B.C.A. – I YEAR

DJA1A : INTRODUCTION TO COMPUTERS

SYLLABUS

Unit I: Introduction to Computers

Introduction to Computers: Introduction – Characteristics of computers- Evolution of computers – Generation of computers – Based on purpose, Based on type of data handling techniques, and according to Functionality – The computer system – Application of computers.

Unit II : Input Devices

Input Devices: Keyboard – Pointing Devices – webcam – Scanners – Optical character Recognition – Magnetic Ink Character Recognition – Bar Code Reader. Output devices: Printers – Plotters – Computer Output Microfilm – Monitors – Voice recognition System – Projectors.

Unit III : Primary Memory

Primary Memory: Memory Representation – Memory Hierarchy – Random Access Memory – Read only Memory – Types of ROM.

Secondary Storage: Classification of secondary storage devices – Storage organization of Magnetic Disk – Storage organization of optical Disk – Magneto Optical Disk – Universal Serial Bus.

Unit IV : Database Fundamental

Database Fundamental: Data, Information and knowledge – Database – Logical Data Concepts – Physical Data Concepts – Database Management System – Need, Benefits of DBMS, Components of DBMS, Database Administrator – DBMS Architecture – Database Models.

Unit V : Internet Basics

Internet Basics: Basic Internet Terms – Internet Addressing – Internet Applications – e-mail, WWW, File Transfer Protocol, Telnet, Internet Relay Chat, Gopher, Chatting, Commerce through Internet , Groups, News, Social Networking, Blog, Videoconference, Online services – E-mail address structure – Sending and Receiving E-mail – Search Engines – Internet and Viruses.

UNIT I: INTRODUCTION TO COMPUTERS

Introduction to Computers: Introduction – Characteristics of computers-Evolution of computers – Generation of computers – Based on purpose, Based on type of data handling techniques, and according to Functionality – The computer system – Application of computers.

1.1 Introduction

We are living in the computer age. Computers can no longer be termed as mere invention, it is a revolution. Computer revolution has found its way into, almost every aspect of human life and living. In India, computers are now being increasingly used and their impact on social, economical, educational and scientific systems is being greatly felt and recognized.

Literally computer is a calculating machine. We know that a computer is a fast calculator, but in reality it is much more than that. A computer is an electronic data processing machine, which receives information, performs basic operation on that information and produces the output or result according to our pre-determined conditions specified in the program.

1.1.1 Computerization and Information Technology (IT)

Like the industrial revolution which changed human society, computer revolution is bringing dramatic shifts in the way we live, perhaps even in the way we think. The computerization has a great impact on the society, people and the organizations. Gaining computer literacy has become top priority of students. Organizations are putting the computers to works, though the geographical world is large, the information technology world has shrunken considerably.

Computerization is essential to survive in the fast world. It is more necessary to introduce automation especially in industries and big offices due to advancement in rapid industrialization and the need of faster communication in

the business dealings. Computerization is changing people's life in areas as diverse as medicine, education, publishing, business, home and transportation.

Computers have completely altered the structure of the business. Large volumes of accounting, and record keeping, data can be manipulated, organized, stored, retrieved and used for specific purposes. Financial projections are made with greater ease.

Word processing saves time for people at all levels of the organization and helps ensure accurate letters, reports and memos. Automated filing uses far less storage space than endless stacks of paper and enables workers to retrieve documents rapidly when they are needed.

Computerizing all phases of activities in transportation, industries and other services should be accelerated to increase productivity, availability to focus on economy, cost effectiveness, security and safety. In short, computerization simply means that the manual work replacement of computer by using its special features. But this alone is not enough to survive in the fast developing and competitive world. To catch up the rest of the world by utilizing the features of computers, the name information technology came into the ground. The achievement of Information Technology (IT) includes.

-Internet (World Wide computer network)

-Intranet (country wide/Zone wide computer network)

-e-mail (sending/receiving messages using internet)

-e-commerce (financial/ bank dealing using internet)

-e-business (business using internet) etc,.

The effects of IT are

- Possible to handle various types of data such as text, audio and video.

- Functions of various means of transmitting data such as telephone, telex, television and computers are integrated.

1.1.2 Computer Related Terminologies

Program: A set of instructions that a computer executes to perform a series of actions or a particular type of work.

Algorithm: A finite sequence of steps for solving a problem or performing a task.

Flowchart: A graphical representation of the path of control of the operations in a program. Symbols such as squares, diamonds and ovals represent various operations. These symbols are connected by lines and arrows to indicate the flow of data or control from one point to another. Flowchart is an aid to show the way a proposed program will work and as a means of understanding the operation of a program.

Machine level language: Program written in terms of only 0 and 1. Machine language is directly executable by a computer. But it is machine dependent that is different computer need different program.

Assembly level Language: Program written in terms of short codes namely mnemonics. To execute this program a translator called assembler is needed to covert from assembly language to machine language. This type of language is also machine dependent. The machine and assembly language are called low level languages.

*High level lang*uag*e* – Program written in languages such as C, C++, java. To execute this program a translator called compiler or interpreter is needed to covert from high level language to machine language.

Source code: Human readable program statements written by a programmer or developer in a high-level or assembly language that are not directly readable by a computer. Source code need to be compiled into object code before it can be executed by a computer.

Debugging: To detect, locate and correct logical or syntactical errors in a program.

Programmer: An individual who writes and debugs computer programs.

End user: The ultimate user of a computer or computer application in its finished form.

1.2 Characteristics of Computer

The following are the characteristics of the computer

Speed   Accuracy   Storage capacity

Diligence   Automation   Consistency

Speed

The speed of a computer is closely related to the amount of data it must process. To classify the speed of different computer systems, the industry has developed the criteria of Million Instructions Per Second (MIPS).

Even the operations of personal computers are measured in thousands of a second or millisecond (10^{-3}), operations for larger processors are measured in microseconds (10^{-6}), nanoseconds (10^{-9}) and picoseconds (10^{-12}) etc.

Accuracy

Accuracy refers to the degree of correctness and exactness of operations performed by a computer. The computer can process data accurately and quickly. Accuracy is a prime consideration in inventing computers. Computers never make errors. The errors attributed to computers are really human errors. The errors may be due to wrongly coded program (code - instructions to computers) or due to mistakes in input data entry.

Storage capacity

Computer systems have total and instant recall of data and almost unlimited capacity to store these data. We can store billions of instructions and perhaps more than thousands of graphic images which are available for instant recall. When properly used, a computer can improve the efficiency of an organization.

Diligence

Computer systems are non-respondent to human factors like fatigue, tiredness or boredom. Computers can work without getting tired for many hours. Even though it has no intelligence, it executes the fed instructions correctly and with accuracy.

Automation

The set of instructions are fed into the computer system along with necessary data. Based on these, the computers automatically perform the instructions continuously without manual intervention, or supervision.

It is more essential to introduce automation especially in industries and big offices due to advancement in rapid industrialization and need for quick interactions in business operations. Computers can do all sorts of automation, if they are programmed suitably and correctly.

Consistency

The capability of adopting all kinds of tasks is termed as consistency. It can perform variety of jobs, ranging from simple arithmetic calculations to complex graphics modeling, animation, depending on the application packages loaded into it. Hence, computer can act exactly like an obedient assistant which can perform tasks repeatedly at high speed with accuracy.

Limitations

The computer can outperform human beings in speed, memory and accuracy but still the computer has limitations. They are

Lack of commonsense - The computer is only a tool. It cannot think. It does not have commonsense and intelligence as we all have. For example, suppose we give the respective ages of two girls as -12 and 10 and ask a boy to find out who is younger? The boy will immediately ask how age can be negative.

But if we give the same data -12 and 10 to a computer it will compare the two numbers and give the answer "The younger girl's age is -12." Why? Because the computer does not have the commonsense to understand that, the age cannot be less than 0.

Inability to correct - When we give instructions to a computer we must give the correct instructions. A computer cannot correct wrong instructions. This follows from the fact that, we cannot replace the human brain by using computers.

Dependency on human instructions - A computer cannot generate any information on its own. It can only do what it is told to do.

Even though the above mentioned limitations are there, research is going on in the field of soft computing, to impart knowledge so that it can think and act independently.

1.3 Evolution of Computer

The computer was born not for entertainment or email but for some computational purpose. Today, we carry more computing power on our smart phones than was available in the early models. The following brief history of computing is a timeline of how computers evolved from their humble beginnings to the machines of today that surf the Internet, play games and stream multimedia in addition to crunching numbers.

Abacus – Ancient counting machine. It is a frame with vertical rods, a horizontal crossbar and beads. It is based on the decimal system: the first vertical bar is for units, the second for tens and so on. The device was first used in china around the year 1300.

1614 – *Napier Rods* – yet another milestone in the quest for manipulating numbers. John Napier invented a system of movable rods, referred to as

Napier rods, which allowed for multiplication, division and square, cube root calculation.

1625 – **Slide Rule** – can to do many different calculations.

1786 – J.H.Muller, an army officer regards the **idea of the Difference Engine**. A difference engine is a special purpose mechanical digital calculator designed to tabulate polynomial functions, logarithmic and trigonometric functions.

1822 – Charles Babbage designed his first mechanical computer, a prototype for his difference engine. Babbage invented two machines the **Analytical engine** and the **Difference engine**. Both were too complex to actually be built. The analytical engine, which Babbage outlined in 1833, used *punched cards* for input.

1842-43 – Ada Lovelace wrote to Babbage about a plan for how the difference engine might calculate Bernoulli numbers. This plan now regarded as the **first computer program.**

1848 – British mathematician George Boole devised binary algebra, soon called **Boolean algebra**. This was essential for a binary computer to be developed.

1896- *Herman Hollerith* founded the tabulating machine company, which went on in 1924, to merge with another company and became *International Business Machine (IBM)*. IBM is one of the most dominant corporations of the computer age and secured his place in history as the *father of information processing*.

1906- Lee de Forest, invented three-element **vacuum tube**.

1937 – Alan Turing invented *Turing Machine*, a theoretical simplified computer.

1941- Atanasoff and Berry completed a special purpose calculator for solving simultaneous linear equations; this was later called *ABC*. Dr. Zuse's *Z3 computer* was designed and was the first automatic, Program-controlled, fully-functional, general purpose digital computer.

1946 – *ENIAC* (Electronic Numerical Integrator And Computer) was one of the first totally electronic, valve-driven and digital computers, designed by Prosper Eckert and John Mauchly.

1947 – Invention of **transistor.**

1948 – **SSEM** (small Scale Experimental Machine) was built based on idea from John Von Neumann, about stored program computers. This was the first computer ever to store both programs and data in memory.

Von Neumann is known for the computer architecture which bears his name – a control unit, an arithmetic and logical unit, memory and input/output. The *Von Neumann architecture* is at the heart of the design of all modern day computers from handhelds to super computer. An alternative architecture known as the *Harvard architecture* where the instructions and data are separately stored is used in some components like digital signal processors used in audio or video signal processing.

1948 – Claude Shannon identified the *bit* as the fundamental unit of data. He also identified it as the basic unit of computation.

1949 – *EDVAC* (Electronic Discrete Variable Computer) was the first to use *magnetic tape* for storage.

1951 – The *UNIVAC I* (UNIVersal Automatic Computer), the first commercial multipurpose computer was delivered.

1953 – A small computer was built using *transistors*, the first time vacuum tubes were not used in a computer.

1954 – The *NORC* (Naval Ordnance Research Calculator) was delivered to the US Navy. It was the first computer to have been called a *super computer*. Magnetic disk storage, high level language FORTRAN and COBOL was developed.

1960-69 – The decade of *networking*. Time sharing concept, packet switching comes in. Object oriented programming, the first modem are notable development in this period.

1970-79 – The innovations, inventions and developments continued in both **hardware and software**. The hardware side saw the birth of the microprocessor, a variety of personal computers, commercial super computers, the hard disk and more. Computers got smaller. The software side saw developments in programming languages – C, SQL and many others. This decade also saw the first generation video game console.

1980-89 – This decade dawn many things that directly constitute the computing experience of today, such as **GUI (Graphical User Interface), the PC, the Mac system.**

1990-99 – Mostly the **Internet** – The World Wide Web, Linux operating system, plug and play, Pentium processor, Yahoo directory, Java, DirectX, Dialup, hotmail, Google etc .

2000 – Till – The **new millennium** – We are enjoying innovation in dotcom bubble, variety of operating system, handheld devices, web browsers,
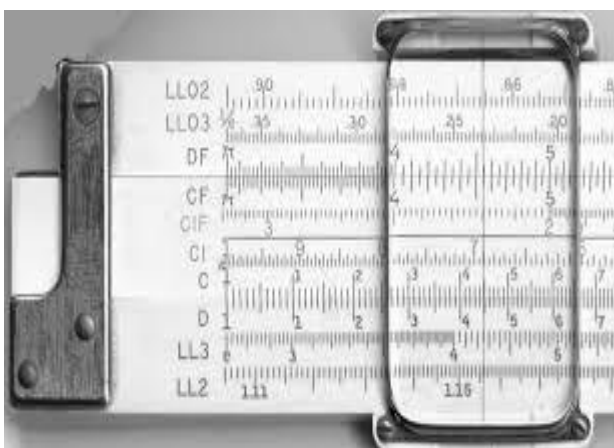
social media, computing techniques3 etc. Figure 1.1 shows the developments in computing environment.
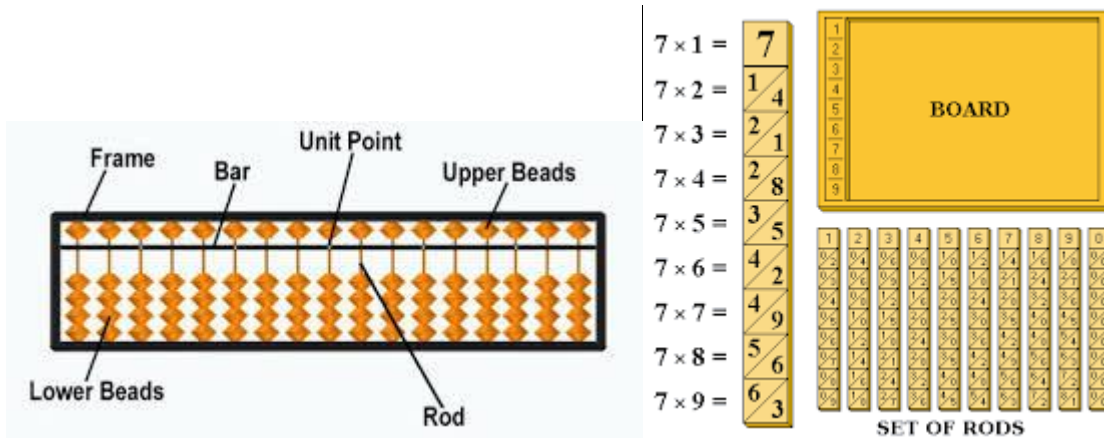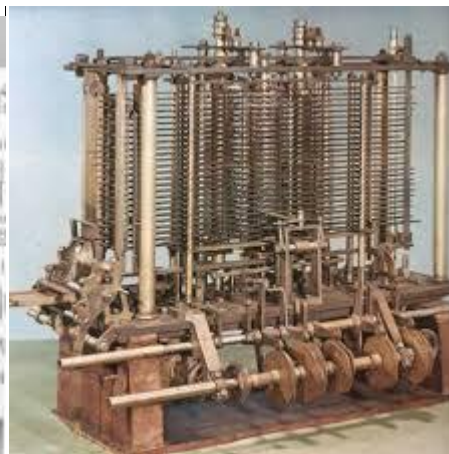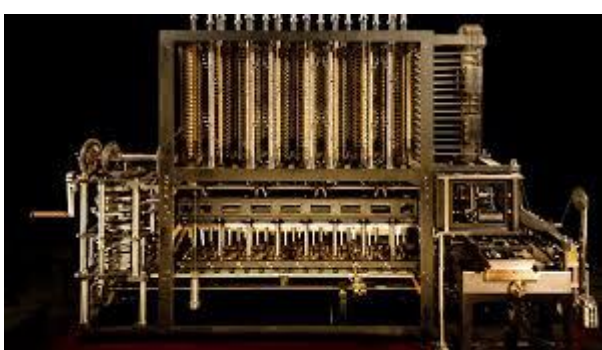
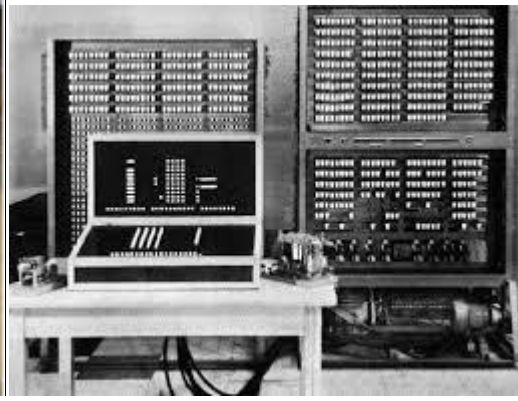

Abacus

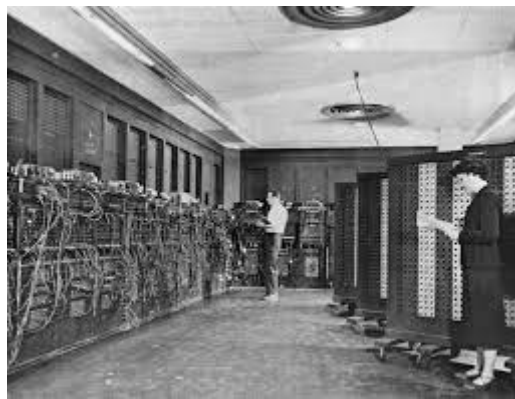Napier Rods

Slide Rule

Analytical Engine

Difference Engine

Turing Machine

ABC Computer   Z3 Computer

ENIAC

Figure 1.1 Developments in computing environment

1.4 Generations of Computer

Computers are classified into five generations based on the size, memory capacity, speed, accuracy, electronic technology used in constructing the computer, the associated system software and applications.

First generation computer

-The period of first generation: 1946-1959

- Built with valves i.e vacuum tubes which were used in old radios and TVs.

Limitations of first generation are: Unreliable, Supported machine   language only (program written in terms of 0's & 1's), Very costly, Generated lot of heat,

Slow input and output devices, Huge size, Need of A.C., Non-portable, Consumed lot of electricity

Example: ENIAC, EDVAC, UNIVAC, IBM-701, IBM-650

Second generation computer

- The period of second generation: 1959-1965

- Built with solid state transistors and diodes.

Features of second generation are: compare to first generation computers, these are reliable, smaller in size, generated less heat, consumed less electricity, improved speed but costly and need of AC.

Supported machine and assembly (program written in terms of short codes namely mnemonics) languages.

Example: IBM 1400 series and IBM 7000 series, CDC 3600.

Third generation computer

- The period of third generation: 1965-1971

- Used integrated circuits (IC's) in place of transistors. A single IC has many transistors, resistors and capacitors along with the associated circuitry. This development made computers smaller in size, reliable and efficient.

- Compare to previous two generations these are more reliable, smaller in size, generated less heat, faster, required lesser maintenance

- Supported high-level languages like FORTRAN-II TO IV, COBOL, PASCAL PL/1, BASIC, and ALGOL-68.

- In this generation remote processing, time-sharing, multi-programming operating system were used.

Example IBM 360, Honey well 200, PDP (Personal Data Processor)

Fourth generation computer

- The period of fourth generation: 1971-1980.

- Used Very Large Scale Integrated (VLSI) circuits. VLSI circuits having about 5000 transistors and other circuit elements and their associated circuits on a single chip made it possible to have microcomputers of fourth generation.

- More powerful, compact, reliable, and affordable. As a result, it gave rise to personal computer (PC) revolution.

- Concept of internet was introduced; Great developments in the fields of networks, Computers became easily available

- In this generation time sharing, real time, networks, distributed operating system were used.

- All the high-level languages like C, C++, DBASE etc., were used. -Example IBM 370, Honey well 6000 series, PDP 11, CRAY-1(Super Computer), CRAY-X-MP (Super Computer)

Fifth generation computers

- The period of fifth generation: 1980-till date.

-In the fifth generation, the VLSI technology became ULSI (Ultra Large Scale Integration) technology, resulting in the production of microprocessor chips having ten million electronic components.

- This generation is based on parallel processing hardware and AI (Artificial Intelligence) software. All the high-level languages like C and C++, Java, .Net etc., belongs to this generation.

- Main features are: More user friendly interfaces with multimedia, Availability of very powerful and compact computers at cheaper rates

- Some computer types of this generation are: Desktop, Laptop, Notebook, Ultra Book and Chrome Book. Figure 1.2 shows the components used in five generations.
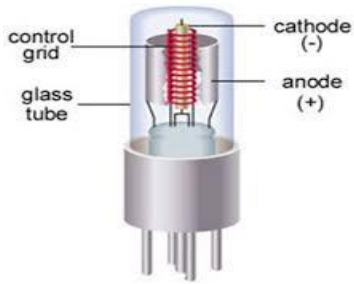
Vacuum Tubes
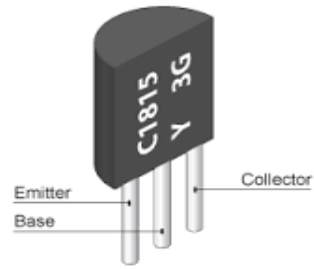
Transistors
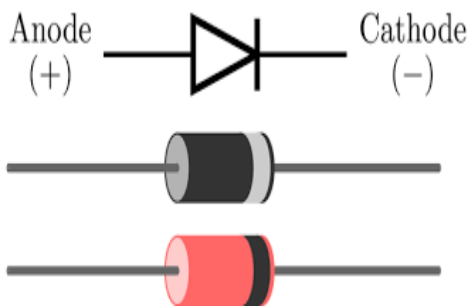
Diodes

Integrated circuits

LSI circuits

VLSI circuits

Figure 1.2 Components used in five generations of computers

1.5 Classification of Computers

1.5.1 Based on Purpose

According to the field of application the digital computers can be divided into two categories namely, *Special purpose computers and General purpose computers*.

