CHAPTER 1
THE COMPUTING SYSTEM

1.1 DEFINITION OF COMPUTER

A computer is an electronic device that can accept data in form of input through an input device, process the data and yield result (output). In other words, it is simply defined as "an electronic device that process data to yield useful information".

1.2 OVERVIEW OF THE COMPUTING SYSTEM

The computing system is made up of the computer system, the user and the environment in which the computer is operated. The schematic diagram of the computing system is presented in Fig. 1.1.
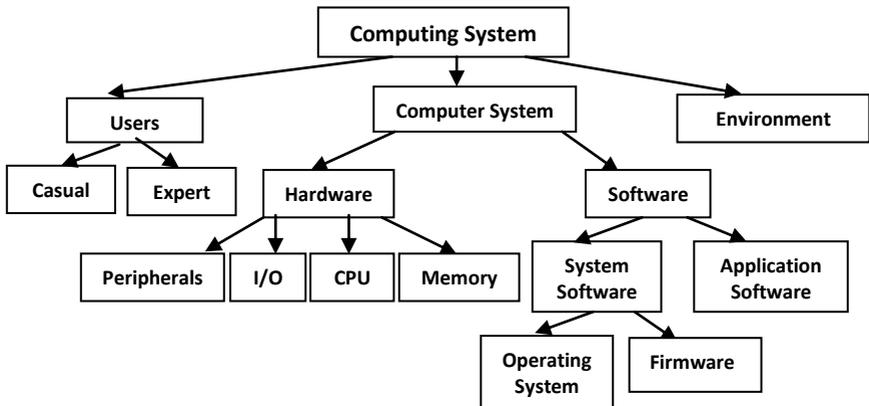

Fig. 1.1 A schematic diagram of the computing system

1.2.1 The Computer System

The computer system is made up of the hardware and the software. The characteristics of a computer system include the following:

i. Speed: The computer can manipulate large data at incredible speed and response time can be very fast.

ii. Accuracy: Its accuracy is very high and its consistency can be relied upon. Errors committed in computing are mostly due to human rather than technological weakness. There are in-built errors detecting schemes in the computer.

iii. Storage: It has both internal and external storage facilities for holding data and instructions. This capacity varies from one machine to the other. The basic unit of a memory location is a byte which is 8 bits (BInary digiTS). Larger units include Kilo (2^{10} = 1024) bytes, Mega (2^{20}) bytes, Giga (2^{30}) bytes and Tera (2^{40}) bytes.

iv. Automation: Once a program is in the computer's memory, it can run automatically each time it is opened. The individual has little or no instruction to give again.

v. Reliability: Being a machine, a computer does not suffer human traits of tiredness and lack of concentration. It will perform the last job with the same speed and accuracy as the first job every time even if ten billion jobs are involved.

vi. Flexibility: It can perform any type of task once it can be reduced to logical steps. Modern computers can be used to perform a variety of functions like on-line processing, multi-programming, real time processing etc.

| Hardware | Software |

Fig. 1.2 The Computer System

1.2.1.1 The Hardware

The computer hardware comprises the input unit, the processing unit and the output unit. The input unit comprises those media through which data is fed into the computer. Examples include the keyboard, mouse, joystick, trackball and scanner. The processing unit is made up of the Arithmetic and Logic Unit (ALU), the control unit and the main memory. The main memory also known as the primary memory is made up of the Read Only Memory (ROM) and the Random Access Memory (RAM). The output unit is made up of those media through which data, instructions for processing the data (program), and the result of the processing operation are displayed for the user to see. Examples of the output unit are the monitor (Visual Display Unit) and the printer.

1.2.1.2 Software

Computer software is the series of instructions that enable the computer to perform a task or group of tasks. A program is made up of a group of instructions to perform a task. Series of

programs linked together make up software. Computer programs could be categorized into system software and application programs.

1.2.2 Computer Users

Computer users are the different categories of personnel that operate the computer. We have expert users and casual users. Categories of expert users include computer (hardware) engineers, computer scientists, programmers, software engineers and system/data analysts.

1.2.3 The Computing Environment

The computing environment includes the building that houses the elements of the computing system namely the computer and the users, the furniture, auxiliary devices such as the voltage stabiliser, Uninterruptible Power Supply (UPS), air conditioners etc.

