Joe. 2010. *Microsoft Office 2010 In Depth*, 1st edition. New Jersey: Que Publishing

Bhatt, Pramod Chandra P. *An Introduction to Operating Systems—Concepts and Practice*. New Delhi: PHI 2003.

Hamacher, V.C., Vranesic, Z.G. and Zaky. S.G. *Computer Organization*, 5th edition. New York: McGraw-Hill International Edition 2002.

Nutt, Gary. *Operating Systems*. New Delhi: Pearson Education 2006.



NOTES

NOTES