A special purpose computer is designed to perform one specific task. Example: Missile guidance computer.

A general purpose computer is one that has the ability to perform various operations. Generally, these are capable of carrying out some general data processing under program control.

1.5.2 Based on type of data handling techniques

Analog Computer

Analog computers work on continuous physical magnitude like length, pressure, temperature, voltage, current, height, mass, radius etc., which are analogous to the numbers under consideration.

There are two categories of analog computers: Direct (Special purpose) and Indirect (General purpose) computers. Direct analog computers are also known as analog devices. A number of instruments are available that work on the principles of an analog direct computer.

For example, the Thermometer, though it is not doing any calculation, it is in a position to precisely measure the temperature of the body by comparing the relative expansion of mercury. The Speedometer is in a position to tell the speed at which the vehicle is running, by comparing the rotational motion of the shaft of the vehicle. Analog indirect computer operate on analog data by performing physical processes on these data.

Digital Computer

A digital computer operates by counting numbers. Digital computers produce precise results with more accuracy. It can accurately represent data using as many positions and numbers as necessary.

Adding machines and pocket calculators are common examples of devices constructed on the principles of the digital computer.

Hybrid computer

Hybrid computers combine the best features of analog and digital computers. It is a combination of high speed of electronic analog computers and the accurate-modest size scientific digital computer, linked together by a communication interface system.

They are usually used for special problems in which input data derived from measurements, is converted into digits and processed by computer. For example a hybrid computer controls national defense, fire control system and passenger flight radar.

The differences between the analog computers and the digital computers are summarized in the *Table 1.1*.

Table 1.1 Analog Versus Digital

Analog system	Digital system
Set up analogy of problem	Breaks down problem into arithmetic operations.
Represents physical variable as continuous measurement of analogous quantity Eg. Presence of pulses, position, voltage, rotation.	Represents numbers as discrete coded pattern. Eg. Any digital data
Operations performed by few single purpose devices.(integrators,	Operations performed by relatively many interchangeable arithmetic

multipliers, adders)	devices (registers, accumulators)
Relatively few devices needed. Low cost and ease of programming.	Many devices needed. High cost and critical programming.
Distinct elements used for operation. (Parallel channels).	Identical elements used in each sequence.(Primary series operation).
Limited accuracy	Unlimited accuracy.
Data storage distributed in various non-interchangeable devices.	Data storage concentrated in space, interchangeable and unlimited in duration
Analog computer serves as model and mirror relations of actual system. Operations usually carried out in actual time of physical system.	Digital computer compounds arithmetic data unrelated to system it represents. Time of operations usually does not correspond to real time.
Represents physical or mathematical quantities.	Represents numbers as well as letters and other symbols.

1.5.3 According to Functionality

Processing capacity of a computer refers to the volume of data that a computer system can process. Formerly a computer's size was a sign of its capacity. With the current state of smallness, dimension of capacity is based on throughput of the computer. Throughput is the quantity of processing that can be performed in a given amount of time. Based on throughput, computer systems can be divided into the following categories:

Micro computers (Personal Computers)

The brain of a microcomputer is the microprocessor; it is a silicon chip containing essential circuits to execute logic or arithmetic operations and to manage the input/output operations. A microprocessor is an integrated circuit

which usually contains millions of transistors squeezed onto a small silicon chip.

When the microcomputers were invented it had very limited processing power and limited choice of input/output devices. Now, they have wider processing capabilities and maintain a wide range of input/output devices.

The older PC had 8 bit processor with speed of 3.7MB and current PC has 64 bit processor with speed in terms of GB. Examples: - IBM PCs, Apple computers

Microcomputers are generally known as personal computers. Desktop computers are the most common type of PC. Notebook (laptop) computers are used by people who need the power of a desktop system, but with portability. Handheld PCs such as PDA, lack the power of a desktop or notebook PC, but offer features for users who need limited functions and small size. The difference between desktops and portables   is, portables can be used while travelling whereas desktops computers cannot be carried around.

The different portable computers are Laptop, Notebooks, Palmtop (hand held), Mobile slate and Wearable computers

Laptop - This computer is similar to a desktop computers but the size is smaller. They are expensive than desktop. The weight of laptop is around 3 to 5 kg.

Notebook - These computers are as powerful as desktop but size of these computers are comparatively smaller than laptop and desktop. They weigh 2 to 3 kg. They are more costly than laptop.

Palmtop (Hand held) - They are also called as PDA. These computers are small in size. They can be held in hands. It is capable of doing word processing, spreadsheets and hand writing recognition, game playing, faxing and paging. These computers are not as powerful as desktop computers.

Mobile slate – A portable device that is in between laptop and PDA is mobile slate or PC slate. This is exactly a slate-shaped PC. It weighs around a kg and is about the size of an A4 sheet.

The PC slate comes with a stylus and a touch enabled screen. The information written on the slate is converted as text. A stylus is used to touch the letters, we need to input on a software keyboard.

Wearable Computers – Wearable PCs are like walkman style. In the wearable hands-free PC, voice control acts as an input, miniature display and head mounted display acts as an output. In this model, applications are integrated into watches and even clothing. (Figure 1.3)

This is the most appropriate instrument for workers and technicians working in difficult situations, whose hands must be free for other work. For example airline technicians and mechanics crawling inside the airline body attending to repairs, physicians who examine patients etc.

Figure 1.3 Wearable Computers

Workstation or Network Computer

Even though microcomputers are widely used, they have some limitations in terms of their available memory, disk space and graphics facilities. Further, their processing speed is also limited.

PCs are made more powerful with improved performance and are called as work stations which are generally used by a single user. They symbolize the bridge between the microcomputers and minicomputers. It is a microcomputer with many of the facilities and abilities. At first it was designed for use by designers and engineers who were in need of extremely powerful processing and output capabilities. Based on their capability, it can be used as a server/ client/terminal.

A computer network is a collection of computers connected together. *Servers* are not designed to be used directly. They make programs and data available for users having access to a computer network.

To use servers, users run desktop programs called *clients*, which contacts the server and retrieve information from the server. Use of desktop clients and centralized servers is called client/server computing.

Although *terminals* look like the personal computers, they have some limitations when compared with PCs. Terminals have only a screen and a keyboard and the electronics that allow them to communicate with the computer to which they are connected. Because they lack the ability to process data on their own, they are called *dumb terminals*.

Mini computers

A minicomputer is a medium-sized computer. That is more powerful than a microcomputer. These computers are usually designed to serve multiple users simultaneously (Parallel Processing). They are more expensive than microcomputers.

Unlike workstations these computers can be used by several users in a time-shared mode. These can also be used as a web server to process the requests and send the requested document through internet.

Main frame computers

Main frame computers offer extensive benefits over mini and microcomputers. Some of these have greater storage facility and processing speed, a larger assortment of input/output devices with multiprogramming and timesharing capability.

The largest mainframes can handle thousands of dumb and intelligent terminals and use more than terabytes of secondary storage. An intelligent terminal has its own processor and do not usually provide any storage. Mainframes are used in large organizations where many users need access to shared data and programs. For example these are used to access the bank account from an automated teller machine. The bank's mainframe handles all the transactions. Complex scientific and engineering problems can also be processed using mainframes.

Super computers

These are sophisticated, expensive and powerful computers to execute complex calculations at the maximum speed. Although some supercomputers are single computer systems, most are comprised of multiple high performance computers working in parallel as a single system. Supercomputers attain their high speed by using special techniques like pipelining and parallel processing.

In pipelining, a primary task is subdivided into smaller subtasks. Each subtask is carried out by the ALU enabling high speed of processing.

Another method to increase the speed of computing is by having multiple processing units. Part of any calculation can be divided up and done in parallel. This method is called *parallel processing*.

Supercomputers are extensively being put to real-time processing environment. Some areas of applications are military applications, aircraft control, modeling complex & dynamic system such as world's weather, U.S

economy, motions of galaxies, spiral arms, map the human genome, DNA structure etc.

The best known supercomputers are built by Cray supercomputers. CRAY XMP-14 is the first Indian supercomputer. Figure 1.4 shows mainframe and super computer models.

**Figure 1.4 Mainframe and Super computer models**

1.6 The Computer System

A computer consists basically of three major units namely input unit, central processing unit and output unit. (*Figure 1.5*)

Input unit

The function of the input unit is to read the information contained in the program and transmit to the central processing unit. The media used for feeding any information may be of different forms, depending on the facilities available and the requirement of the situations.

Example of Input devices include Keyboard, Mouse, Magnetic Ink Character Reader (MICR), Punched card reader, Paper tape reader, Magnetic tape reader, Tele typewriter, Light pen, etc.,
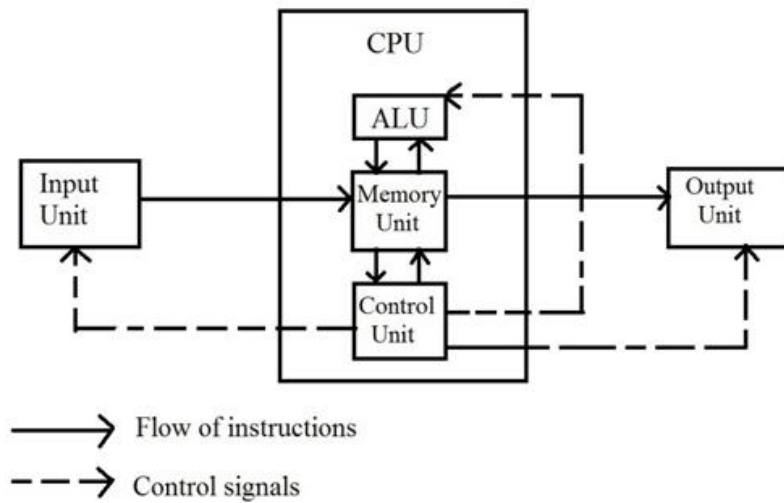
Figure 1.5 Block diagram of a computer

Central Processing Unit (CPU)

CPU is the heart of the computer system since the actual computation work is carried out in this. It links the input and output units together. Internally it consists of a Memory unit, an Arithmetic Logic Unit (ALU) and Control unit. Memory is the section where the instructions and data can be stored. All the intermediate and final results are also stored in the memory.

Arithmetic and Logic Unit

The arithmetic operations such as +, - ,*, / etc as well as logical comparison are carried out by the ALU and the results obtained are transferred back to the memory unit for future retrieval.

Control unit

Control unit is playing the overall co-coordinating role in the entire system. It has direct control over the input and output units. As a program is being fed into the input unit, the information is converted into binary digits and stored in the memory, after the initiation from control unit. If the processing

requires any arithmetic calculation or logical operations to be carried out, the control unit transfer the relevant data to ALU.

Output unit

On completion of storing all the results of the processed data in the memory, the control unit transfers the relevant results from the memory unit, to the output unit in the converted form, which human beings can understand easily.

Some common output devices are printer, plotter and speaker.

1.7 Application of Computers

Computers are used in all walks of our life. They have become so deeply embedded in information processing and communication systems.

They are used for management applications in factories and companies, in schools to teach students, in banks to maintain client's account, in engineering offices for the design of buildings or machine components, in airport for reservation etc. and this list is endless.

Generally computers are used extensively in

Scientific and engineering applications

Satellites and space flight control, military weapons systems, earthquake calculations, architectural design and drafting.

Business, administrative applications

Payroll, financial accounting, sales analysis and forecasting, Bank accounting, handling cheques and credit systems.

Humanities applications (Entertainment)

Computer games, house applications, music analysis and composition, computer art, computer aided instruction and linguistic.

Further more, the applications can be classified into the following six types of activities:

Data acquisition (relevant collection)   Data management

 Analysis   Decisions   Actions   Monitoring.

1.7.1 Computers in Business

The new trends in IT integrated into business are

Transaction processing – Most business activities involve transactions with suppliers, employees or customers. Computers make these transactions easy to use.

Home-based or Mobile Workers – It is easy to connect a computer at home, to the computer at an office thereby enabling to work at home. These telecommuters (working at home) enable working at any time, without the traditional commute. Faster communication makes the home-based business easier to start and operate.

Desktop Publishing – Producing finished business literature is one of the uses of computers in business. Using desktop publishing programs and well featured word processing programs, people can create sales letters, brouchers, pricelists, newsletters and even very large manuals.

Financial Analysis - Financial analysis is performed in the company, from top management to all levels. Many tools are available to analyze investments, sales, expenses, markets and other aspects of business.

1.7.2 Computers in Industry

A computer has a major impact in industries where products are designed and manufactured. Computer Aided Design (CAD) is used to design products. Computer Aided Manufacturing (CAM) is used to produce them. The IT developments in industries are

Product Design – Designing complicated products can require the efforts of thousands of people work together. This team work is greatly enhanced through the use of computerized design.

Automization – Computers are used to run robots to create, finish, assemble and test products and their components.

Architectural walk-through – To help visualize how a building will really look like, when it's build computerized walk-through are prepared. One application of this technology on the web is Home building and Home decoration site. This offers information tips and advice on

- Build offline; plan online and dream a home.

- Estimating material costs and details of materials such as type of floor, roof etc.

- Tips and information about choosing the right color scheme, offers a range of shades, estimation of the amount of paint required and the cost of the paint varieties.

- Find an architect and dealers of the materials needed.

- Designing and decorating the room online.

- Remodeling the home, rearrange the furniture virtually with the handy room planner.

- Another application offers walk-through an auditorium or stadium, to select a seat for an event. Before purchase the ticket, we can even sit in the seat, to see what the angle of view will be like.

In case of business or industry two fundamental computing systems are

- Office Automation

- Management Information Systems

Office automation is defined as using computer and communications technology to help people for better usage and management of information. Office automation technology includes all types of computers, telephones, electronic mail and office machines that use microprocessors or other high technology components.

Office automation systems are comprised of many distinct subsystems like text management systems, business analysis systems, document management systems and network & communication systems.

Text management systems: These are used for tasks like writing memory notes, letters and other short documents, printing envelops and labels, preparing preprinted forms such as invoices, composing complex documents such as proposals and reports, retrieving and editing documents such as contracts, creating and display documents like newsletters etc.

Business analysis systems: Managers need solid data from which to extract the information necessary to make good decisions for the business.

Other software tools for performing analysis that are commonly used in large companies are decision support system which helps the knowledge worker to extract information from the various MIS (Management Information System) data base and reporting systems, analyze it and then formulate a decision or a strategy for business planning.

An *expert system* is a computer system that can store and retrieve data with special problem solving expertise.

An *executive support system* is an information system that consolidates and summarizes ongoing transactions within the organization. It provides the management with all the information it requires at all times from internal as well as external sources.

Document management systems: Office automation demands that, data be immediately accessible and instantaneously retrievable. For that reason, we are slowly moving away from paper and toward document forms that can be stored on the computer.

Network and communication management system: Today, knowledge workers have many ways to communicate with one another, primarily by voice, fax and e-mail. They can communicate in real time, via phone or computer. The

network and communication management systems include telephone, e-mail, voice messaging systems, teleconferencing and fax machines.

1.7.3 Computers in Home:

People use home computers generally for the work as they do in their offices. Word processor, data base programs, personal information management programs, accounting and income tax calculation programs are useful to do several activities for home users.

People use home computers for education and information collection. When the computer is connected to internet, encyclopedias, dictionaries, telephone directories, medical references, movies and animation are instantly available.

One can do research work or assignment, join special interest groups, download files, play games or make airline reservation without leaving home.

Sales catalogs are now being computerized and many catalogs are even being put on the web, so we can do online shopping. We can buy anything and pay through Internet banking.

One can read the news paper, magazines, shop advertisement or even read reviews of the movies in town, through internet.

The multimedia computers provide home entertainment. Most computer games are simulations, puzzle solving, animations, cartoons with controls; some other depends on eye-hand coordination.

Computer games can simulate card games, board games, sporting events, intergalactic battles or something else, real or imaginary. Interesting games can take people on a tour to solar system, a flight simulator allows both children and adults to experience being a pilot and learn about the principles of flying.

Computer games are replacing too many real world activities. Instead of playing music on guitars, one can play music on boom boxes. In general computer games improve the creativity and concentration.

1.7.4 Computers in Education

As technology has advanced, educational software has become a major influence at all levels from elementary schools to universities. Computers have opened many streams in the field of education. Some of them are

- Carrying out research in the remote fields in the absence of computer is impossible

- Experimenting various exercises pertaining to all subjects

- Using computers as teaching aid in class rooms

- Self learning by students, using various packages to gain knowledge. The tools such as Computer Aided Instruction (CAI), Computer Based Training (CBT) are useful for students.

Computers allow students to learn based on the drill-and-practice principle. The CBT software allow students to learn at their own pace, in small steps and give feedback about how much they have learned. This method have the principles namely, individual rate, small steps and positive feedback. Similar to this, is tutorial software, provides direct instruction in a clearly specified subject.

Drill-and-practice software and *tutorial software* are often referred to as CAI software. One of the major advantages of CAI is individualized learning. That is, individual students can learn their own pace. Teachers can spend their time, working with individual students. This is almost impossible in typical presentation and discussion class rooms.

Education for Kids – Puzzles, crosswords and word games increase problem solving and comprehensive skills. There are many programs available, to teach basics such as counting, maths, letters, spelling etc., through interactive tools.

There is no lower limit to use the computer. One can start off even when the kid is just a year old. The program called 'lapware', gives bright graphics.

They have to be responsible to any movements on the keyboard and the mouse. For example, images on the screen changes every time the child touches the keyboard.

E-learning – It means learning through web, putting up lessons on the corporate intranet, video broadcasting, CD-ROM based training in conjunction with real time online personal help, or a combination of these elements. One can learn almost anything, with e-learning from computer basics to programming, new technologies and more.

Training programs in business and other organizations are huge and expensive. To save money and improve performance, companies are using CBT, to train people on procedures and techniques they need to know in their jobs. Assume that an organization has several branch offices and the training is in head office. The advantages of e-learning are many.

- An organization can have employees anywhere in the world, and still make the same training content available to them.

- Save time on travelling to attend training sessions

- The content can also be made available quickly so that training can happen in time for the projects to start.

- The employees can access the learning content at anytime of the day or night.

1.7.5 Computers in Movies

Computers are used in movies to create and enhance the productivity and performance. Computers are used for, creating special effects in movies, editing movies, creating full length movies with cartoon characters, multimedia presentations, composing-editing-recording-reproducing music and sound effects.

Examples of movies which use artificial effects are

Science fiction movies – Starwars, E.T, Matrix

Special effects in movies – Titanic, Terminator

Cartoon movies – A Bug Life, Toy Story, Antz

Other techniques are, computer controlled lighting, control the images and sounds using computers, computer animation by projecting series of still images (frames), etc.

UNIT II : INPUT DEVICES

Input Devices: Keyboard – Pointing Devices – webcam – Scanners – Optical character Recognition – Magnetic Ink Character Recognition – Bar Code Reader. Output devices: Printers – Plotters – Computer Output Microfilm – Monitors – Voice recognition System – Projectors.

2.1 Input Devices -Introduction

Information is entered into a computer through input devices. An input device converts input information into a suitable binary form acceptable to a computer. The input units have direct access to the CPU.

The commonly used input devices are keyboard and mouse. Punched card reader, Paper tape reader, magnetic tape, Magnetic disk are the oldest input devices.

Input devices which do not require typing of input information are Mouse, joystick, light pen, graphic tablet (graphics input devices), touch screen and track balls. These allow users to select one of the items or images displayed on the screen. Therefore, these devices are called pointing devices.

2.2 Keyboard

Keyboard is an input device consisting of a set of typewriter like keys that enables the user to enter data into a computer. Though it is similar to typewriter keys, it also contains some additional keys.

Key layouts: The arrangement of keys on the keyboard is termed as the key layouts. *QWERTY and Dvorak-Dealey* are the two types of layouts which are universally used.

The standard layout of letters, numbers and punctuation in QWERTY keyboard indicates the five keys on the top row of letters QWERTY. Another keyboard design which has letters positioned for speed typing is the Dvorak-Dealey. This keyboard is developed by August Dvorak and William L.Dealey and hence named after them. The basic goal is to faster the alteration of hands in

typing. This keyboard arrangement places all vowels in the home row under the left hand's fingertips and the consonants used most often in the right hand's home row.

Keyboard keys:

The original PC keyboard has 84 keys and enhanced keyboards came with 101 keys and some with 104 keys. Latest keyboard consists of 107 keys, with additional hot keys (23 hotkeys) for multimedia applications.

The keyboard keys are classified into the following types:

Alphabetic keys	Numeric keys	Cursor-movement keys
Arrow keys	Modifier keys	Function keys
Special keys	Hot keys	

Alphabetic keys: A-Z

Alphabetic characters 'a' to 'z' are typed using these keys. Shift plus characters or "Caps lock" key ON is used to type capital letters.

Numeric keys: 0-9

Numeric keys are located on the keyboard at two places. One set is in the top row above alphabetic keys and the other set of these keys are on the right hand side of the keyboard which can be used with "Num lock" key ON.

Cursor-movement keys:

Cursor is an on-screen blinking character that shows where the new character will appear. Cursor-movement keys are used to move the cursor on screen. The standard keys of this type are Arrow keys, Home/End, Tab, Page up/Page down.

Arrow keys

The arrow keys move the cursor in the direction indicated by the arrow.←, →, ↑, ↓.

Home – to move the cursor to the beginning of a line

End – to move to the end of a line

Tab – key moves the cursor to predefined tab stops

Page up – to jump to the previous screen

Page down – to jump to the next screen

Modifier keys

Shift – it is used for typing capital letters and special characters. In combination with both Ctrl and Alt keys, it is sometimes used to execute a particular command.