1.3 HISTORICAL OVERVIEW OF THE COMPUTER
1.3.1 A Brief History of Computer Technology

A complete history of computing would include a multitude of diverse devices such as the ancient Chinese abacus, the Jacquard loom (1805) and Charles Babbage's analytical engine (1834). It would also include a discussion of mechanical, analog and digital computing architectures. As late as the 1960s, mechanical devices, such as the Merchant calculator, still found widespread application in science and engineering. During the early days of electronic computing devices, there were much discussion about the relative merits of analog vs. digital computers. In fact, as late as the 1960s, analog computers were

routinely used to solve systems of finite differential equations arising in oil reservoir modelling. In the end, digital computing devices proved to have the power, economics and scalability necessary to deal with large scale computations. Digital computers now dominate the computing world in all areas ranging from the hand calculator to the supercomputer and are pervasive throughout society. Therefore, this brief sketch of the development of scientific computing is limited to the area of digital electronic computers.

The evolution of digital computing is often divided into generations. Each generation is characterized by dramatic improvements over the previous generation in the technology used to build computers, the internal organization of computer systems, and programming languages. Although not usually associated with computer generations, there has been a steady improvement in algorithms, including algorithms used in computational science.

1.3.2 Generations of Electronic Computers

The following history has been organized using these widely recognized generations as mileposts.

1.3.2.1 First Generation (1942-1955)

In 1942, John P. Eckert, John W. Mauchly and their associates at the Moor School of Electrical Engineering of University of Pennsylvania decided to build a high – speed programmable electronic computer to calculate artillery firing tables for the United States Army's Ballistic Research Laboratory. This machine became known as Electronic Numerical Integrator and Calculator (ENIAC). ENIAC was built between July 1943 and November 1945 and was the fastest and largest machine of its time.

The features of ENIAC are:
- ➢ It is has several electromechanical parts
- ➢ Used about 18,000 vacuum tubes
- ➢ Very bulky (uses about 1800 square feet of floor space and weighs 30 tonnes)
- ➢ Consume about 180,000 watts of electrical power
- ➢ Gears easily wear out
- ➢ Has punched card I/O, 1 multiplier, 1 divider/square rooter, and 20 adders using decimal ring counters
- ➢ Program were embodied in separate —units‖ and plugged together to form a —route for the flow of information
- ➢ Numerical word size was 10 decimal digits.
- ➢ It could multiply two of these numbers at a rate of 300 per seconds, by finding the value of each product from a multiplication table stored in its memory 1000 times faster than mechanical relay calculator
- ➢ Commonly accepted as the first successful high-speed Electronic Digital Computer (EDC)
- ➢ Efficient in handling the particular programs for which it had been designed

In 1945, an abstract study of computation showed that a computer should have a very simple, fixed physical structure, and yet be able to execute any kind of computation by means of a proper programmed control without the need for any change in the unit itself. Von Neumann contributed a new awareness of how practical, yet fast computers should be organized and built. These ideas, usually referred to as the stored - program technique, became essential for future generations of high-speed digital computers and were universally adopted. The stored program technique involves many features of computer design and function

besides the one that it is named after. In combination, these features make very – high – speed operation attainable.

- ➢ Von Neumann was able to come up with stored program technique with a very high speed operation
- ➢ The machine instruction called —conditional control transfer allows the program sequence to be stopped and start again at any point.
- ➢ Instruction programs are stored together with data in the same memory units, so that when needed, instructions could be arithmetically changed in the same way as data
- ➢ Computing and programming became much faster, more flexible and more efficient
- ➢ Regularly used subroutines are kept in the Libraries for subsequent use
- ➢ It uses Random – Access – Memory (RAM)
- ➢ It has punched card or punched tape I/O device RAMs of 1000 word capacity and access time of 0.5*10-6 seconds
- ➢ They were much smaller than ENIAC
- ➢ Needed a lot of maintenance (gear easily wear out)
- ➢ Reliability Of Operation (ROD) is about 70 – 80 %
- ➢ Usually programmed in Machine Language

The first generation of modern programmed electronic computers by Von Neumann to take the above advantage was built in 1947, EDVAC (Electronic Discrete Variable Computer).

1.3.2.2 Second Generation (1955-1965)
Early in the 50's two important engineering discoveries (Transistor and magnetic core memory) changed the image of the electronic computer field, from one of the fast but unreliable hardware to an

image of relatively high reliability and even more capability. The size of the computers was decreased considerably by replacing vacuum tubes with transistors. Examples of second generation computers are IBM 7000 series, IBM 1400 series and CDC 164.