Ctrl, Alt – these keys does not have any purpose. However, combined with different keys, it performs some useful operation. Some examples are

Ctrl + left arrow – moves to the previous word

Ctrl + Shift + Home – selects all the text to the left of the cursor to the beginning of the document

Special keys

Punctuation keys – consisting of comma, semicolon, colon etc

Arithmetic keys - , +, _, *, /

Backspace – it moves the cursor to go back by one character and delete it

Spacebar – is used to leave space between words

Enter –In text editing, pressing Enter key moves the cursor to a new line. In programming it is used to execute the external command.

Delete - to erase the wrongly entered character

Insert – this key is used to insert characters

Print screen – to print the current page of the screen

Scroll lock – to prevent the screen from scrolling

Pause – Temporarily suspend the current task. If the output of the command exceeds more than one screen, we can see screen by screen by pressing pause.

Break – to escape from the unhandled error or used for abnormal termination.

SysRq – keyboard interrupt key

Function keys: F1-F15

These keys are used for specific purposes as defined in the program we are using. The newer keyboard consists of extra functional keys for increased productivity. Example includes surfing the internet, running a CD audio player or opening the mail box etc.

Windows 107 keys

A few functions in window 95 don't fall readily under the caps of 104-keys keyboard. Three additional keys (two windows key and a popup menu key) are added to better match the operations of windows.

The two window keys, identified by the zooming window logo on their caps, serve as attention keys to pop into the windows task manager. One is located on the left of the left Alt key and the other to the right of the right Alt key. The third key serves to select the item at which the mouse cursor points, which is located to the right of the right side window key.

In 107-keys keyboard, the three additional buttons are called sleep, wakeup and power. The sleep button puts the computer in suspend mode. The wakeup button pulls it out of suspend mode. The power button works like a toggle switch for suspend and wakeup. While the computer is on suspend mode, pressing any other key will also bring it out of suspend mode.

Hot keys

Hot keys are specially meant for multimedia applications. They are normally placed above the function keys. They look like small round/oval shaped buttons. Some of the hot keys are

Volume control keys – Mute, Volume increase, Volume decrease

CD player keys – Preview, Play, Next, Stop, Eject

Recorder keys – Review/Play, Record/Stop for audio/Video recording

Edit Keys: Cut, Copy, Paste

Other short cut keys: Rotate Close, My Document, Game, Calculator, Menu, WWW, accessing E-mail, instant messenger etc. These keys allow the multimedia user for faster web surfing and greater efficiency when working with applications.

2.3 Pointing Devices

A pointer represents a small symbol on the screen. It usually appears on the screen in Graphical User Interface (GUI) environment. A pointing device is an input device. It is used to control the movement of the pointer or cursor on the screen. It can also be used for:

Sending command signals to the computer.

Selecting items on the screen.

Selecting commands from commands menu and Drawing.

Mouse

Mouse is the most commonly used pointing device, originally referred to as X-Y Position Indicator. It is used to control the cursor or pointer on the screen and to give commands to the computer. The mouse is attached with the computer by a cable or wireless connection. .

Types of mouse

Mechanical

It has a rubber or metal ball at the bottom that can roll in all directions. Mechanical sensors within the mouse detect the direction. The rolling ball moves the screen pointer according to the movement.

Optomechanical

It uses optical sensors to detect the movement of the ball.

Optical

It uses a laser to detect the mouse movement. The mouse is moved along a special mat with a grid so that the optical mechanism has a frame of reference.

An optical mouse has no mechanical moving parts. They respond more quickly and precisely than mechanical and optomechanical mouse.

Functions performed by a mouse:

Pointing	Clicking	Double clicking
Moving	Clicking and Dragging	Scrolling

are the functions performed by the mouse. It has a left button, a right button, and a scroll wheel.

Pointing: To point at something on screen, move the mouse until the pointer points on it.

Clicking: To select an object on the windows screen. For left click on screen item, press and release the left mouse button once. For right click, point with the mouse and then press and release the right mouse button.

Double clicking: To start an application. The left button of the mouse is pressed twice in quick succession by keeping the pointer on the desired object.

Moving: To move the pointer on the desired direction. It helps to point the object on the screen. No button is pressed while moving the mouse.

Clicking and dragging: To move an object from one place to another. The selected object moves along with the mouse pointer when it is clicked and dragged. Hold the left button of the mouse down and move the mouse so that click and drag is performed. Dragging is also used to draw an item on the screen.

Scrolling: Scrolling the page up and down with the help of a scroller. The small scroller is placed in between the two buttons.

The mouse also lets users zoom in and out quickly.

Trackball

A trackball is a pointing input device. It performs functions like a mouse but it is a stationary device. It has moveable ball on its top. The ball is rotated or rolled with fingers (or palm of the hand) to move the pointer on the screen. Like

mouse, a trackball also has buttons that is used to send the commands to computer.

The trackball is usually available with laptop computers. It is fixed on its keyboard. It is also available as separate input device.

Touch Pad

Touch pad is also known as track pad. This is a pressure-sensitive pointing device. This stationary device has no moving parts. It is a small, flat surface (or sensitive pad) over which a user slides fingertip to move the pointer on the screen.

Touch Pad also has one or two buttons. These buttons are located near the pad. These buttons work like mouse buttons. Touch Pad is normally used with laptops.

Pointing Stick

A pointing stick is a pressure-sensitive pointing device. It looks like a pencil eraser. It exists between keys of keyboard. It is used to control the movement of a pointer on the screen. The pointer in the screen moves in the direction in which the pointing stick is pushed. It is normally used with laptop.

Joystick

Joystick is a pointing device. It consists of a vertical handle or hand-held stick, which is mounted on a base. This stick is used to control the movement of pointer on the screen. Joystick also contains one or two buttons. The button of the joystick is pressed to activate certain event or action. Joystick is basically used to play video games. It is also used in some computer-aided design (CAD) system.

Touch Screen

A touch screen is a special video display screen. Input is given to the computer directly by touching the screen with user fingertip. Most touch screen computers use sensors in the computer's screen to detect the touch of a finger.

Usually touch screens used to make selection from a menu of choices displayed on the screen. Touch screen is commonly used in ATM. They are also used in restaurants, supermarkets, departmental stores etc.

In some computer systems, pen-based or light pen input devices are used. The user touches the screen with a pen. However, the term touch screen implies a system that accepts input data by touching the screen with the user's fingertip.

Light Pen

The light pen is a hand-held pointing input device. It looks like a pen. It is a light-sensitive input device. It is connected with the computer by a wire. The tip of pen contains a light-sensitive element. This device also contains a button.

Input is given to the computer directly by touching the screen with pen. When the pen touches the screen, it detects or senses a position on display screen. The user points the object to be selected on the display screen and then presses a button.

Light Pen device is usually used by engineers and graphic designers. This input device requires special monitor. Light pen is used in PDA and other types of hand-held computers.

### Digitizer or Graphics Tablet

Digitizer is used for drawing new images or tracing old images. Usually, it is used for engineering drawings. A digitizer is also known as graphics tablet. It consists of a flat, rectangular board (or pad) and an input device stylus or puck. Each location on the rectangular board points to corresponding position on the computer screen. The stylus or puck is connected to the board / pad by a wire.

Stylus looks like a pen. It is used for drawing images on a specially designed graphics screen or digitizer. Puck looks like a mouse. Puck is also

called cursor. Usually, it is used for tracing old images. Figure 2.1 shows some examples of the pointing devices

Joystick

Light Pen

Trackball
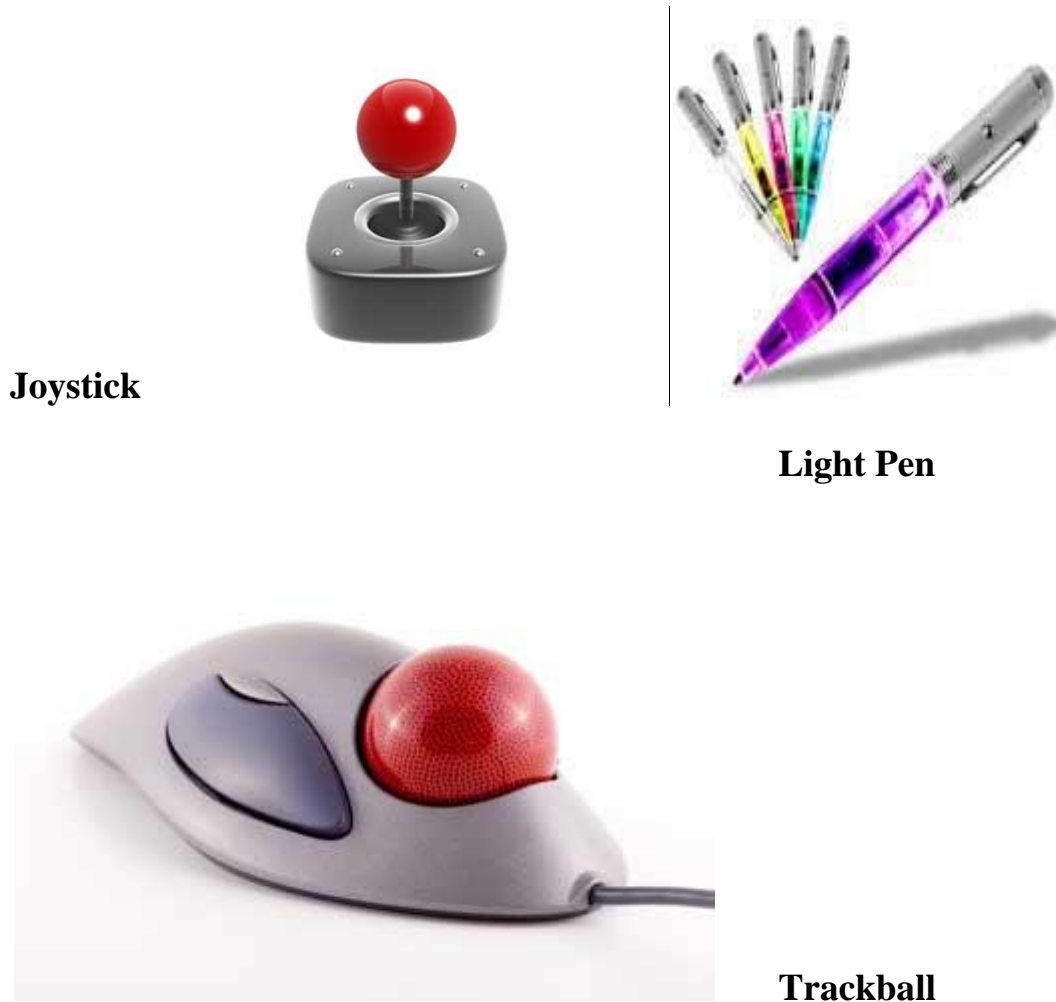
Figure 2.1 Examples of pointing devices

2.4 Webcam

A webcam is an input device that captures digital images. A webcam – short for 'web camera' – is a digital camera that's connected to a computer.(Figure 2.2) It can send live pictures from wherever it's sited to another location by means of the internet

Unlike a digital camera and digital camcorder, a webcam does not have any built-in storage. Instead, it is always connected to a computer and uses the computer hard drive as its storage.

Today, most webcams are either embedded into the display with laptop computers or connected to the USB or FireWire port on the computer. Some others are wireless (wifi).

Other features might include:

- An integral microphone

- The ability to pan and tilt

- In-built sensors that can detect movement and start recording

- A light that, when on, will let us to know that the camera is in use.

There's a wide range of things that we can do with a webcam. Webcams are mostly used in videoconferencing and for security surveillance. Other uses include video broadcasting, social video recording and computer vision. The most common is to video chat over the internet using Skype.

Figure 2.2 Web cam

2.5 Scanning Devices

Much of the data we want to computerize already exists as hardcopy (paper) documents. Instead of retyping this data, the scanning devices can scan (read) documents and transfer it, into the computer.

Scanner

Scanner is one of the input devices which are used to transfer drawings, graphs, photographs or text, inside the computer in the form of digital images. The scanner does an optical scanning mechanism over the information to be stored. Once the digital representation of the picture is obtained various editing operations can be performed on it.

A scanner works by passing light at the hard copy being scanned and measuring intensity of light reflected back using an optical sensor. The amount of light that is reflected back indicates how light or dark the image is, at each point.

There are various types of scanners. The most common are Drum, flatbed and hand scanners. Drum scanners use photomultiplier tubes (PMT). Flatbed scanners use a glass pane and a bright light. Hand scanners are dragged across the image manually.

Flatbed scanner

The flatbed scanner is also known as a desktop scanner. The most popular type of desktop scanner is flatbed scanner, so called because of its flat, glass platen (or bed) which serves as both the scanning area and surface for laying objects down to be scanned. Most flatbeds are used for scanning reflective art.

A flatbed scanner works in a similar fashion to a photocopier. The image

to be scanned is placed on a piece of glass. The scanning beam then moves across the image. It can also be found mounted with printers.

Flatbed scanners are available in various size and capability. Desktop full color scanners, large floor model scanners are some of them. They vary in resolution, Platform size, and color depth. The highest the resolution of the scanner the sharper will be the image reproduced.

Drum Scanners

Drum scanners use Photo Multiplier Tube (PMT) technology for greater dynamic range and color accuracy. Drum scanners can produce more scans per hour than a flatbed scanner. This offer features such as direct conversion to CMYK, auto sharpening, batch scanning, greater dynamic range, and huge image scanning areas.

2.6 Optical Character Recognition (OCR)

It is an input device. It is so designed as to read numeric and alphabetic character from printed documents.(Figure 2.3) However, the recent OCR's can read handwritten characters.

All OCR systems include an optical scanner for reading text, and sophisticated software for analyzing images. Most OCR systems use a combination of hardware and software to recognize characters, although some inexpensive systems do it entirely through software. It is a common method of digitizing printed texts so that they can be electronically edited, searched, stored more compactly, displayed on-line, and used in other machine processes.

Early versions needed to be trained with images of each character, and worked on one font at a time. Advanced systems are capable of producing a high degree of recognition accuracy for most fonts. Some systems are capable of reproducing formatted output that closely approximates the original page including images, columns, and other non-textual components.

They can be used for:

- OCR is being used by libraries to digitize and preserve their holdings.

- Data entry for business documents, e.g. passport, invoice, bank statement and receipt
- Automatic number plate recognition
- Automatic insurance documents key information extraction
- Extracting business card information into a contact list
- More quickly make textual versions of printed documents, e.g. book scanning
- Make electronic images of printed documents searchable, e.g. Google Books
- Assistive technology for blind and visually impaired users free guide
- Billions of magazines and letters are sorted every day by OCR machines, considerably speeding up mail delivery.

Figure 2.3 OCR Reader

2.7 Optical Mark Recognition *(OMR)*

OMR (optical mark recognition) is a form of automated data input. Marks are made on a specially printed paper forms which are then read by an OMR reader. The data is then sent to a computer for processing. One of the most

common uses of OMR is the use of "bubble sheets" in multiple choice examinations. Students mark their answers on specially printed sheets using either a pencil or a special marker. The sheets are then collected in and the data read using a sheet feed OMR scanner. (Figure 2.4)

Other uses are

- Attendance registers in schools
- Survey forms for market research
- Time sheets in factories
- Lottery Forms

The advantages are this technology is easy to use and requires little training and this helps us to easily detect mistakes in an answer sheet.

Figure 2.4 OMR Reader

2.8 Magnetic ink character reader (MICR)

Magnetic ink character recognition code (MICR Code) is a character-recognition technology used mainly by the banking industry to ease the processing and clearance of cheques and other documents.

The MICR encoding called the *MICR line*, is at the bottom of cheques and other vouchers and typically includes the document-type indicator, bank code, bank account number, cheque number, cheque amount, and a control

indicator. The technology allows MICR readers to scan and read the information directly into a data-collection device. (Figure 2.5)

The ink used in the printing is a magnetic ink or toner, usually containing iron oxide. The MICR text is passed before a MICR reader. The ink in the plane of the paper is first magnetized. Then the characters are passed over a MICR read head, a device similar to the playback head of a tape recorder. As each character passes over the head it produces a unique waveform that can be easily identified by the system.

The use of MICR allows the characters to be read reliably even if they have been overprinted or obscured by other marks, such as cancellation stamps and signature. This provides a high level of security because any attempt to alter the magnetic ink printout with normal ink by writing over it will be ignored.

Figure 2.5 MICR Reader

2.9 Bar Code Reader

Bar code reader, also called a price scanner or point-of-sale (POS) scanner. Bar Code Reader is a device used for reading bar coded data (data in the form of light and dark lines). Bar coded data is generally used in labeling goods, numbering the books etc. It scans a bar coded data, converts it into an alphanumeric value which is then fed to the computer to which bar code reader is connected. (Figure 2.6)

There are five basic kinds of barcode readers -- pen wands, slot scanners, Charge-Couple Device (CCD) scanners, image scanners, and laser scanners.

- A pen wand is the simplest barcode reader. It contains no moving parts. A pen wand scanner has to remain in direct contact with the bar code, must be held at a certain angle, and has to be moved over the bar code at a certain speed.

- A slot scanner remains stationary and the item with the bar code on it is pulled by hand through the slot. Slot scanners are typically used to scan bar codes on identification cards.

- A CCD scanner has a better read-range than the pen wand and is often used in retail sales. Typically, a CCD scanner has a "gun" type interface and has to be held not more than one inch from the bar code. Each time the bar code is scanned; several readings are taken to reduce the possibility of errors. The disadvantage of the CCD scanner is that it cannot read a bar code that is wider than its input face.

- An image scanner, also called as camera reader, uses a small video camera to capture the image of the bar code and then uses sophisticated digital image processing techniques to decode the bar code. It can read a bar code from about 3 to 9 inches away and generally costs less than a laser scanner.

- A laser scanner does not have to be close to the bar code for scanning. It uses a system of mirrors and lenses to allow the scanner to read the bar code regardless of orientation, and can easily read a bar code up to 24 inches away. To reduce the possibility of errors, laser scanners may perform up to 500 scans per second. Specialized long-range laser scanners are capable of reading a bar code up to 30 feet away.

Figure 2.6 Barcode Reader

2.10 Output Devices

The peripheral devices used to present the results computed by the computer in a human readable form are called output devices. The output devices receive results and other information from the computer and provide them to users.

The computer sends information to an output device in a binary form. An output device converts it into a suitable form convenient to users such as printed form, display on screen, voice output, etc. The commonly used devices are monitors, printers and plotters. What you see on the monitor will be printed on the printer. This is known as "What You See is What You Get" (WYSIWYG).

2.10.1 Printers

Printers are more commonly used output device. Printers are used to produce printouts or hard copies of the output data from the computer. The various factors to be considered in selecting a printer are printing speed, on-paper quality, color capabilities, print engine, media of handling and the cost. According to the method of printing, printers can be classified as follows.

Serial printers (character printers)

Line printers

Page printers

The serial printers can print a character at a time and its speed is determined by characters per second. It is also classified into two categories according to the technology employed. They are,

Impact printers   Non-impact printers

An impact printer contacts with the paper, by pressing an inked ribbon against the paper using a hammer or pins for printing. Example: Dot matrix printers and Line printers

Non-impact printers do not use a striking device to produce characters on the paper, but it sprays ink on the paper. The following tree (Figure 2.7) summarizes the various classifications of printers with example.
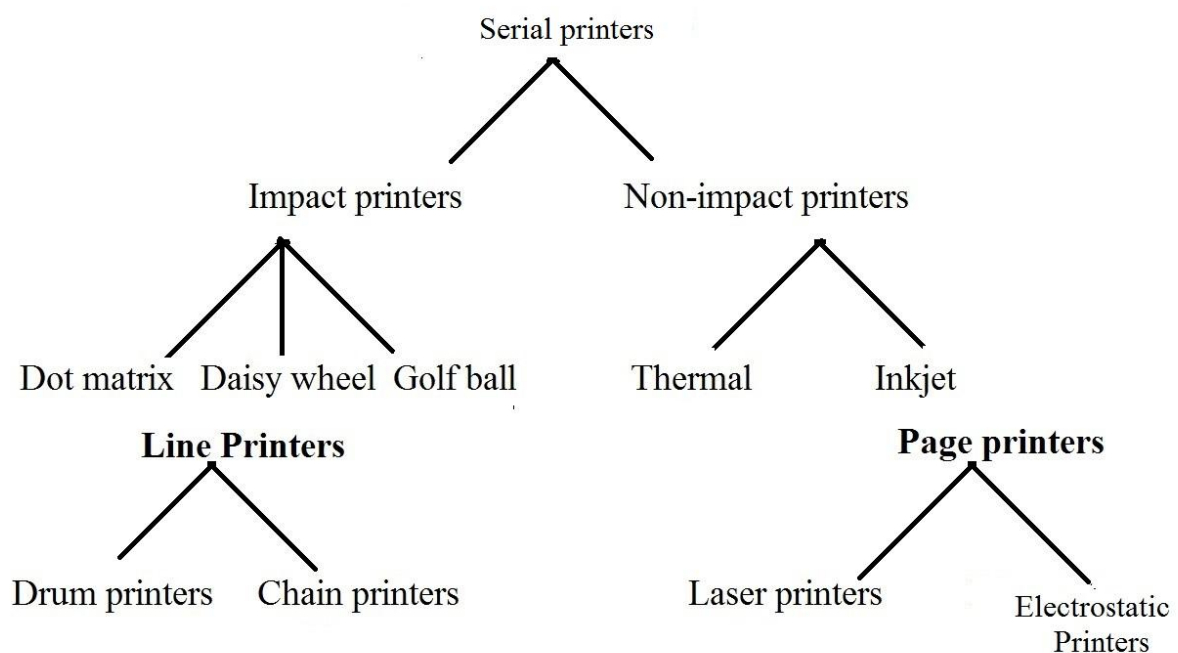

Figure 2.7 Classification of printers

Impact Printers

Dot matrix printers

The dot-matrix printer uses print heads containing from 9 to 24 pins. These pins produce patterns of dots on the paper to form the individual characters. The 24 pin dot-matrix printer produces more dots that a 9 pin dot-matrix printer, which results in much better quality and clearer characters. The more pins, the clearer the characters on the paper. The pins strike the ribbon

individually as the print mechanism moves across the entire print line in both directions.

The user can produce a color output with a dot-matrix printer by changing the black ribbon with color stripes ribbon. Dot-matrix printers are inexpensive and typically print at speeds of 100-600 characters per second.

Daisy-wheel printers

It is called daisy-wheel printer because the print mechanism looks like a daisy; at the end of each "Petal" is a fully formed character which produces solid-line print. A hammer strikes a "petal" containing a character against the ribbon, and the character prints on the paper. It prints typically 25-55 characters per second.

Line printers

Line printers use special mechanism that can print a whole line at a time. They can typically print in the range of 1,200 to 6,000 lines per minute. Drum, chain, and band printers are line printers.

Drum printer

A drum printer consists of a solid, cylindrical drum that has raised characters in bands on its surface. The number of print positions across the drum equals the number available on the page. This number typically ranges from 80-132 print positions. The drum rotates at a rapid speed. For each possible print position there is a print hammer located behind the paper. These hammers strike the paper, along the ink ribbon, against the proper character on the drum as it passes. One revolution of the drum is required to print each line. This means that all characters on the line are not printed at exactly the same time, but the time required to print the entire line is fast enough to call them as

line printers. Typical speeds of drum printers are in the range of 300 to 2000 lines per minute.

Chain printers

A chain printer uses a chain of print characters wrapped around two pulleys. Like the drum printer, there is one hammer for each print position. Circuitry inside the printer detects the correct character at the desired print location on the page. The hammer then strikes the page, pressing the paper against a ribbon on the character located at the desired print position. An impression of the character is left on the page. The chain keeps rotating until all the required print positions on the line have filled. Then the page moves up to print the next line. Speeds of chain printers range from 400 to 2500 characters per minute.

Band printers

A band printer operates similar to chain printer except that it uses a band instead of a chain and has fewer hammers. Band printer has a steel band divided into five sections of 48 characters each. The hammers on a band printer are mounted on a cartridge that moves across the paper to the appropriate positions. Characters are rotated into position and struck by the hammers. Font styles can easily be changed by replacing the band or chain.

Non-impact printers

Ink-jet printers

Ink-jet printers work in the same fashion as dot-matrix printers. The dots are formed by tiny droplets of ink. Ink-jet printers form characters on paper by spraying ink from tiny nozzles through an electrical field that arranges the charged ink particles into characters at the rate of approximately 250 characters per second. The ink is absorbed into the paper and dries instantly. Various

colors of ink can also be used. These printers produce less noise and print in better quality with greater speed.

Thermal printer

Thermal printing (or direct thermal printing) is a digital printing process which produces a printed image by selectively heating coated thermo chromic paper, or thermal paper. When the paper passes over the thermal print head, the coating turns black in the areas where it is heated, to produce an image. Two-color direct thermal printers can print both black and an additional color (often red) by applying heat at two different temperatures.

Page Printers

Laser printers

A laser printer works like a photocopy machine. It produce images on paper by directing a laser beam at a mirror which bounces the beam onto a drum. The drum has a special coating on it to which toner (an ink powder) sticks. Using patterns of small dots, a laser beam conveys information from the computer to a positively charged drum to become neutralized. From all those areas of drum which become neutralized, the toner detaches. As the paper rolls by the drum, the toner is transferred to the paper printing the letters or other graphics on the paper. A hot roller bonds the toner to the paper.

Laser printers use buffers that store an entire page at a time. When a whole page is loaded, it will be printed. The speed of laser printers is high and quietly without producing much noise.

2.10.2 Plotters

A plotter is a special output device used to produce hardcopies of graphs and designs on the paper. A plotter is typically used to print large-format graphs

or maps such as construction maps, engineering drawings and big posters. Different types of plotters are

Drum plotter   Flatbed plotter   Inkjet Plotter   Cutting Plotter

Drum Plotter

A drum plotter is also known as Roller Plotter. It consists of a drum or roller on which a paper is placed and the drum rotates back and forth to produce the graph on the paper. It also consists of mechanical device known as Robotic Drawing Arm that holds a set of colored ink pens or pencils. The Robotic Drawing Arm moves side to side as the paper are rolled back and forth through the roller to create a perfect graph or map. Drum Plotters are used to produce continuous output, such as plotting earthquake activity.

Flatbed Plotter

A flatbed plotter is also known as Table Plotter. It plots on paper that is spread and fixed over a rectangular flatbed table. The flatbed plotter uses two robotic drawing arms, each of which holds a set of colored ink pens or pencils. The drawing arms move over the fixed paper and draw the graph on it. Typically, the plot size is equal to the area of the flat bed. It is used in the design of cars, ships, aircrafts, buildings, highways etc. Flatbed plotter is very slow in drawing or printing graphs. The large and complicated drawing can take several hours to print. The main reason of the slow printing is due to the movement of mechanical devices.

Mechanical plotters have been replaced by thermal, electrostatic and ink jet plotters. These systems are faster and cheaper. They also produce large size drawings.

Inkjet Plotter

The inkjet plotter creates an image by spraying small droplets of ink on to paper. A popular choice for advertising agencies and graphic designers, inkjet plotters are used generally for large outputs, such as banners, billboards and

large sign boards often seen along roadsides. They are available in thermal or piezoelectric models. Thermal inkjet plotters use heat to apply droplets of ink, while piezoelectric plotters use charged crystals to apply the ink. Inkjet plotters typically produce better quality graphics than other types.

Cutting Plotter

The cutting plotter is a large scale cutting device that produces ready-cut mylar or vinyl lettering and graphics. Automated plotter knives cut the sheet of the material lying on the plotter's flat surface area, carving out the design stored in the computer. Used for sign making, billboard advertising and vehicle stickers graphics. This devices offer far greater speed and precision than, traditional method of creating sign lettering and logos by hand.

2.10.3 Computer Output Microfilm

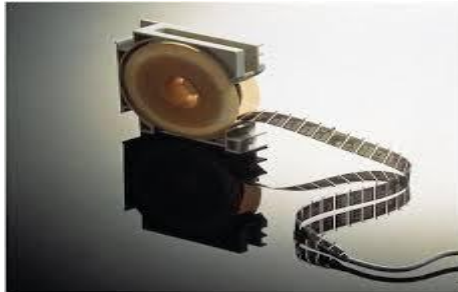**Computer output microfilm** (**COM** Computer Output Microfilm is a system that converts stored data directly to microfilm or microfiche.(Figure 2.8) COM technology has been used for document and newspaper archival since the 1920s. And, with the advent of the personal computer, computer-produced microforms are becoming more popular for non-library use, such as storing catalogs and patient records.

COM systems are still used today, mostly by organizations who need to store payroll, accounting, insurance, inventory, or employee data. Yet because most of these organizations have outputted the Computer Output Microfilm to microfiche, they have to manually search for a record and use a reader printer to save out a particular file.

A more feasible option is to convert Computer Output Microfilm to digital image via microfiche scanning services. Using high quality/high production microfiche scanners, we convert the COM microfiche to digital. The

end result of this entire process is a digital image which we could easily search the file containing records by name, account number, date, and/or social security number. Obviously this saves time and efficiency.

Figure 2.8 Computer Output Microfilm

2.10.4 Monitors

Display systems are the focal point of any human-machine interface. Display systems have been used in digital watches, laptops, mobile phones, digital cameras, PC monitors, TV etc; this allows seeing exactly, what the PC is doing as it works. The output on the monitor is called softcopy output.

While monitors come in a variety of shapes, designs, and colors, they can be broadly categorized into three types. They are

CRT (Cathode Ray Tube) monitors
LCD (Liquid Crystal Display) monitors
LED (Light Emitting Diodes) monitors
Can be either be monochrome or color monitor.

A monochrome monitor is a type of CRT computer display which was very common in the early days of computing. Monochrome monitors actually display two colors, one for the background and one for the foreground. The colors can be black and white, green and black, or amber and black.

Color monitors can display color ranging from 16 to over 1 million different colors. These are sometimes called RGB monitors since it uses RGB as base colors (Red, Green, and Blue).

CRT monitors

These monitors employ CRT technology, which was used most commonly in the manufacturing of television. With these monitors, a stream of intense high energy electrons is used to form images on a fluorescent screen. A cathode ray tube is basically a vacuum tube containing an electron gun at one end and a fluorescent screen at another end.

The features to be considered are

Monitor *screen size* is measured diagonally across the screen, in inches.

The *resolution* of the monitor is the maximum number of pixels it can display horizontally and vertically (such as 800 x 600, or 1024 x 768, or 1600 x 1200). Most monitors can display several resolutions below its maximum setting.

Pixels (picture elements) are the small dots that display the image on the screen.

The spacing of the screen's tiny phosphor dots is called the *dot pitch* (dp), typically .28 or .26 (measured in millimeters). A screen with a smaller dot pitch produces sharper images.

LCD monitors

A flat panel display usually uses an LCD screen to display output from the computer. The LCD consists of several thin layers that polarize the light passing through them. The polarization of one layer, containing long thin molecules called liquid crystals, can be controlled electronically at each pixel, blocking varying amounts of the light to make a pixel lighter or darker.

Flat panel displays are much lighter and less bulky than CRT monitors, and they consume much less power. The display size of a flat panel is expressed in inches. Cold cathode fluorescent (CCFL) back-lighting is used in LCDs.

Organic Light Emitting Diode (OLED) monitors provide higher contrast and better viewing angles than LCDs but they require more power when displaying documents with white or bright backgrounds. Types of OLED are PMOLED and AMOLED Passive and Active Matrix Organic Light Emitting Diode respectively.

Plasma Display is another type of flat panel display that utilizes small cells containing electrically charged ionized gas.

LED monitors: These are flat panel, or slightly curved displays which make use of light-emitting diodes for back-lighting. The advantages of LED monitors are that they produce images with higher contrast, have less negative environmental impact when disposed, more durable than CRT or LCD monitors, does not produce much heat and feature a very thin design.

2.10.5 Voice Recognition System

Speaker

Speaker is used to produce audio output that can be heard by the listener. The sound produced by speakers is defined by frequency and amplitude. The frequency determines how high or low the pitch of the sound is. A speaker system's ability to accurately reproduce sound frequencies is a good indicator of how clear the audio will be. Amplitude, or loudness, is determined by the change in air pressure created by the speakers' sound waves.

Speakers typically come in pairs, which allow them to produce stereo sound. This means the left and right speakers transmit audio on two completely separate channels. By using two speakers, music sounds much more natural since our ears are used to hearing sounds from the left and right at the same time.

For portable laptops and mobile devices, the speakers are built into the device. Some desktop computers have speakers permanently installed to the monitor.

Headphone

Headphone is used to hear sound output from the computer. They are similar to speakers, except that, only the person who is wearing the headphone can hear.

2.10.6 Projectors

This device captures the display of a computer screen and project a large version of it onto a flat surface. Cathode Ray Tube (CRT) projector, Liquid Crystal Display (LCD) projector and Digital Light Processing (DLP) projector are the various types of projector.

CRT projector - In the early days of projectors, CRT projectors were commonly used. It has three tubes, one for each of the primary colors. Due to their large size, low light output and the frequent need to converge and align the images projected from each of the three tubes, they are no longer commonly used

LCD projector - LCD projectors work by utilizing polarized mirrors that pass and reflect only certain colors of light. This causes each channel of red, green and blue to be separated and later re-converged via a prism after passing through an LCD panel that controls the intensity and saturation of each color. An LCD projector can achieve greater brightness at lower energy consumption and are smaller than CRT projectors

DLP projector - DLP projectors can be classified as one-chip or three-chip projectors. One-chip DLP projectors can produce more than 16 million colors while three-chip models can produce more than 35 trillion colors. This allows DLP projectors to reproduce more natural and lifelike images

UNIT III : PRIMARY MEMORY

Primary Memory: Memory Representation – Memory Hierarchy – Random Access Memory – Read only Memory – Types of ROM.

Secondary Storage: Classification of secondary storage devices – Storage organization of Magnetic Disk – Storage organization of optical Disk – Magneto Optical Disk – Universal Serial Bus.

3.1 Computer Memory - Introduction

The function of the memory is to store information. It stores program, data, results or any other kind of information. Various types of memories are

> Main memory or primary memory
>
> Secondary memory or Auxiliary memory
>
> Cache memory

3.2 Primary memory

The memory which is built as a part of the CPU is called the primary memory or main memory. The memory capacity representation is listed in *Table 3.1*

The main memory can be classified into two types as magnetic core memory and semiconductor memory.

Magnetic core memory

A series of electromagnets are used to represent the characters or instructions.

Table 3.1 Memory capacity representations

Bit	A binary digit 0,1
Byte	8 bits or 1 character
kiloByte(KB)	1024 Bytes
MegaByte(MB)	1024 KB
GigaByte(GB)	1024MB

TeraByte(TB)	1024GB
PetaByte(PB)	1024TB
ExaByte(EB)	1024PB
ZettaByte(ZB)	1024EB
YottaByte(YB)	1024ZB
ProntoByte(PrB)	1024YB
GeopByte	1024PrB

Semiconductor memory

These memories are built by using the flip flops as its basic cell. RAM and ROM are the two types of semiconductor memory. There are several variations in RAM and ROM as in *Figure 3.1*.
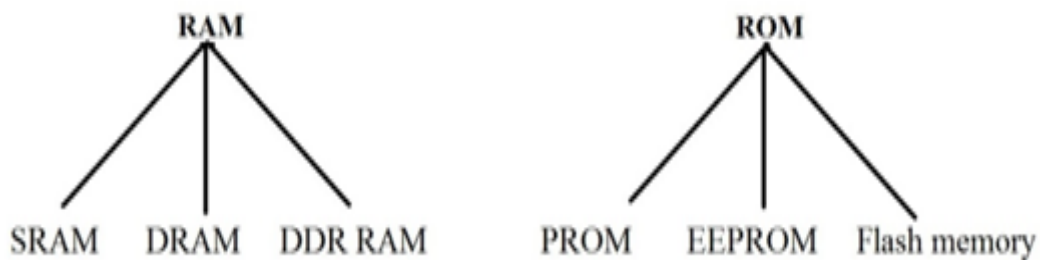

Figure 3.1 Categories of RAM and ROM

3.3 Memory Hierarchy

The hierarchical arrangement of storage in current computer architectures is call the memory hierarchy. The layers of memory in memory hierarchy in most computers are shown in Table 3.2. The important point to know about the memory hierarchy is the tradeoffs between speed and size -- the faster the memory the smaller it is.

CPU Registers
L1 Cache memory
L2 Cache memory
L3 Cache memory
Main memory - RAM
Rotating magnetic memory
Optical memory

Table 3.2 Layers of memory in memory hierarchy

The memory hierarchy in most computers is as follows.

- Memory locations built directly into the CPU are called *registers*. Processor registers are the special memory locations that are used to hold the data currently being processed. Each register is typically as wide as the processors data bus. The processor can read and write data values into these locations much more quickly than other memory. The size of the registers is measured in number of bits, sometimes called word size.

- Storing and retrieving data to and from RAM is one of the time consuming process. To avoid this complexity, a memory namely *Cache memory* in the CPU was introduced. Cache memory is similar to RAM, except that it is extremely fast.

  The most important type of cache memory is the *CPU cache*. A small form of memory located directly on the chip itself. It is much smaller, but can be accessed much faster than the main memory.

A computer can have several different levels of cache memory (L1, L2, L3). The level numbers refers to distance from CPU where Level 1 is the closest. All levels of cache memory are faster than RAM. The cache closest to CPU is always faster but generally costs more and stores less data than other level of cache.

- *RAM* - All instructions and storage addresses for the processor must come from RAM. Although RAM is very fast, there is still some significant time taken for the CPU to access it (this is termed *latency*). RAM is stored in separate, dedicated chips attached to the motherboard and it is much larger than cache memory.

- Magnetic and optical disk - Slow but very large storage capacity.

3.4 Random access memory (RAM) or Read/ Write memory

RAM comprises the main memory of the computer system. The RAM chips are arranged in rows and installed in the mother board of the computer. There are two types of RAM namely Dynamic RAM and Static RAM.

Types of RAM

Static Random Access Memory (SRAM)

It consists of flip-flops, a bistable circuit composed of four to six transistors. Once a flip-flop stores a bit, it keeps that value until the opposite value is stored in it. SRAM gives fast access to data, but it is physically relatively large. It is used primarily for small amounts of memory called registers in CPU and for fast "cache" memory.

Dynamic Random Access Memory (DRAM)

DRAM needs only one transistor cell. The data capacity is at least six times more than with SRAM. The high density feature and the associated cost-benefit ratio of DRAM make it suitable to form the core memory of PC's.

The two categories in DRAM are SDRAM (Synchronous DRAM) and RDRAM (Rambus DRAM).

Double Data Rate RAM

DDR RAM is already in use in video cards, and now used on computer motherboards as main RAM. DDR RAM will immediately double the memory bandwidth available to the processor.

By using this type of memory, the user can expect 5 to 7 percent increase in memory intensive applications such as Photoshop and CAD based applications.

3.5 Read Only Memory (ROM)

The data are stored permanently in the memory. ROM stores the programs that are used to boot the computer. ROM's are used extensively in calculators and peripheral devices such as laser printers, whose fonts are often stored in ROMs.

The difference between the RAM and the ROM is listed in the *Table 3.3*.

Table 3.3 Differences between RAM and ROM

RAM	ROM
The information are stored and retrieved by programmer's choice.	Constructed by the manufacturer. Information stored in this memory cannot be modified by the programmer (In EPROM and PROM modifications are possible.)
Volatile memory i.e. information is lost during power off.	Non-volatile memory.
Used to store intermediate results.	Permanent function like $\sin(x)$, $\cos(x)$ etc are stored for further uses.
Complex and expensive memory	Simple and cheap type of memory
Easy to change.	ROM's are expensive to change.

Types of ROM

Programmable Read Only Memory (PROM).

Used for, the small quantity of user defined applications. PROMs are manufactured as blank chips on which data can be written. The programming job is carried by the user.

Erasable Programmable Read Only Memory (EPROM)

In Erasable Programmable Read Only Memory, the contents of memory locations can be erased by passing the Ultra Violet light. This type of memory is useful for design and development application, where very frequent changes in memory contents are required.

Electrically Erasable programmable Read only memory (EEPROM)

The memory locations are cleared electrically and other processes are similar to the EPROM processes.

3.6 Classification of Secondary Storage devices

There are two main technologies used to store data, namely magnetic and optical storage. Magnetic storage devices include magnetic disk such as floppy disk, hard disk, removable hard disk and magnetic tape. Optical storage devices include CD-ROM, DVD, WORM, magneto optical disk.

Magnetic storage devices record data as magnetic fields. Optical storage techniques make use of the pinpoint precision that is possible only with laser beams.

The secondary storages are also called as external or auxiliary memory or backing storage. These devices are used to hold mass information which can be desirably transferred at any time. Two types of secondary memories, which have larger capacity are, Magnetic bubble memory and Charge coupled memory.

The advantage of magnetic bubble memory is that the large amount of data can be stored in a less area. They are considerably faster than

semiconductor memories. Charge coupled devices (CCD) are volatile memory. In CCD's presence or absence of electric charge represents a bit.

There exist different types of secondary storage devices, each of them suitable for a specific purpose. They mainly differ in the following aspects: Technology used to store data, Portability of storage device and Access time.

3.7 Storage organization of Magnetic Disk

The most common physical device for storing files is the magnetic disk. Actually, a disk typically contains several rotating disks, or platters. The surfaces of the platters are covered in metal oxide, and read/written by electromagnetic recording heads.

The surface of a platter is organized as a number of concentric tracks. Each track is divided into sectors. (Figure 3.2) The information held in one sector, a block, is the unit of transfer between the disk and primary memory. The time taken to access a particular block consists of:

- Time to move the heads to the right track - the seek time.
- Time waiting for the sector to come round to the head - the latency.
- Time to actually transfer the data - the block transfer time.

In addition, an empty index is also placed on the disk. The data is stored in sectors and tracks and the PC locates specific data with the help of the index. The placing of tracks, dividing them into sectors, and placing of an empty index is called *formatting* of the disk. Without formatting, the disk will not accept any data.

The basic unit of data storage in magnetic disk is called cluster. A cluster includes two or more sectors. The smallest space that a file can occupy is a cluster. Depending on the size of the file, it can take, track of the clusters, that hold the contents of each file.
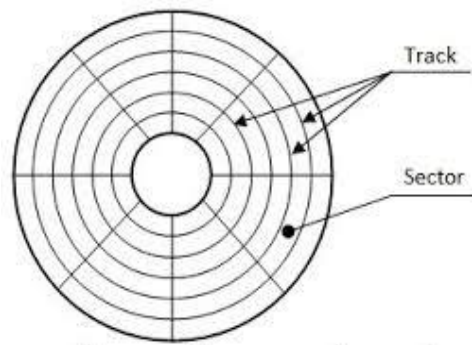
Figure 3.2 Track, Sector

3.7.1 Floppy Disk

It is a small, random access disk which can be used for both input and output operations. Floppy disks are often referred to as diskettes.

The disk is made up of flexible plastic and coated with magnetic oxide. For protection, it is normally contained within a plastic slave. It can be readily loaded into and unloaded from a drive unit. The heads on a floppy disk unit make contact with the disk surface when reading or writing.

The standard sizes of floppy disks are, 8 inches and 5.25 inches, and 3.5 inches. The storage capacity of 8 inch floppy disk ranges from 80 KB to 1,212 KB, 5.25 inch floppy disk ranges from 160KB to 1.2 MB, 3.5 inch floppy disk ranges from 400 KB to 2.88 MB.