The features are:

- ➢ Transistor invented at Bell Telephone Laboratories replaced vacuum tubes in the design of switching circuits
- ➢ Magnetic core memory replaced the cathode ray tube and delay – line memories
- ➢ RAM capacities increased from 8000 to 64000 words in commercially available machine with access time of 2 to 3 ms (milliseconds)
- ➢ Very expensive to purchase or even to rent
- ➢ Expensive to operate because of the cost of expanding programming
- ➢ Machine independent High Level Language such as ALGOL, COBOL and FORTRAN were introduced to simplifying programming
- ➢ Allow batch processing in which problems are prepared and then held ready for computation on a relatively cheap medium
- ➢ When a process finished, it dumps the whole problem (program and results) on one of the peripheral storage units and starts on a new problem
- ➢ Magnetic drum, magnetic disk packs, or magnetic tapes were usually used
- ➢ Allow line – sharing such that computer processes many jobs, run as if the other jobs did not exist, thus keeping each customer satisfied

1.3.2.3 Third Generation (1965-1975)

In the 1960's, efforts to design and develop the fastest possible computer with the greatest capacity reached a turning point with the Livermore Advanced Research Computer (LARC) Machine, built for the Livermore Radiation Laboratories of the University of California by the Sperry–Rand corporation, and the stretch Computer by IBM. The invention of solid state integrated circuit technology brought about miniaturized circuits. The third generation of computers was pioneered by IBM. The result of this was IBM 360 and later IBM 370 for multiprogramming. Until the early 1960's, every computer model was generally designed independently, and sometimes individual machines were custom modified for a particular customer. IBM changed this forever, when IBM-360 family of computers were announced in April 1964 with models 20 - 91.

The IBM System/360 Model 50 had approximately ten times the internal power of the Model 30 when performing a mix of computations. It was the most powerful unit in the medium price range with communication facilities.

The features are:
- ➤ Integrated Circuits (ICs) began to replace the discrete transistor circuits
- ➤ Semiconductor memories began to augment and ultimately replaced the ferrite's core in main memory designs
- ➤ Variety of techniques for concurrent or parallel processing were introduced, such as multiprogramming and multiprocessing

- ➢ **Automatic sharing of the facilities or resources of a computer system such as the processors and memory space were developed.**
- ➢ **Computer of various capabilities and accessories of various capabilities were produced such as card feeders, page printers, cathode-ray-tube displays and graphing devices.**
- ➢ **The system were widely used in businesses such as accounting (payroll, inventory control, ordering supplies, billing), hospitals, libraries such as National Medical Library, chemical abstracts system.**
- ➢ **Gains in applications programming language technologies removed many problems.**
- ➢ **Multipurpose and larger ranges of application systems were produced.**
- ➢ **Less costly**
- ➢ **Reduced size**
- ➢ **Much more reliable and compatible**
- ➢ **Internal components increased**
- ➢ **Development of minicomputers**
- ➢ **The emergence of the software industry**

1.3.2.4 Fourth Generation Computers (1975 — Till Date)

The present day computers that you see today are the fourth generation computers that started around 1975. The improvement of solid state electronics contributed immensely to the design and manufacture of the fourth generation computers. The use of this technology made possible the balkanization of hitherto large components into smaller ones, also the use of bubble memory became prominent. Other features of this Computer generation are as follows:

- ➤ **Very Large Scale Integration (VLSI) replaced ordinary Integrated Circuits (ICs) of the**
- ➤ **third generation**
- ➤ **Emergence of the Microprocessor, and invariably the Microcomputers / personal**
- ➤ **computers**
- ➤ **Capabilities of computers were fully realized.**
- ➤ **Computer became common market items especially in advanced countries**
- ➤ **Small and inexpensive**
- ➤ **Computers became available for recreation purposes such as playing of games**
- ➤ **Internal memory employed Metal-Oxide Semiconductor (MOS) which is faster, and cheaper than core.**
- ➤ **Emergence of super computers as well as super microcomputers.**

1.3.2.5 Fifth Generation (Present and Beyond)

The Fifth Generation Computer Systems project (FGCS) was an initiative by Japan's Ministry of International Trade and Industry, begun in 1982, to create a fifth generation computer which would perform much calculation using massive parallel processing. It relies heavily on the technique of Artificial Intelligence (AI). With increased processing power, computers can understand spoken words, identify objects, classify objects, imitate human reasoning and perform human capabilities by sensing its surroundings using different types of sensors. In all the fifth generation computers, the import is on endowing the computer with the ability to take decisions in various circumstances as well as simulating the human brain through its myriad of neurons.