Higher storage capacity floppy disk categories are Zip disks (100/250MB), Superdisk (120/250MB) and (High Capacity Floppy Disk) HIFD disk (250MB).

3.7.2 Hard disk

Hard disks are designed to store very high volume of data. Hard disks can store gigabytes to terabyte of data and they form an integral part of the computer. Most operating systems are stored in hard disks and all contents

(application software, documents, images, music files etc) in the computer are also stored in the hard disk.

The data is stored on a metal platter in the form of magnetic spots. The metal platter is sealed inside a disk drive that protects the platter and as well enables reading and writing to the disk. Hard disk may contain several platters forming hard disk packs that increase the storage capacity. These provide fast access for both reading and writing.

Four common hard disk storage technologies are:

- ATA (Advanced Technology Attachment)

- FireWire

- SCSI(Small Computer System Interface)

- RAID (Redundant Array of Inexpensive Disks) and Fiber Channel

3.8 Storage organization of Optical Disk

Optical disk, is any media, read using a laser assembly. The most common types of optical media are Blu-ray, CDs, and DVDs. Computers can read and write to CDs and DVDs using a CD Writer or DVD Writer drive, and a Blu-ray is read with a Blu-ray drive.

Optical memory is used for storing large volumes of data like audio, text, graphics and video. An optical disk is a removable disk that uses laser to read and write data. A laser in this context means an electromagnetic wave with a very specific wavelength within or near the visible light spectrum. Different types of discs require different wavelengths.

An optical drive that can work with multiple types of discs will therefore contain multiple lasers. The mechanism to read and write data consists of a

laser, a lens to guide the laser beam, and photodiodes to detect the light reflection from the disc.

The optical mechanism for reading CDs and DVDs are quite similar, so the same lens can be used for both types of discs. The mechanism for reading Blu-ray discs, however, is quite different. An optical drive that works with all types of discs will therefore have two separate lenses: one for CD/DVD and one for Blu-ray. (Figure 3.3)
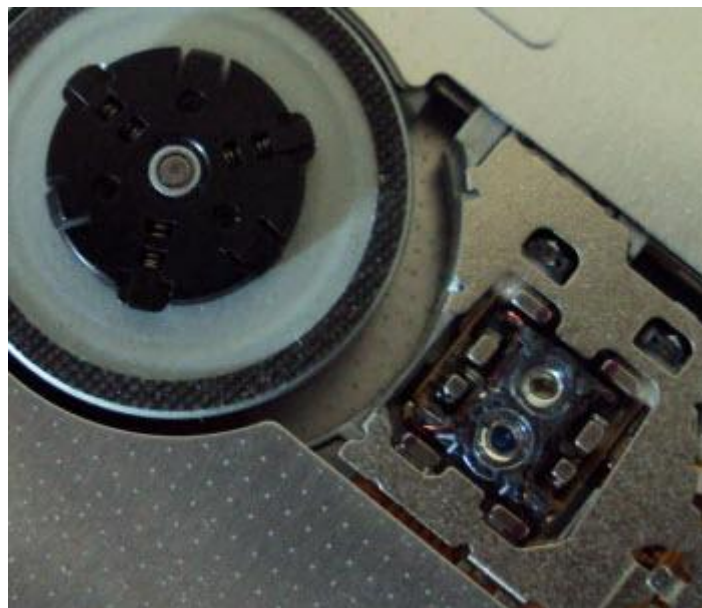
Figure 3.3 An optical disc drive with separate lenses for CD/DVD and for Blu-ray discs

In addition to the lens, an optical drive has a rotational mechanism to spin the disc. Optical drives were originally designed to work at a constant linear velocity (CLV) - this means that the disc spins at varying speeds depending on where the laser beam is reading,

This constant speed is very important for music CDs and movie discs, since we want to listen to music or watch a movie at the regular speed. For other applications, however, such as reading or writing other types of data, working at

this speed is not needed. Modern optical drives can often spin much faster, which results in higher transfer speeds.

An optical drive also needs a loading mechanism. Two general types are in use:

1. A tray-loading mechanism, where the disc is placed onto a motorized tray, which moves in and out of the computer case.
2. A slot-loading mechanism, where the disc is slid into a slot and motorized rollers are used to move the disc in and out.

Tray-loading mechanisms for optical drives in desktop computers tend to be rather bulky. (Figure 3.4)

Figure 3.4 Typically tray-loading optical drive for desktop computers

For laptops, the tray-loading mechanism is much smaller. (Figure 3.5)

Figure 3.5 Typical tray-loading optical
drive for laptop computers

3.8.1 Compact Disk (CD)

Compact Disk is an optical memory device. It is a non-erasable disk that stores the digitized audio information. The standard system uses 12cm disks and they can record more than 60 minutes of playing time without any interruption.

CD-ROM – CD-ROM is a read only disk. It is used to hold prerecorded text, graphics and sound. A CD-ROM disk can hold up to 650MB of data.

Mini CD – CD's are available in a smaller 8cm diameter that can store upto 185 MB of data. A regular CD-ROM drive and CD-writers can be used to read and write on the mini-CD.

CD-R 74 / 80 / 90 / 99 - In CD-Rs it may found that one of the above numbers after the CD-R acronym. These numbers refer to the length of uncompressed audio that can be written to the disc in minutes. 74 minutes is the industry standard which can contain about 630MB of data. 80 minutes CD-R can hold about 703 MB of data, a 90 minute about 790 MB, and 870 MB for a 99 minute CD-R.

One popular form of recordable CD is *PhotoCD,* a standard developed by Kodak for storing digitized photographic images on a CD. Many film-developing stores have PhotoCD drives that can store photos and put them onto a CD.

CD-RW – Compact Disk - ReWritable also called as erasable optical disk allow users to record and erase data so that the disk can be used over and over again.

3.8.2 Digital Versatile Disk (DVD)

DVD – Digital Versatile Disk (or Digital Video Disk) – It is a High density CD-ROM disks, which are read by laser and which have both write-once and rewritable capabilities. There are many variations on the DVD theme.

There are two physical sizes: 12 cm (4.7 inches) and 8 cm (3.1 inches), both of 1.2 mm thick. These are the same form factors as CD. A DVD disc can be single-sided or double-sided. Each side can have one or two layers of data. This can be able to store 4.7 – 17 GB.

DVD-R – DVD disks that allow one time recording by the consumer. Two types of reusable disks are DVD-RW (DVD rewritable) and DVD-RAM (DVD Random Access Memory), both of which can be recorded and erased more than once. There are two formats namely –R, +R.

There are small technical differences in that the +R format allows for features such as a faster disc eject compared to -R or additional menu functions, but these differences are negligible in the wider scope of things. Both types of disc offer 4.7GB of storage.

M-Disc DVD+R - The quite unique M-Disc DVD, made out of a rock like data layer, the disc is designed to last for 1000 years. It has even been military tested to withstand the very toughest conditions. The M-Disc has been design to be read in standard DVD drives. An M-Disc cannot be written in a standard DVD burner, but requires a dedicated writing drive.

DVD-RW and DVD+RW (DVD-5)

The RW stands for Re-Writable, and it works in the same way as a CD-RW. The user can store data to the disc one time or over a period of time, and can delete data and replace it with new content as required. A DVD-RW has a thicker coating so it can withstand more wear and tear, and is rated for up to 1000 write operations. The RW format has a storage capacity of 4.7GB.

DVD-DL (DVD-9)

The DL stands for Dual Layer. This type of DVD disc has almost double the storage capacity of a DVD-5 format, with a total space availability of

8.5GB. Physically a DVD-9 disc has the same dimensions as a CD or DVD-5, but will have a thicker protective lacquer coating to protect the data side.

Some other formats are DVD-10 (12cm, DS/SL): 8.75 GB, DVD-18 (12cm, DS/DL): 15.90 GB- SS/DS means single-/double-sided, SL/DL means single-/dual-layer. Table 3.4 shows the storage capacity of various optical medium.

Table 3.4 Storage capacity of various optical medium.

Device	Type	Size
• CD-ROM • CD-R • CD-RW	• Read Only • Write once then Read only • re-Writable	650 - 900 MB
• DVD-ROM • DVD-R • DVD-RW • DVD-RAM	• Read Only • Write once then Read only • re-Writable • re-Writable	4.7 - 9.4 GB
• Blu-ray (BD) disc • HD DVD (obsolete)	Re-Writable and Read Only versions available. Uses a blue laser, that is able to recognize smaller pits and lands, which allows for the pits and lands to be more closely packed, and so store more data	25 - 128 GB

3.9 Magneto-optical (M-O) Disk

A storage method used in rewritable optical disk drives that combine magnetic and optical recording techniques. The disk is coated with film that initially is uniformly magnetized. A laser beam is used to demagnetize a small spot on the film by heating it above a critical temperature and a local magnetic field determines the direction in which the spot is magnetized when it cools. To read the information, the disk is scanned by polarized light from a low-power laser. The plane of polarization of the light reflected from a magnetized surface is rotated according to the direction of the magnetic field. This rotation, though

small, can be detected and the original binary signal can be reproduced. In early M-O disk drives data had to be erased separately before it could be rewritten, but direct rewriting is now possible.

The M-O technique achieves recording densities similar to those of other optical stores and much higher than has been achieved by magnetic recording.

3.10 **Universal Serial Bus** (**USB**)

The Universal Serial Bus (USB) is technology that allows a person to connect an electronic device to a computer. It is a fast serial bus.

It is mostly used on personal computers. USB is also used on other devices, like the PlayStation 3, PlayStation 4, the Xbox 360, and others. USB connects different devices using a standard interface. Connecting point is called port. (Figure 3.6)

Most people use USB for computer mice, keyboards, scanners, printers, digital cameras, and USB flash drives. There are over six billion USB devices around the world.

The standard was made to improve plug and play devices. This means that a device can be plugged into a free socket, and simply work. The computer will notice the device. The computer sometimes installs special software so that the device can directly be used. The device can be removed after it stops being used. This technology is called "hot swapping". "Hot swapping" means it can be plugged and unplugged while the power is on. The computer does not need to be turned off for people to change the devices.

Now, USB has replaced several older standards. Those include the parallel port, serial port and SCSI. Some special purposes still need those old connection methods.

Figure 3.6 USB Port

Several USB transfer speeds are available to provide faster data transfer rates and support more number of peripheral devices. They are USB 1.X, USB 2.0, also known as hi-speed USB, USB 3.0, also known as Super Speed USB. USB 3.1, also known as Super Speed, USB Type-C.

UNIT IV : DATABASE FUNDAMENTAL

Database Fundamental: Data, Information and knowledge – Database – Logical Data Concepts – Physical Data Concepts – Database Management System – Need, Benefits of DBMS, Components of DBMS, Database Administrator – DBMS Architecture – Database Models.

4.1 Database fundamental: Data, Information and Knowledge

- Data represents unorganized and unprocessed facts.
    - Usually data is static in nature.
    - It can represent a set of discrete facts about events.
    - Data is a prerequisite to information.
    - An organization sometimes has to decide on the nature and volume of data that is required for creating the necessary information.
- Information
    - Information can be considered as an aggregation of data (processed data) which makes decision making easier.
    - Information has usually got some meaning and purpose.
- Knowledge
    - By knowledge we mean *human understanding of a subject matter that has been acquired through proper study and experience*.
    - Knowledge is usually based on learning, thinking, and proper understanding of the problem area.
    - Knowledge is not information and information is not data.
    - Knowledge is derived from information in the same way information is derived from data.
    - We can view it as an understanding of information based on its perceived importance or relevance to a problem area.

o It can be considered as the integration of human perceptive processes that helps them to draw meaningful conclusions.

Figure 4.1 shows the relationship between data, information and knowledge.
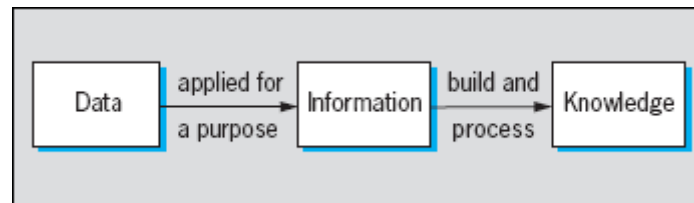
Figure 4.1 Data, Information and Knowledge

- Data is unprocessed facts and figures without any added interpretation or analysis. "The price of crude oil is $80 per barrel."

- Information is data that has been interpreted so that it has meaning for the user. "The price of crude oil has risen from $70 to $80 per barrel" gives meaning to the data and so is said to be information to someone who tracks oil prices.

- Knowledge is a combination of information, experience and insight that may benefit the individual or the organization. "When crude oil prices go up by $10 per barrel, it's likely that petrol prices will rise by 2$ per litre" is knowledge.
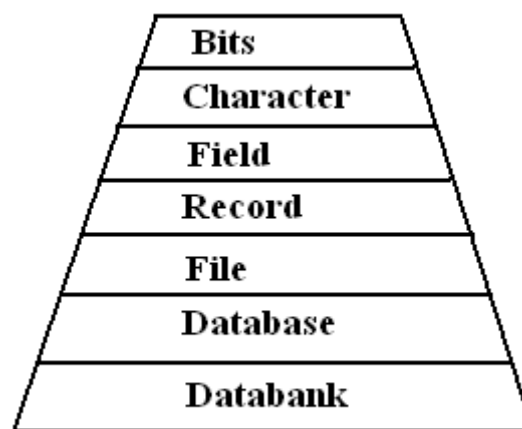
4.2 Hierarchy of data organization

In database system concept, the single data or *data element* is referred to as a *field* and *information* is referred to as a *record*. Collection of related records forms a *file.*

A *database* is a collection of the occurrences of multiple records, containing the relationship between records and fields. One more concept is a *databank,* which refers to a collection of databases.

For example, Register number, Mark1, Mark2, Mark3, Total, Grade are individual fields. A student record is the collection of all the above data fields. Suppose, if there are 40 students in a class, then the collection of 40 records will form the file called student-academic-file. Grouping of related file namely, Student-academic-file and student-personal-file (consists of Father's-name, occupation, address etc.,) , will form a Student database. In a college, one must maintain the databases of all the department student's details, to form a databank. (Figure 4.2)

Figure 4.2 Hierarchy of data organization

A database is a collection of information that is organized so that it can easily be accessed, managed, and updated. In one view, databases can be classified according to types of content: bibliographic, full-text, numeric, and images. Usually the database is organized in the form of table with rows and columns.

4.3 Data Modeling, Key, Normalization

There are a lot of new terms to learn when we begin to study the database.

Entity: An entity is a person, place, thing or concept about which, data can be collected. Examples include Employee, house, Car.

Attributes: Each entity is made up of a number of 'attributes' which represent that entity, also known as property of an entity. An attribute describes the facts, details or characteristics of an entity.

Example:

Entity = Customer. Attributes for the entity customer are Customer ID, First Name, Last Name, Date of Birth, Address, Phone number etc.

Relationship: Link between entities.

Example: Employee and family – One-to-one relationship

Employee and house – One- to-many relationship

Family members and car – Many-to-one relationship

To set up a database, we should work out:

- The entities - The attributes - The entity relationships

This process is called **'data modeling'**

Primary Key

A primary key is a main reference key for the table. As its name suggests, it is the primary key of reference for the table and is used throughout the database to help establish relationships with other tables. The primary key must contain unique values, must never be null and uniquely identify each record in the table.

As an example, a studentid might be a primary key in a student table. Primary keys are mandatory for every table and each record must have a value for its primary key. (Figure 4.3)

Primary Keys

| StudentId	firstName	lastName	courseId
L0002345	Jim	Black	C002
L0001254	James	Harradine	A004
L0002349	Amanda	Holland	C002
L0001198	Simon	McCloud	S042
L0023487	Peter	Murray	P301
L0018453	Anne	Norris	S042

Figure 4.3 Primary Key

Foreign Key

A foreign key is generally a primary key from one table that appears as a field in another where the first table has a relationship to the second. In other words, if we had a table A with a primary key X that linked to a table B where X was a field in B, then X would be a foreign key in B.

An example might be a student table that contains the courseid the student is attending. Another table lists the courses on offer with courseID being the primary key. The 2 tables are linked through courseID and as such courseID would be a foreign key in the student table. (Figure 4.4)
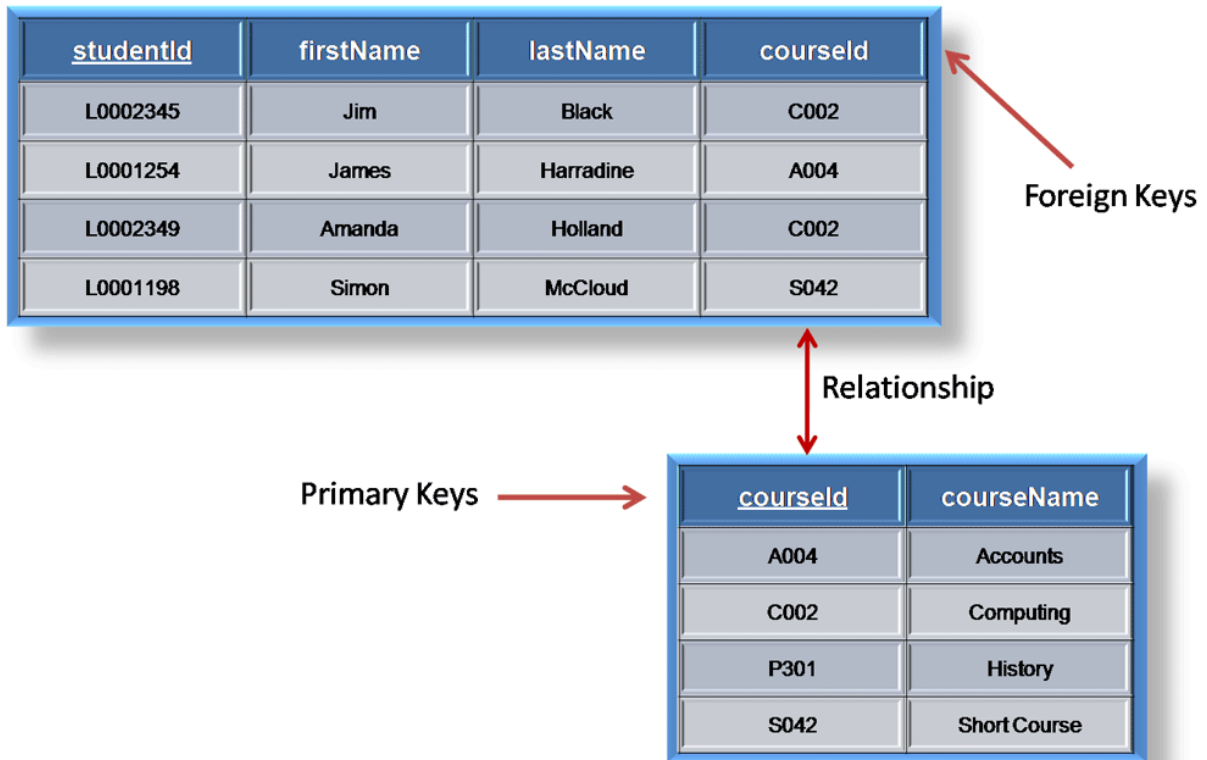
Figure 4.4 Foreign Key

Normalization

- It is the process of organizing data into related table.

- To normalize database, we divide database into tables and establish relationships between the tables.

- This is performed by creating relationships among tables through primary and foreign keys.

- It reduces redundancy. It is done to improve performance of query.

- Steps of normalization:

First Normal form

-Entities of the table must have unique identifier or entity key.

Second Normal Form

- All the attributes of the table must depend on the entity key for that entity.

Third Normal Form

-All attributes that are not part of the key must not depend on any
other non-key attributes.

De-normalization

The process of adding redundant data to get rid of complex join, in order
to optimize database performance. This is done to speed up database access by
moving from higher to lower form of normalization.

4.4 Logical Data Concepts

A logical data model describes the data in as much detail as possible,
without regard to how they will be physical implemented in the database.
Features of a logical data model include:

- Includes all entities and relationships among them.
- All attributes for each entity are specified.
- The primary key for each entity is specified.
- Foreign keys (keys identifying the relationship between different entities)
  are specified.
- Normalization occurs at this level.

The steps for designing the logical data model are as follows:

- Specify primary keys for all entities.
- Find the relationships between different entities.
- Find all attributes for each entity.
- Resolve many-to-many relationships.
- Normalization.

The Figure 4.5 is an example of a logical data model.
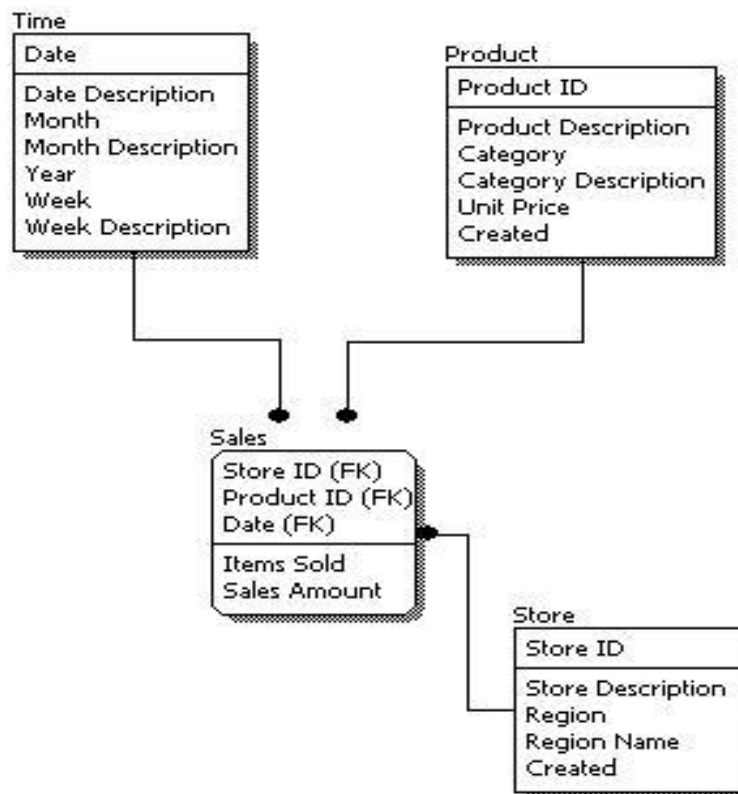
Figure 4. 5   Example of Logical Data Model

4.5 Physical Data Concepts

Physical data model represents how the model will be built in the database. A physical database model shows all table structures, including column name, column data type, column constraints, primary key, foreign key, and relationships between tables. Features of a physical data model include:

- Specification all tables and columns.
- Foreign keys are used to identify relationships between tables.
- Denormalization may occur based on user requirements.
- Physical considerations may cause the physical data model to be quite different from the logical data model.

- Physical data model will be different for different RDBMS. For example, data type for a column may be different between MySQL and SQL Server.

The steps for physical data model design are as follows**:**

- Convert entities into tables.
- Convert relationships into foreign keys.
- Convert attributes into columns.
- Modify the physical data model based on physical constraints / requirements.

The figure 4.6 is an example of a physical data model.
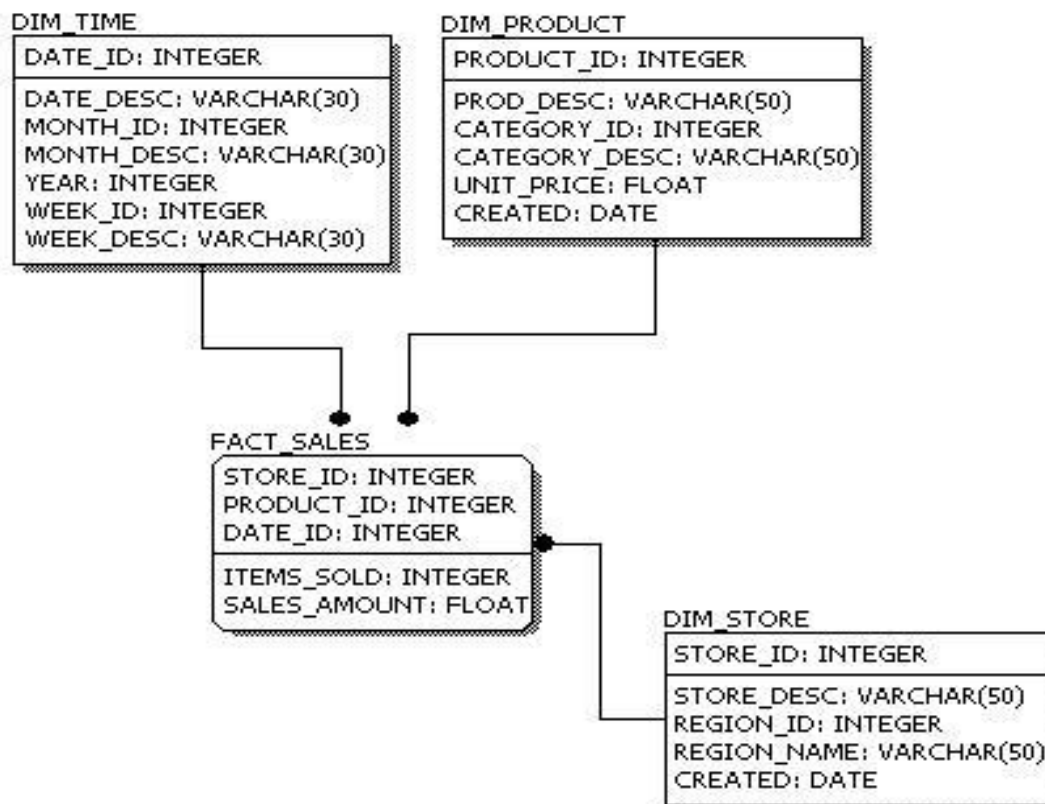
Figure 4.6 Example of Physical Data Model

Comparing the physical data model shown above with the logical data model diagram, we see the main differences between the two:

- Entity names are now table names.

- Attributes are now column names.

- Data type for each column is specified. Data types can be different depending on the actual database being used.

4.6 Data Base Management System (DBMS)

A Database is a collection of related data organized in a way that, data can be easily accessed, managed and updated. A DBMS is software that allows creation, definition and manipulation of database. DBMS is actually a tool used to perform any kind of operation on data in database. DBMS also provides protection and security to database. It maintains data consistency in case of multiple users. Some examples of popular DBMS are MySql, Oracle, Sybase, Microsoft Access and IBM DB2 etc.

4.6.1 Need of DBMS

 * A DBMS is important because it manages data efficiently and allows users to perform multiple tasks with ease.

 * It stores, organizes and manages a large amount of information within a single software application.

 * Use of this system increases efficiency of business operations and reduces overall costs.

 * It is important to businesses and organizations because they provide a highly efficient method for handling multiple types of data.

 * Without DBMS, tasks have to be done manually and take more time.

 * Data can be categorized and structured to suit the needs of the company or organization.

* Multiple users can use the system at the same time in different ways. For example, a company's human resources department uses the database to manage employee records, distribute legal information to employees and create updated hiring reports. A manufacturer might use this type of system to keep track of production, inventory and distribution. In both scenarios, the DBMS operates to create a smoother and more organized working environment.

4.6.2 Benefits of DBMS

The DBMS has a number of benefits compared to traditional computer file processing approach. The following are main benefits

Controlling Data Redundancy:

In traditional computer file processing, each application program has its own data. In this case, the duplicated copies of the same data are created at many places. In DBMS, all the data of an organization is integrated into a single database. The data is recorded at only one place in the database and it is not duplicated. For example, the principal's faculty file and the faculty payroll file contain several items that are identical. When they are converted into database, the data is integrated into a single database so that multiple copies of the same data are reduced to-single copy. By controlling the data redundancy, **we** can save storage space. Similarly, it is useful for retrieving data from database using queries.

Data Consistency:

If a data item appears only once, any update to its value has to be performed only once and the updated value (new value of item) is immediately available to all users.

If the DBMS has reduced redundancy to a minimum level, the database system enforces consistency. It means that when a data item appears more than once in the database and is updated, the DBMS automatically updates each occurrence of a data item in the database.

Data Sharing:

In DBMS, data can be shared by authorized users of the organization. Many users can be authorized to access the same set of information simultaneously. The remote users can also share same data. Similarly, the data of same database can be shared between different application programs.

Data Integration:

In DBMS, data in database is stored in tables. A single database contains multiple tables and relationships can be created between tables (or associated data entities). This makes easy to retrieve and update data.

Integrity Constraints:

Integrity constraints or consistency rules can be applied to database so that the correct data can be entered into database. The constraints may be applied to data item within a single record or they may be applied to relationships between records.

Examples:

The *examples of integrity* constraints are:

(i) 'Issue Date' in a library system cannot be later than the corresponding 'Return Date' of a book.

(ii) Maximum obtained marks in a subject cannot exceed 100.

(iii) Registration number of BCS and MCS students must start with 'BCS' and 'MCS' respectively etc.

There are also some standard constraints that are intrinsic in most of the DBMSs. These are shown in Table 4.1.

Data Security:

Data security is the protection of the database from unauthorized users. Only the authorized persons are allowed to access the database. Some of the users may be allowed to access only a part of database i.e., the data that is related to them or related to their department. Mostly, the Data Base

Administrator (DBA) or head of a department can access all the data in the database. Some users may be permitted only to retrieve data, whereas others are allowed to retrieve as well as to update data. The database access is controlled by the DBA. He creates the accounts of users and gives rights to access the database. Typically, users or group of users are given usernames protected by passwords.

Constraint Name	Description
PRIMARY KEY	Designates a column or combination of columns as Primary Key and therefore, values of columns cannot be repeated or left blank.
FOREIGN KEY	Relates one table with another table.
UNIQUE	Specifies that values of a column or combination of columns cannot be repeated.
NOT NULL	Specifies that a column cannot contain empty values.
CHECK	Specifies a condition which each row of a table must satisfy.

Table 4.1 Standard Constraints

Data Atomicity:

A transaction in commercial databases is referred to as atomic unit of work. For example, when we purchase something from a point of sale (POS) terminal, a number of tasks are performed such as;

- Company stock is updated.

- Amount is added in company's account.

- Sales person's commission increases etc.

All these tasks collectively are called an atomic unit of work or transaction. These tasks must be completed in all; otherwise partially completed tasks are rolled back. Thus through DBMS, it is ensured that only consistent data exists within the database.

Database Access Language:

Most of the DBMSs provide SQL as standard database access language. It is used to access data from multiple tables of a database.

Development of Application:

The cost and time for developing new applications is also reduced. The DBMS provides tools that can be used to develop application programs. For example, some wizards are available to generate Forms and Reports. Stored procedures (stored on server side) also reduce the size of application programs.

Creating Forms:

Form is very important object of DBMS. A Form provides very easy way (user-friendly interface) to enter data into database, edit data, and display data from database. The non-technical users can also perform various operations on databases through Forms without going into the technical details of a database. The Forms can be created very easily and quickly in DBMS, Once a Form is created, it can be used many times and it can be modified very easily. The created Forms are also saved along with database and behave like a software component.

Report Writers:

Most of the DBMSs provide the report writer tools used to create reports. The users can create reports very easily and quickly. Once a report is created, it can be used many times and it can be modified very easily. The created reports are also saved along with database and behave like a software component.

Control over Concurrency:

In a computer file-based system, if two users are allowed to access data simultaneously, it is possible that they will interfere with each other. For example, if both users attempt to perform update operation on the same record, then one may overwrite the values recorded by the other. Most DBMSs have sub-systems to control the concurrency so that transactions are always recorded" with accuracy.

Backup and Recovery Procedures:

In a computer file-based system, the user creates the backup of data regularly to protect the valuable data from damaging due to failures to the computer system or application program. It is a time consuming method, if volume of data is large. Most of the DBMSs provide the 'backup and recovery' sub-systems that automatically create the backup of data and restore data if required. For example, if the computer system fails in the middle (or end) of an update operation of the program, the recovery sub-system is responsible for making sure that the database is restored to the state it was in before the program started executing.

Data Independence:

The separation of data structure of database from the application program that is used to access data from database is called data independence. In DBMS, database and application programs are separated from each other. The DBMS sits in between them. We can easily change the structure of database without modifying the application program. For example we can modify the size or data type of a data items (fields of a database table).

On the other hand, in computer file-based system, the structure of data items is built into the individual application programs. Thus the data is dependent on the data file and vice versa.

Advanced Capabilities:

DBMS also provides advance capabilities for online access and reporting of data through Internet. Today, most of the database systems are online. The database technology is used in conjunction with Internet technology to access data on the web servers.

4.7 Components of DBMS

A DBMS consists of several components. Each component plays very important role in the database management system environment. The major components of DBMS are (Figure 4.7)
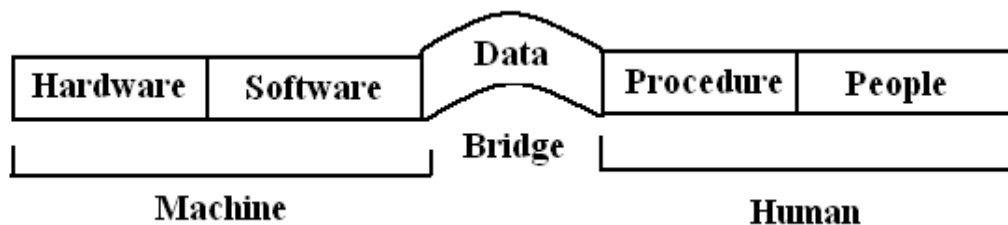
Software

Hardware

Data

Procedures

Database Access Language

Figure 4.7 Components of DBMS environment

Software

The main component of a DBMS is the software. It is the set of programs used to handle the database and to control and manage the overall computerized database

DBMS software itself, is the most important software component in the overall system

Operating system, network software used to share the data of database among multiple users.

Application programs developed in programming languages such as C++, Visual Basic that are used to access database in DBMS. Each program contains statements that request the DBMS to perform operation on database. The operations may include retrieving, updating, deleting data etc.

Hardware

Hardware consists of a set of physical electronic devices such as computers (together with associated I/O devices like disk drives), storage devices, I/O channels, electromechanical devices that make interface between computers and the real world systems etc, and so on. It is impossible to implement the DBMS without the hardware devices. In a network, a powerful computer with high data processing speed and a storage device with large storage capacity is required as database server.

Data

Data is the most important component of the DBMS. The main purpose of DBMS is to process the data. In DBMS, databases are defined, constructed and then data is stored, updated and retrieved to and from the databases. The database contains both the actual (or operational) data and the metadata (data about data or description about data).

Procedures

Procedures refer to the instructions and rules that help to design the database and to use the DBMS. The users that operate and manage the DBMS require documented procedures on hot use or run the database management system. These may include.

Procedure to install the new DBMS.

To log on to the DBMS.

To use the DBMS or application program.

To make backup copies of database.

To change the structure of database.

To generate the reports of data retrieved from database.

Database Access Language

The database access language is used to access the data to and from the database. The users use the database access language to enter new data, change the existing data in database and to retrieve required data from databases. The user writes a set of appropriate commands in a database access language and submits these to the DBMS. The DBMS translates the user commands and sends it to a specific part of the DBMS called the Database Jet Engine. The database engine generates a set of results according to the commands submitted by user, converts these into a user readable form called an Inquiry Report and then displays them on the screen. The administrators may also use the database access language to create and maintain the databases.

The most popular database access language is SQL (Structured Query Language).

Users

The users are the people who manage the databases and perform different operations on the databases in the database system. There are three kinds of people who play different roles in database system

Application Programmers

Database Administrators

End-Users

Application Programmers

The people who write application programs in programming languages (such as Visual Basic, Java, or C++) to interact with databases are called Application Programmer.

Database Administrators

A person who is responsible for managing the overall database management system is called database administrator or simply DBA.

End-Users

The end-users are the people who interact with database management system to perform different operations on database such as retrieving, updating, inserting, deleting data etc.

4.8 Data Base Administrator (DBA)

A database administrator (DBA) directs or performs all activities related to maintaining all databases required for development, testing, training and production usage, managing share resources used amongst applications. Some other responsibilities are

* Administrates all database objects, including tables, views, indexes, stored procedures, functions, packages etc.

* Enforces and maintains database constraints to ensure integrity of the database

* Installation and configuration of DBMS server software and related products

* Upgrading and patching/hot-fixing of DBMS server software and related products

* Assists with impact analysis of any changes made to the database objects

* Migrating database to another server

* Evaluate DBMS features and DBMS related products

* Ensure that the site is running the products that are most appropriate

* Ensure that any new product usage or release upgrade takes place with minimal impact

*Establish and maintain sound backup and recovery policies and procedures.

* Take care of the Database design and implementation.

* Implement and maintain database security (create and maintain logins, users and roles, assign privileges).

* Performance tuning and health monitoring on DBMS, OS and application.

* Setting up Server Level, Database Level and Operating System Alerts

* Implementation of robust maintenance plan

* Setup and maintain documentation and standards.

*Perform reviews on the design and code frequently to ensure the standards are being adhered to

*Plan growth and changes (capacity planning).

* Do general technical troubleshooting and give consultation to development teams.

* Troubleshooting on DBMS and Operating System performance issue

* Documentation of any implementation and changes (database changes, reference data changes and application User Interface changes etc)

* Be able to provide a strategic database direction for the organization.
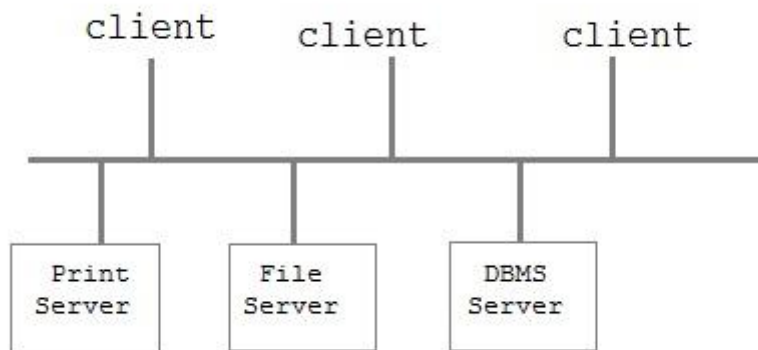
4.9 DBMS Architecture

Database architecture is logically divided into two types.

Logical two-tier Client / Server architecture

Logical three-tier Client / Server architecture
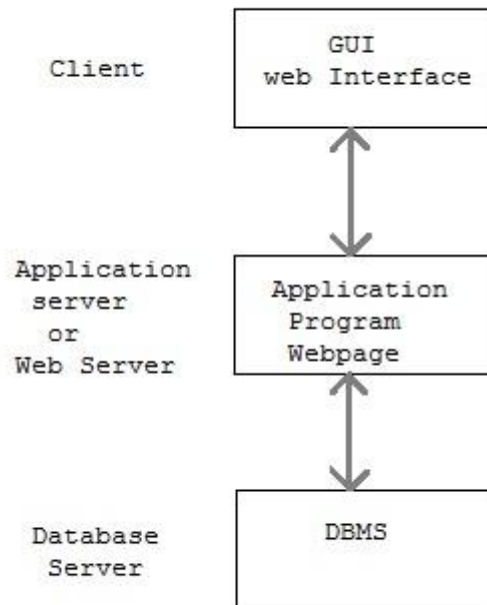
Two-tier Client / Server architecture

Figure 4.8 Two-tier Client / Server architecture

Two-tier Client / Server architecture is used for User Interface program and Application Programs that runs on client side (Figure 4.8). An interface called ODBC (Open Database Connectivity) provides an API (Application Programming Interface) that allows client side program to call the DBMS. Most DBMS vendors provide ODBC drivers. A client program may connect to several DBMS's. In this architecture some variation of client is also possible for example in some DBMS's more functionality is transferred to the client including data dictionary, optimization etc. Such clients are called Data server.

Figure 4.9 Three-tier Client / Server architecture

Three-tier Client / Server database architecture is commonly used architecture for web applications. Intermediate layer called Application server or Web Server stores the web connectivity software and the business logic (constraints) part of application used to access the right amount of data from the database server (Figure 4.9). This layer acts like medium for sending partially processed data between the database server and the client.

4.10 Data Base Models

A database model is a logical structure of a database and fundamentally determines the manner in which the data can be stored, organized, and manipulated. Common logical data models for databases include:

- Flat model
- Hierarchical model

- Network model

- Relational model

- Object model

- Entity–relationship model

Flat model

The flat model is the earliest, simplest data model. It simply lists all the data in a single table, consisting of columns and rows. In order to access or manipulate the data, the computer has to read the entire flat file into memory, which makes this model inefficient for all but the smallest data sets Figure 4.10.

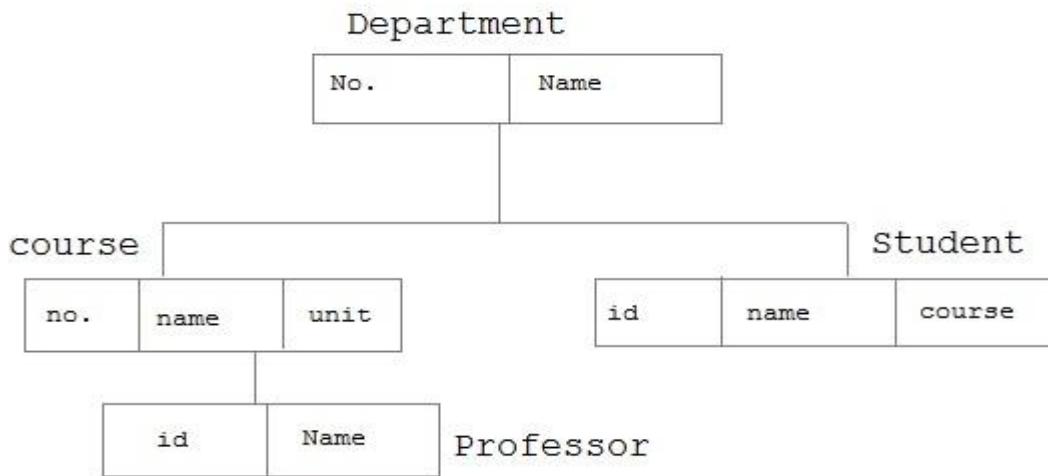|	Route No.	Miles	Activity
Record 1	I-95	12	Overlay
Record 2	I-495	05	Patching
Record 3	SR-301	33	Crack seal

Figure 4.10 Flat File Model

Hierarchical model

In this model each entity has only one parent but can have several children. At the top of hierarchy there is only one entity which is called root.
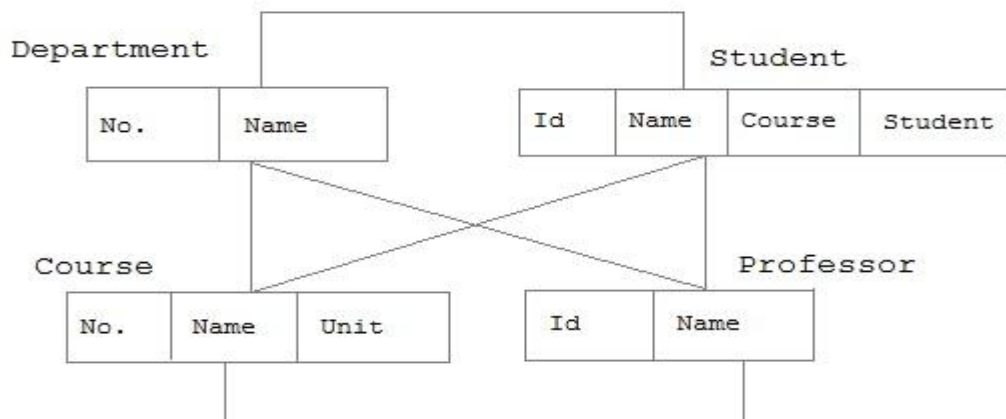
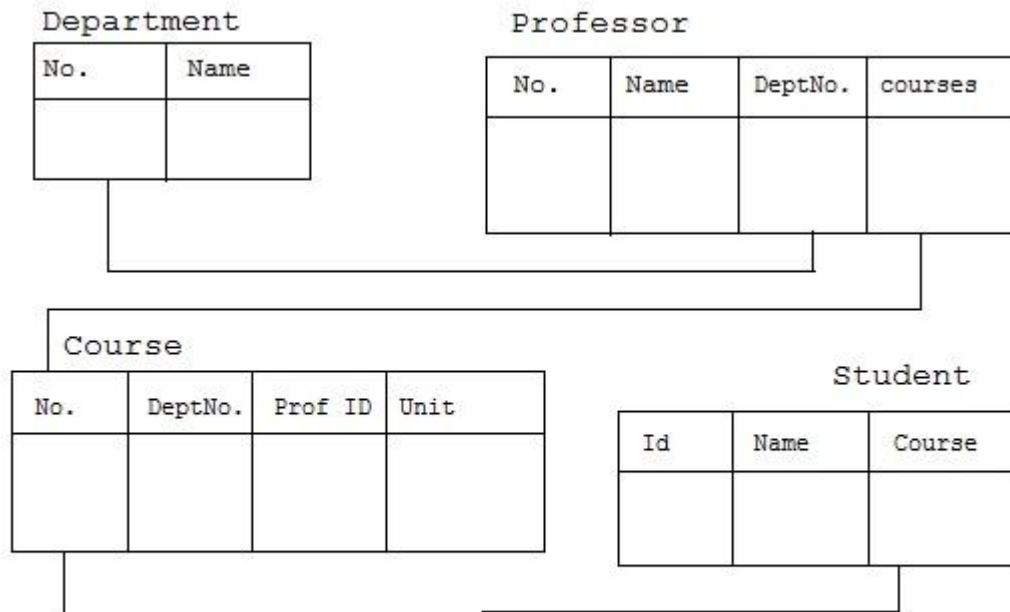**Figure 4.11 Hierarchical Model**

Network Model

In the network model, entities are organized in a graph, in which some entities can be accessed through several paths. (Figure 4.12)

Figure 4.12 Network Model

Relational Model

In this model, data is organized in two-dimensional tables called relations. The tables or relation are related to each other. (Figure 4.13)

Figure 4.13 Relational Model

Object model

This model defines a database as a collection of objects, or reusable software elements, with associated features and methods. (Figure 4.14)There are several kinds of object-oriented databases:

A *multimedia database* incorporates media, such as images, that could not be stored in a relational database.

A *hypertext database* allows any object to link to any other object. It's useful for organizing lots of disparate data, but it's not ideal for numerical analysis.

The object-oriented database model is the best known post-relational database model, since it incorporates tables, but isn't limited to tables. Such models are also known as hybrid database models.
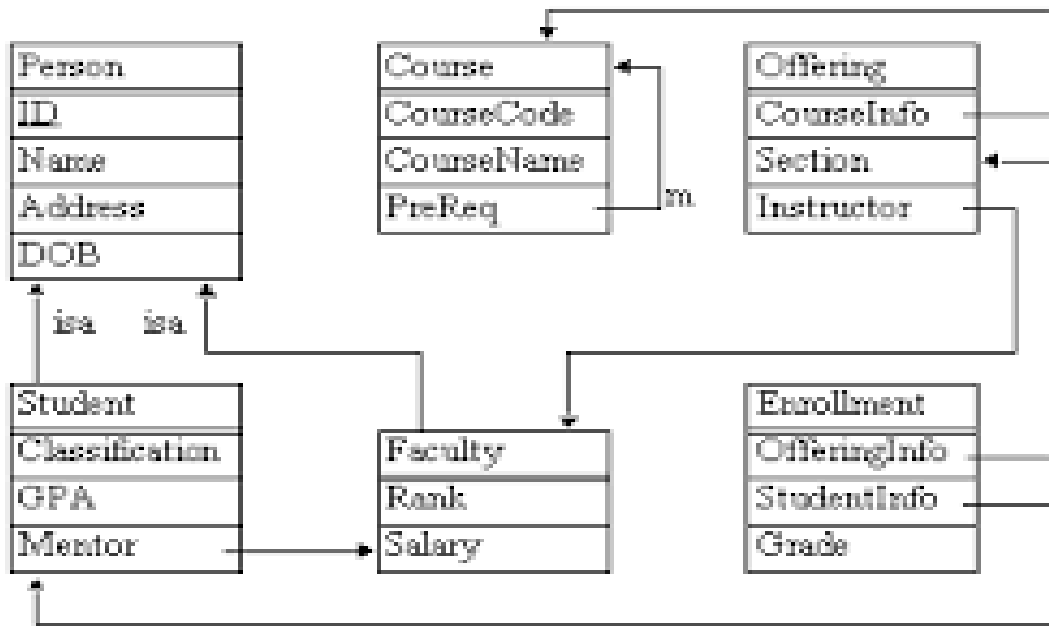
Figure 4.14 Object model

Entity-relationship model

This model captures the relationships between real-world entities much like the network model, but it isn't as directly tied to the physical structure of the database. Instead, it's often used for designing a database conceptually (Figure 4.15).

Here, the people, places, and things about which data points are stored are referred to as entities, each of which has certain attributes that together make up their domain. The cardinality, or relationships between entities, is mapped as well.
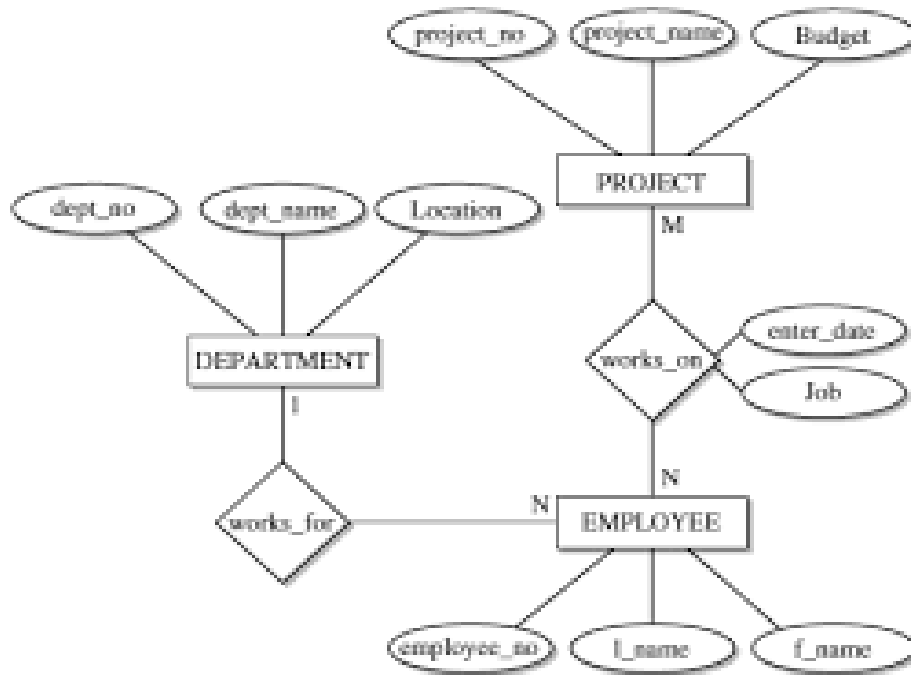
Figure 4.15 Entity-relationship model

UNIT V : INTERNET BASICS

Internet Basics: Basic Internet Terms – Internet Addressing – Internet Applications – e-mail, WWW, File Transfer Protocol, Telnet, Internet Relay Chat, Gopher, Chatting, Commerce through Internet , Groups, News, Social Networking, Blog, Videoconference, Online services – E-mail address structure – Sending and Receiving E-mail – Search Engines – Internet and Viruses.

5.1 Internet Introduction

The Internet is a worldwide, publicly accessible series of interconnected computer networks that transmit data by packet switching using the standard Internet Protocol.

The U.S. department of defense established a program called Advanced Research Projects Agency Network (ARPANET) in 1969, to provide a secure communications network for defense related research. After undergoing several transformations, the internet has become what it is today.

Vinton Cerf and Robert Kahn developed the first description of the TCP protocols in 1973. The term "Internet" was first used in 1974 to describe a single global TCP/IP network detailed in the first full specification of TCP written by Cerf and his colleagues. The first TCP/IP-Wide Area Network was created on January 1, 1983 when all hosts on the ARPANET were switched over from the older protocols to TCP/IP.

5.2 Basic Internet Terms

Computers are connected together to form a network. The communication among them is originally attained by the network operating system. As the number of computers connected to the network increased and the computers running on the different operating system needed to communicate with each other, the necessity of standardization in the communication arised.

Standardization in communication is achieved in the areas of,

Communication protocol, addressing scheme and service providers

Communication protocol

Information is transmitted over the internet in the form of packets. When a piece of information is sent to another user on the internet, the information is broken down into packets of fixed size and sent. In the destination machine, the packets are reassembled as information.

A protocol is simply an agreed set of standards that allows computers to communicate one another. The name of the protocol may change according to the internet service. The protocol (rules) used to transfer packets across the internet is called Internet Protocol (IP) and the protocol that breaks the information into packets and reassembles it, is Transmission Control Protocol (TCP). Hence the standard for communication on the internet is called TCP/IP.

TCP/IP transfers information in small units called packets. Each packet includes,

- The identification of the source computer from where the data arises

- The destination computer to which it is addressed

- The data itself

- The error checking mechanism to confirm that the individual packet was accurately and completely sent and received.

In order to speed up the data transmission, the TCP/IP can break the large file into multiple packets, each sent over the different paths in the network. These packets are then reassembled at the other end into one file, and saved in

the destination computer. The TCP/IP determines addressing mechanisms and the routing of packets.

Addressing scheme

Some form of identification is needed, to send information between computers. The web address which provides this identification is called Uniform Resource Locator (URL). It is composed of a specific string of characters, consisting of,

- a protocol   - a host name   - a port

- a directory   - a file name

Each page on the web has its own unique URL. Even some elements on a page have separate URL's. For example a graphic in a page have a URL, which is slightly different from the page URL.

Example:

http:// www.netscape.com/new/index.html#download

http – specifies the protocol, in this case hyper text transfer protocol (the language that www uses to tie the pages together). The protocol is followed by a colon.

www.netscape.com – part of the address called a domain name, is preceded by two slashes. This is the name assigned to a computer, located somewhere in the world that is permanently connected to the web. Here www stands for World Wide Web; netscape is the name of the organization hosting the site; .com is the identification of the type of organization hosting the site (called top most domain).

/new/index.html is the path to a file, meaning that the document index.html is stored in a folder named /new.

#download is called a fragment ID. This is the name of a specific part of the document.

Internet Service Providers (ISP)

To access the internet one need an ISP. The ISP is connected to the internet backbone, which is the permanent cabling of the internet. This backbone may consist of copper wire, fiber optic cable, microwave and even satellite connections between any two points. The Internet Service Providers govern the internet. They are the authorized people to provide the internet service.

In our country BSNL (Bharat Sanchar Nigam Limited) is one of the service providers. Some private concerns who provide this service are Aircel , Reliance Communications, Tata DoCoMo , and Tata Indicom. Private Service providers have to obtain a national wide ISP license from Department Of Telecommunication (DOT).

5.3 Internet Addressing

An Internet address uniquely identifies a node on the Internet. Internet address may also refer to the name or IP of a Web site (URL). The term Internet address can also represent someone's e-mail address.

5.3.1 IP Address

IP addressing is a hardware-independent convention which in principle allows every computer attached to the Internet to be given a

unique *logical* address. IP addresses are currently 32-bit binary strings, in *dotted decimal* such as 223.58.1.10

The 32-bit address is broken up into four 8-bit sequences each of which is converted to decimal. So the above address in binary is:

11011111 00111010 00000001 00001010

IP Address Formats

Internet addresses combine – a routing portion, known as the network part – a name portion known as the host part „ How to split an Internet address into the network part and the host part has changed over time. Originally IP addresses were divided into five *classes* as shown below. Classes A, B and C are the most important: the initial bits determine which class an address belongs to, and the classes differ in how much of the address is taken up with the *network address* and how much with the *host address*.

Offsets

0	8	16	24
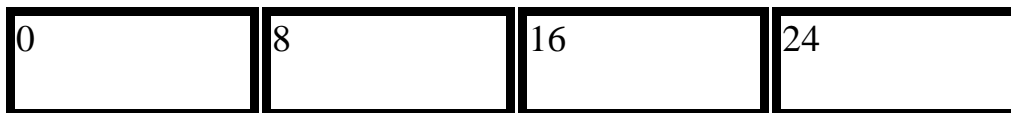
Class A

0 Network	Host

Addresses 1.0.0.0 to 127.255.255.255

Class B

10 Network	Host

Addresses 128.0.0.0 to 191.255.255.255

Class C

110 Network	Host

Addresses 192.0.0.0 to 223.255.255.255

Class D

1110 Multicast address

Addresses 224.0.0.0 to 239.255.255.255

Class E

11110 Reserved for future use

Addresses 240.0.0.0 to 255.255.255.255

The designers of the Internet Protocol defined an IP address as a 32-bit number and this system, known as Internet Protocol Version 4 (IPv4), is still in use. However, because of the growth of the Internet and the predicted depletion of available addresses, a new version of IP (IPv6), using 128 bits for the address has been established. The address size was increased from 32 to 128 bits, thus providing up to 2^{128} (approximately 3.403×10^{38}) addresses. This is deemed sufficient for the foreseeable future.

The Internet Assigned Numbers Authority (IANA) manages the IP address space allocations globally and delegates five Regional Internet Registries (RIRs) to allocate IP address blocks to local Internet registries (Internet service providers) and other entities.

Currently, classless addressing is in use providing arbitrary host and network part lengths.

5.3.2 Domain Name

In general, a domain name represents an Internet Protocol (IP) resource, such as a personal computer used to access the Internet, a server computer hosting a web site, or the web site itself or any other service communicated via the Internet. It is an identification string used in various networking contexts and application-specific naming and addressing purposes. Domain names are formed by the rules and procedures of the Domain Name System (DNS).

Domain names are organized in subordinate levels (sub domains) of the DNS root domain, which is nameless. The first-level set of domain names are the top-level domains (TLDs), including the generic top-level domains (gTLDs), such as

com -commercial (business)   edu – education

net – organizations directly involved in internet operations

org – organization   gov – government

mil –military

and the country code top-level domains (ccTLDs) such as

in- India   us – United States   de - Germany

Each TLD has a separate registry managed by a designated organization under the direction of the Internet Corporation for Assigned Names and Numbers (ICANN).

Below these top-level domains in the DNS hierarchy are the second-level and third-level domain names that are typically open for reservation by end-users who wish to connect local area networks to the Internet, create other publicly accessible Internet resources or run web sites. The registration of these

domain names is usually administered by domain name registrars who sell their services to the public.

A fully qualified domain name (FQDN) is a domain name that is completely specified in the hierarchy of the DNS, having no parts omitted. Most commonly domain names are written in lowercase in technical contexts.

Domain names are used to establish a unique identity. Organizations can choose a domain name that corresponds to their name, helping Internet users to reach them easily.

ICANN identifies the following categories of TLDs:

- Country-code top-level domains (ccTLD) -- Each ccTLD identifies a particular country and is two letter long. The ccTLD for the United States, is *.us*

- Infrastructure top-level domain -- There is only one TLD in this group, ARPA (Address and Routing Parameter Area). The Internet Assigned Numbers Authority (IANA) manages this TLD for the IETF (Internet Engineering Task Force).

- Sponsored top-level domains (sTLD): These are overseen by private organizations.

- Generic top-level domains (gTLD) -- These are the most common and familiar TLDs. Examples include "com" for "commercial" and "edu" for "educational." Most gTLDs are open for registration by anyone, but there is also a subgroup that is more strictly controlled.

For example, http*://whatis.techtarget.com*: com is the top-level domain name; *techtarget* is the second-level domain name; and *whatis* is a subdomain

name. All together, these constitute a fully-qualified domain name (FQDN); the addition of *HTTP://* makes an FQDN a complete URL.

A generic domain is a name that defines a general category, rather than a specific or personal instance. For example, the name of an industry, rather than a company name. Some examples of generic names are *books.com*, *music.com*, and *travel.info*. Companies have created brands based on generic names, and such generic domain names may be valuable.

5.3.4 Uniform Resource Locator (URL)

A Uniform Resource Locator (URL) (commonly informally referred to as a web address is a reference to a web resource that specifies its location on a computer network and a mechanism for retrieving it.

A URL implies the means to access an indicated resource. URLs occur most commonly to reference web pages (http), but are also used for file transfer (ftp), email (mailto), database access (JDBC), and many other applications.

The format combines the pre-existing system of domain names with file path syntax,   where slashes are   used   to   separate directory and file names. Conventions already existed where server names could be pretended to complete file paths, preceded by a double slash (//).

http://www.domainname.com/folder-name/web-page-file-name.htm

5.4 Internet Applications

E-Mail

E-mail stands for electronic email. It is the most popular service provided by the Internet. It provides the fast and efficient way to send and receive messages through Internet. One message can be sent to many persons with a single e-mail. Different types of files can also be sent as attachment through e-mail.

World Wide Web (WWW)

The World Wide Web is commonly known as *web*. It is a network of web servers that stores web pages. The web pages are connected to each other using

hyperlinks. The user can jump from one page to another by clicking the hyperlinks. The web pages are accessed using web browsers. The HTTP (Hyper Text Transfer Protocol) protocol is used for communication between browser and web servers.

File Transfer Protocol

FTP stands for File Transfer Protocol. It is a way for transferring files from one computer to another. The process of transferring a file from a server (or remote computer) to local computer is called *downloading*. Similarly, the process for transferring a file from local computer to the server on the Internet is called *uploading*.

Different FTP client programs are available for uploading and downloading files to and from the server. The most commonly used FTP client programs are WS-FTP and cute FTP.

Telnet

Telnet is an abbreviation for Terminal Network. It is software. It is used to connect to a remote or host computer for accessing information. Through this service, the user can also access information on the Internet.

When a user runs this software on their computer, it provides a prompt on the user's computer screen. The user can access the host computer by giving commands through this prompt. When a command is sent to the host computer, information is accessed from host and displayed on user's computer screen. The user's computer linked to the remote computer will act as a terminal.

Internet Relay Chat

IRC is short for **Internet Relay Chat** and is a popular chat service. IRC enables users to connect to a server using a software program or web service and communicate with each other live. For example, the Computer Hope chat room uses an IRC server to allow its users to talk and ask computer questions live.

To connect and chat with other IRC users, we must either have an **IRC client** or a web interface that connects us to IRC servers. There are numerous software IRC clients that enable users to connect and communicate to other IRC servers.

On the Web, certain sites like Talk City or IRC networks such as the Undernet provide servers and helps us to download an IRC client to our PC.

We can start a chat group (called a channel) or join an existing one. There is a protocol for discovering existing chat groups and their members. Depending on the type of network, nicknames can be reserved (registered) or just used during the session.

Popular ongoing IRC channels are #hottub and #riskybus. A number of channels are set up and conducted in foreign languages. The most common IRC networks are IRCnet (mostly European), EFnet (mostly North American), Undernet, and Dalnet. Popular IRC clients include mIRC for Windows, IRCle for Mac OS, and irc2 (the original client) for UNIX-base operating systems.

Gopher

A gopher is a menu-based information retrieval system. It is used for retrieving files and programs on the Internet. Gopher allows access to files found on FTP servers and web servers.

Chatting

Internet provides the facility to its users to chat with people online all over the world. Different programs like MSN Messenger, Yahoo Messenger, AOL etc. are available for chatting on the Internet. Chat rooms allow the users to participate in a chat on the Internet. Chat rooms are locations on the web that provide facilities to the users to chat with each other online over the Internet. They send messages by typing with keyboard and receive messages from other end instantly. Some chat rooms support voice chats and video chats where people talk with each other and see them also.

Commerce through Internet

The smooth operation of a business depends on the efficient exchange of information between different parts of business, between business and outside world.

E-commerce is nothing but business through Internet. This is a dynamic set of technologies, applications and business process that links enterprises, consumers and business communities through electronic interchange. Simply the process of achieving business priorities through smart and efficient use of Internet technology is called as E-commerce.

E-commerce is a multifaceted concept which includes the following features.

- Conduct of business via communication networks

- Management through electronic methods such as electronic data interchange and automated data collection system

- Electronic transfer of funds between buyers and suppliers

- Creates interdepencies between the company and supplier & customers related to the company

- Internet based marketing

- Intranet based information (A private network designed for information management within an organization)

- Extranet based information (An extension of organization's intranet by using WWW technology to facilitate communication with the corporation's suppliers and customers.)

Various services integrated by E-commerce are

- *Communications services* which supports the data transfer from the originator to the recipient

- *Data management services* which defines the exchange format of the information

- *Security services* which authenticate the source of information, guarantee the integrity of the information received, prevent from unauthorized users and document the acknowledgement from the recipient.

IT Implementation Trends in business

Information Technology (IT) has certainly changed the traditional business methodology. With the advent of IT business process management and automation was achieved with full potential thereby reducing human errors. The following are some of the IT implementation trends in business.

To improve customer service - Business Process Management (*BPM*) and automation, Customer Relationship Management (**CRM),** Business Intelligence *(BI)*, Decision Support System (*DSS*).

IT enabled BPM has and more scientific approach that has improved, the turn-around time, with high processing speeds. CRM, BI and DSS are other IT enabled business services.

CRM is to service customers, retain existing customers, integrate multiple communication channels (mobile, e-mail, web etc) so that their customer's support requests reach a common, central touch point. An e-CRM includes Sales Force Automation (SFA), marketing applications and customer service and support applications.

BI and DSS are the need for executives, to know about the customer, business solutions, create right type of report for faster decision making.

IT/ITES – Business solutions are easily achieved with Information *Technology/ IT Enabled Services that applies* complex data warehousing tools to store information, data mining technologies to retrieve information. Online training program for employees, video and web based conferencing, call centre service, consolidated datacenter are some other ITES in business.

SCM, PLM, KM are some of the ITES in business.

SCM – Supply Chain Management – SCM facilitates the companies to examine their manufacturing, distribution, among multiple suppliers and business partners.

For manufacturing customized software packages such as basic accounting package for the accounts department, inventory management packages for the store and separate software for production planning etc are used. Most industries are having common **ERP** (Enterprise Resource Planning) which serves as central part of business strategy and planning.

Advantages of using SCM are greater visibility into work-in-process, cycle times, productivity, equipment utilization, inventory management, delivery effectiveness etc.

PLM – Project Lifecycle Management – The lifecycle spans from design of product to manufacturing it without defects. The important stages of PLM are

 - Designing   - Authoring   - Manufacturing   - Product testing

This has to be achieved by reduced production time and cost without compromising quality. At each stage customized software packages are used for improving efficiency.

Typically, stakeholders of manufacturing companies are distributed across geographies. However, the ability to come up with frequent design changes with reduced production cycle time is very important for manufacturing companies to stay ahead of curve. PLM provides solutions to this, which include 3D design platforms, product design authoring, manufacturing process creation and management, product testing etc.

KM – Knowledge Management – KM uses web technologies to distribute their brand name, goodwill and ideas to employees and customers through website, presentations, documentations etc. The company's trademarks, patents,

potential customers, annual reports, sales figures will documented and are analyzed for future presentations using this knowledge.

All organizations have two kinds of knowledge – explicit and tacit. Explicit knowledge can be documented, such as the company's patents, trademarks, customer lists and sales figures. Tacit knowledge is what the employees gain over the years, after spending significant time in their respective fields. KM is the ability to capture and utilize this tacit knowledge that is identify, capture, share and manage the tacit knowledge in the minds of the employees.

The easiest way of doing this is, by using web technologies. The gained knowledge can be distributed to employees through website, chat sessions, presentations or streaming audio/video.

Business – Intelligence, Analytics and Reengineering

ITES has extended its services in BI, BA and BPR. BI uses data mining and web mining technologies. BA enables future predictions based on decision support system. BPR is Business Process Reengineering that redesigns the workflow for improving business efficiency.

BI/BA tools – Business Intelligence and Business Analytics. BI systems can help organizations mine real time data for valuable and accurate insights. BA is no longer a technology that only large enterprises can offer. Today, even SME (Small to Medium scale Enterprises) can also gain from this technology, because of high benefit. BI & BA both are meant to help the planning. But BA is more flexible than BI. BI resembles more of typical data base operations, where we query facts, whereas BA allows us to dig at a deeper level and model what-if scenarios as well as optional outcomes. Example: BI – Oracle, MySQL server, BA – SAP, IBM's cognos Insight.

All ITES uses latest technologies like mobile computing and cloud computing. Also social media like whatsapp, facebook, twitter, hike, instagram plays important role in promoting business.

BPR – Business Process Re-engineering – It means redesigning the workflow to make the business more efficient.

Technologies like cloud computing, mobile computing and social media make it possible for aspiring entrepreneurs to try their ideas with minimal investment.

E-commerce and its benefits

Placeless office/ E-marketplace – Reduces market place .Many organizations are turning to e-commerce as a viable option for marketing and selling the products and services. This can be done by using an e-marketplace such as e-Bay or the company's own website is used.

Paperless office – Reduces paper usage. Paper and printing are significant cost to every organization. To reduce this, most of the organizations are having Internet/Intranet based centralized application that has digitized workflow to move towards paperless office.

Currency less transactions /Online payment systems - Reduces currency usage. Trend of shopping online are on a steady rise. More and more users are now using online shopping to save time and efforts. So, for an organization it is vital for proper payment system. For payment, several card mode systems, e-cheque, payment via ATM, online banking or e-banking, mobile banking etc are available. The most important aspects related to any payment system are reliability, secure, ease of use and cost effectiveness.

News Groups

Newsgroups are Internet-based discussion group, similar to a bulletin board system (BBS), where people post messages concerning whatever topic around which the group is organized.

Newsgroups are typically found on USENET, a network of discussion groups where millions of users read postings, or articles, using software called a newsreader. Users can then make comments and ask questions in response to the postings. Thousands of newsgroups exist, covering a wide range of topics. Newsgroups typically fall into a few basic categories. There are newsgroups that have to be pre-approved and cover a specific topic, and there are alternative newsgroups that can be created by anyone and cover any topic of their choosing. Newsgroups are also categorized as either moderated—posts have to be approved—or immoderate.

Newsgroup names are made up of parts, separated by dots, that indicate the topics covered in the newsgroup. The parts of the name go from least specific to most specific. For example, the group name comp.sys.apple2 discusses computers (a general topic), systems (a computer topic), and Apple II (a specific computer system). Some common group subjects are comp (computers), rec (recreation), and sci (science). Other newsgroups follow a similar order, including alternative groups, which start with alt.

Social Networking

A social networking service (also social networking site or **SNS**) is a platform to build social networks or social relations among people who share similar interests, activities, backgrounds or real-life connections. A social network service consists of a representation of each user (often a profile), his or her social links, and a variety of additional services such as career services. Social network sites are web-based services that allow individuals to create a public profile, create a list of users with whom to share connections, and view and cross the connections within the system. Most social network services are web-based and provide means for users to interact over the Internet.

The main types of social networking services are those that contain category places (such as former school year or classmates), means to connect

with friends (usually with self-description pages), and a recommendation system linked to trust. Popular methods are Face book, Google+, LinkedIn, and Instagram.

Blog

A blog (short for weblog) is a personal online journal that is frequently updated and intended for general public consumption. Blogs are defined by their format: a series of entries posted to a single page in reverse-chronological order. Blogs generally represent the personality of the author or reflect the purpose of the Web site that hosts the blog.

The author of a blog is often referred to as a blogger. People maintained blogs long before the term was coined, but the trend gained momentum with the introduction of automated published systems, most notably Blogger at blogger.com.

Videoconference

A videoconference is a live connection between people in different locations for the purpose of communication, usually involving audio and often text as well as video. At its simplest, videoconferencing provides transmission of static images and text between two locations. At its most sophisticated, it provides transmission of full-motion video images and high-quality audio between multiple locations.

The tangible benefits for businesses using videoconferencing include lower travel costs and profits gained from offering videoconferencing as an aspect of customer service. The intangible benefits include the facilitation of group work among geographically distant teammates and a stronger sense of community among business contacts, both within and between companies. Videoconferencing software is available as standard computer equipment. For example, Microsoft's NetMeeting is included in Windows 2000.

Online services

We can enjoy several benefits with a PC and internet connection. Some of them are

Communication - email, instant messaging, message boards, online conferencing, blogs, micro-blogging e.g. Twitter, newsgroups, chat, e-communities, social networking

Real Time Information - train timetables, news services, traffic reports, weather

Commerce - shopping, banking, auctions

Government - online tax returns, e-voting, applications for services (e.g. university entry, business incorporations, land registration), revenue collection, health services like NHS Direct

Education - VLEs, online learning, school web sites, school league tables, online revision, simulations, online teaching

Business - video conferencing, collaborative working, business networks

Entertainment - multi-user games, radio players, sports, books

Download Services - software, upgrades, music, film

Web Storage - Google docs, Drop Box

Online mapping and route planning - Google maps, AA Route planner, RAC route planner

5.5 E-mail address structure

E-mail allows the user to communicate information, through the internet. An E-mail address identifies a person and the mailbox for the purpose of exchanging messages. Most of the e-mail address contains three basic parts, while others can be more complex.

Account Name

- The first part of the email address shows the name of the account holder, or the sender, followed by the "@ (at)" sign. It is usually the name used to log in to the account and is also called a username.

Domain

- The name of the domain that hosts the account will follow the username and "at" sign. If a person has a Google email account, the address would read "username@gmail.com."

Domain Extension

- The last letters following the domain name and the dot will be an abbreviation of the type of organization the address originates from. A commercial domain uses ".com," government uses ".gov," educational institutions use ".edu," organizations use ".org," and networks use ".net."

Country Code

- If an email sender is from a foreign country, the email address may have two letters at the end indicating which country it is coming from. This code may replace the domain extension or be added after. For example, the United Kingdom is .uk, Ireland is .ie, France is .fr, and Australia is .au.

Sub-Domain

- The level of complexity will depend on the email server. Some organizations and companies may add a sub-domain followed by a dot between the "at" sign and the domain name. Others may add an extra part in the account name.

5.6 Sending and Receiving E-mail

In a desktop based client like outlook, windows live mail etc, the Figure 5.1 below shows how email is sent and received.

The email is composed on the email client (C1).

When we hit send, the email is transferred to the Email server (S1) using the SMTP (Simple Mail Transfer Protocol).
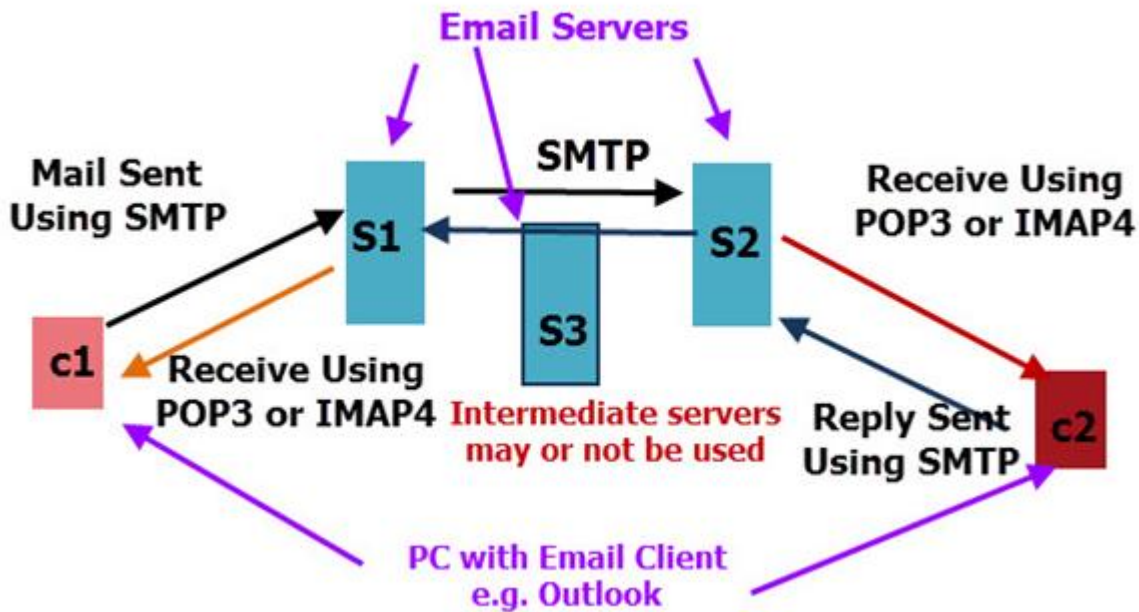
Figure 5.1Sending and receiving E-mail on Internet using desktop E-mail Clients

The email server (S1) belongs to our email provider e.g.Yahoo.

The email server (S1) needs to deliver the email to the destination server (S2).

Message transfer between email servers is done using SMTP.

The destination server (S2) holds the recipients mailbox.

There are two choices that sever S1 can make:

1. It can deliver the email directly to the destination server S2.
2. It can send the Email to an Intermediate server S3.

Direct Delivery – If the sever S2 uses direct delivery then it needs to find the email server responsible for the recipient mailbox. This it does using a DNS lookup and MX records.

Indirect Delivery – If it chooses to use an intermediate server (S3) known as an email forwarder then the intermediate server must then choose either a direct delivery or to forward to another server.

Receiving the Email

The sent email will eventually be placed in a mailbox on the server S2. The

server S2 is the email server of the email provider for the recipient**.**

The recipient isn't notified that he has email. To find out if he has new email the recipient must connect to the email server S2 and check his mailbox.

All desktop based email clients will have a default email check interval when they check for new email, as well as a manual option (Figure 5.2)

Figure 5.2 Desktop based E-mail client options

There are two protocols used for transferring email from an email server to a client –POP3 (Post Office Protocol) and IMAP4 (Internet Message Access Protocol).

SMTP is not used for receiving email on an email client.POP3 is the oldest protocol, IMAP4 is the newer and much better protocol, and is the preferred option today.

The main difference between the two is that POP3 was designed to move the email from the email server to the email client and delete it from the email server. This means that once a POP3 client has connected to the email server it leaves the mailbox on the server empty and the only copy of the email is now on the email client.IMAP4 was designed to leave all email on the email server and only copy the emails to the email client.

5.7 Search Engines

The World Wide Web is full of information. All kinds of articles, reports, news, tables, databases, books, reviews, etc. are stored in this huge global network. To find the most relevant resources, search engines are used to locate and return the documents that match the specific criteria from the search query.

On a user level, the procedures are very simple. All the user has to do is to input necessary key word of the search. In addition, most search engines offer additional features that allow the user to narrow down the search. Such feature is typically referred to as advanced search. With this the user can not only indicate what the search engine must fetch, but can also specify when, what kind, where, language, filtering and other criteria. These features are common and they are included into every search engine.

The method that the search engines used to decide which page match with the query is a real science. Search engines have their own technologies, algorithms, and formulas. Generally, search engines use some parameters and criteria to determine relevancy. Those that are widely know include the occurrence of the search query term on the page text, the title, the meta tags, in the liking text inside the page, in the URL address (Uniform Resource Locator), the headings, as well as bold and italicized font, and in the text outside the site but linking to it. Google's Page Rank is famous for valuing the external linking, while MSN relies more on the actual content of the page and not the links.

Again, all search engines are unique, and have their own formulas and algorithms, but they perform the same function; they fetch relevant results that match the user's query.

In order to build up a huge database of the web-sites, search engines use their crawlers or spiders, or bots. A crawler is an automated web-browser, which follows every link it sees. These crawlers visit a site and all its internal pages, gather them, and bring back to the server for index. Thus, a search engine keeps a saved version of all indexed pages on the server; these saved pages are referred to as cache.

Once the page is indexed, the search engine algorithms and formulas analyze the content of the pages. The technology checks the occurrences within the page and checks the external links from without to this page to determine the relevancy for a certain keyword. The combination of the various criteria gives the page a certain ranks under this keyword or key-phrase query, and then the engine returns the results to the user.

In addition to key word search, searching methods adopted by search engine are Boolean search, Concept search, Full text index, Fuzzy search, Phrase search, Query-by-example search etc.

Some popular search engines are Google, Bing, Yahoo, Ask, Aol.

Meta search engines

Web Meta searchers provide a single interface that enables users to search many different search engines, indexes and databases simultaneously.

Some popular Meta search engines are Dogpile, Sputtr, Clusty, IxQuick/StartPage, Mamma.

Special Search engines

To do academic research or purely scientific research, these Sites can be an invaluable aid.

Google Scholar (http://scholar.google.com/): Google Scholar is a free academic search engine that indexes academic information from various online web resources.

CiteSeer[x] (http://citeseerx.ist.psu.edu): CiteSeer[x] is a digital library and an online academic journal that offer information within the field of computer science. It indexes academic resources through autonomous citation indexing system. This academic database is particularly helpful for students seeking information on computer and information sciences.

5.8 Internet and Viruses

5.8.1 Computer Virus

A VIRUS (Vital Information Resources Under Siege) is a malicious program designed to disperse copies of itself to other computers and disrupt those computer's normal operations. A computer virus usually attaches to or inserts itself in an executable file or the boot sector (the area that contains the first instructions executed by the computer when it is started or restarted) of a disk. Although some viruses are merely disruptive, others can destroy or corrupt data or cause an operating system or applications program to malfunction.

Virus is the generic term used to describe a group of destructive computer programs. The three most common types of destructive programs are the Worm, Trojan horse and the Logic bomb.

A *worm* is a program that replicates itself. It creates an image of itself either in a file or a particular location on the disk. It also propagates itself over a network, reproducing as it goes.

A *Trojan horse* program is hidden inside another, apparently useful program. While the useful program is running, this program does something nasty, like erase the File Allocation Table (FAT) and directory. This type of program may be distributed through USB drive, an external disk, Internet downloads, e-mail etc.

A Bomb is a piece of code embedded in a program or the operating system itself that waits for a particular event to occur such as specific date. When that event occurs, the logic bomb goes off, doing some kind of damage.

Millions of computer malware programs are known; they can be spread via removable disk or drive, networks or Internet websites and services.

5.8.2 Internet and viruses

The Internet brings a new dimension to the virus problem. Earlier days, viruses generally spread from system to system on physical media, often the floppy disk. This is a fundamentally slow way for viruses to spread and they are bad at reproduction. The Internet changes all this.

Most of the danger on the Internet currently comes from old viruses exploiting new paths for transmission. There are basically two ways they can do this: innocent and malicious distribution.

Innocent virus distribution

Sharing software over the Net is simple and easy; a simple mouse click attaches a program to an e-mail, and it is just as easy to detach and run it. People can place a program on their web page almost as simply, and this can be downloaded by anyone anywhere. Any one of these programs could be infected.

What kinds of viruses could these practices spread? Purely boot sector viruses are out. Parasitic file viruses work well in this environment, although many (but no means all) users are cautious about obtaining programs from places they do not trust.

The viruses that *really* win in the Internet environment are the macro viruses. They are attached to data, not code, making them harder to avoid. An increasing number of documents on the Net are available as Word files, for example, with no alternative format, and Word documents are frequently exchanged via e-mail.

The only solution here is to obtain viewer programs which read the data in the file but ignore the macros. Such programs are available for Word and Excel among others. Never open a file that does not trust with the application that created it.

Malicious virus distribution

Viruses may also be spread by malicious individuals, knowingly passing on infected programs. Virus authors and others find the Internet perfect for giving a new virus a start in life, by means of hundreds of unsuspecting Internet users; by infecting an attractive-looking file that then gets placed in a public download area, the virus can spread far in a short time.

Viruses typically are programs. But that's not always the case. In case of Internet actually the virus is not only a program, it also is available in other forms also. The following are the some forms of virus through Internet.

Virus through e-mail

There have been an increasing number of hoaxes and scares stories about email viruses in recent years.

A Virus as an E-mail Hoax

Sometimes we may receive, forward chain letter, a virus warning, or the email of a desperate parent whose child needed an organ donation?, we probably fell for a hoax and helped to spread it.

In this case the virus is the e-mail itself, although it's not really an infection. The damage, however, is the same as with other viruses: clogging up of inboxes and mail servers. E-mail viruses generally cause a waste of time, resources, and energy.

Don't forward hoax email, unless we're sure it's genuine. Hoax-Slayer is a good resource to get informed about email hoaxes and eventually verify the credibility of an e-mail we received.

A Virus in an E-mail Attachment

While forwarding an e-mail doesn't cause any major damage, an infection with a "real" computer virus is a lot more critical. There are "harmless" variants that just forward themselves to the contacts. However, a virus may also be programmed for example, to destroy specific files on our system.

Most viruses are delivered through an e-mail attachment. Attachments that contain viruses are either executable programs (file types: .com, .exe, .vbs, .zip, .scr, .dll, .pif, .js) or macro viruses (file types: .doc, .dot, .xls, .xlt). The safest way to avoid them is to not open e-mail attachments.

It is safe to open Word documents in alternative programs that don't support macros, such as WordPad or Open Office.

Simply downloading and reading the email will not do any harm. As long as the attachment is not opened, the virus won't hatch. If we don't know the sender and if the e-mail text is suspicious, just delete the e-mail along with the attachment.

A Virus in the Email Body

Sometimes the malicious content can be found in the body of an e-mail. Today, HTML is a common element of e-mails as it is used to embed pictures and links. However, HTML can also be used to embed scripts that execute automatically and subsequently infect our computer with a virus. That's why many mail programs, by default, block HTML and make click a button to display content of trustworthy sources.

Also, URLs can be a virus in disguise. We may see a harmless link that either leads to a website that executes a malicious script or links to a completely different URL where we automatically download a self-executing virus. The same rule as for attachments applies: never view or access links from suspicious sources.

Macro viruses

Macro viruses have become common in Internet. Most of these viruses are written in the scripting languages for Microsoft programs such as Word and Excel and spread throughout Microsoft Office by infecting documents and spreadsheets. Since Word and Excel were also available for Mac OS, most could also spread to Macintosh computers. Although most of these viruses did not have the ability to send infected email messages, those viruses which did take advantage of the Microsoft Outlook COM interface

A virus may also send a web address link as an instant message to all the contacts on an infected machine. If the recipient, thinking the link is from a friend (a trusted source) follows the link to the website, the virus hosted at the site may be able to infect this new computer and continue propagating.

Viruses from Face book

There are a variety of viruses that can be obtained from using Face book. One kind is a worm. It can copy itself within our computer as well as from one machine to another. A worm infects the computer as well as the Internet and slows things down significantly.

Zeus is a Trojan virus that gets into the computer without our knowledge. It remains dormant until we access our bank account, and then it has access to our log-in information and is able to steal from the account.

Text Book:

Computers – Basics to Advancements, Dr.P.Velmani & Dr.V.Lakshmi Praba, 1/e, Chess Educational Publishers, Chennai.

Reference Books:

1. Introduction to Computer Science, ITL, Education Solutions Limited, 2/e, Pearson
2. Introduction to Computers, Peter Norton, 7/e, TMH.

Prepared by

Dr. Mrs. P.VELMANI

Assistant Professor, Department of Computer Science,

The M.D.T.Hindu College, Tirunelveli.