The features are:
- ➢ Emergence of new programming techniques such as object-oriented programming to
- ➢ improves software development and production.
- ➢ Comparatively reduced cost of hardware with increased computing power
- ➢ Conveniences and efficiency increased
- ➢ Experiencing huge growth in the area of computer networking, computer mail,
- ➢ electronic publishing
- ➢ Artificial intelligence and robotic
- ➢ Knowledge-based systems
- ➢ Automation of industrial processes
- ➢ Expert systems
- ➢ Computer assisted design and manufacturing
- ➢ Parallel computing

1.4 CLASSIFICATION OF COMPUTER

Although there are no industry standards, computers can generally be classified based on the following parameters: signal type, purpose and capacity.

1.4.1 Classification Based on Signal Type

There are basically three types of computers based on signal type. These are Digital, Analog and Hybrid computers.

1.4.1.1 The Digital Computer

This represents its variables in the form of digits. The data it deals with, whether representing numbers, letters or other symbols, are converted into binary form on input to the computer. The data undergoes a processing after which the binary digits are converted back to alphanumeric form for output for human use.

12

Because of the fact that business applications like inventory control, invoicing and payroll deal with discrete values (separate, disunited, discontinuous); they are best processed with digital computers. As a result of this, digital computers are mostly used in commercial and business places today.

1.4.1.2 The Analog Computer

It measures rather than count. This type of computer sets up a model of a system. The common type represents its variables in terms of electrical voltage and sets up circuit analog to the equation connecting the variables. The answer can be obtained either by using a voltmeter to read the value of the variable required, or by feeding the voltage into a plotting device. Analog computers hold data in the form of physical variables rather than numerical quantities. In theory, analog computers give an exact answer because the answer has not been approximated to the nearest digit. Whereas, when we try to obtain the answers using a digital voltmeter, we often find that the accuracy is less than that which could have been obtained from analog computer.

It is almost never used in business systems. It is used by scientists and engineers to solve systems of partial differential equations. It is also used in controlling and monitoring of systems in such areas as hydrodynamics and rocket production. There are two useful properties of the analog computer once it is programmed:

- It is simple to change the value of a constant or coefficient and study the effect of such changes.
- It is possible to link certain variables to a time pulse to study changes with time as a variable, and chart the result on an X-Y plotter.

1.4.1.3 The Hybrid Computer

In some cases, the computer user may wish to obtain the output from the analog computer as processed by a digital computer or vice versa. To achieve this, a hybrid machine may be set up where the two are connected and the analog computer may be regarded as a peripheral of the digital computer. In such a situation, a hybrid system attempts to gain the advantage of both the digital and the analog elements in the same machine. This kind of machine is usually a special-purpose device which is built for a specific task. It needs a conversion element which accepts analog inputs, and outputs digital values. Such converters are called digitisers. There is also a need for a converter from digital to analog with the advantage of giving real-time response on a continuous basis. Complex calculations can be dealt with by the digital elements, thereby requiring a large memory, and giving accurate results after programming. They are mainly used in aerospace and process control applications.

1.4.2 Classification by Purpose

Depending on their flexibility in operation, computers are classified as either general purpose or special purpose.

1.4.2.1 General-Purpose Computers

General-purpose computers are computers designed to handle a wide range of problems. Theoretically, a general-purpose computer can be adequate by means of some easily alterable instructions to handle any problems that can be solved by computation. In practice however, there are limitations imposed by memory size, speed and the type of input/output devices. Examples of areas where general purpose computers are employed include the following:

- Payroll
- Banking

- **Billing**
- **Sales analysis**
- **Cost accounting**
- **Manufacturing/Scheduling**
- **Inventory control**

The general Attributes of general-purpose computers include:

- **General-purpose computers are more flexible than special purpose computers. Thus, the former can handle a wide spectrum of problems.**
- **They are relatively less efficient than the special-purpose computers due to such problems as the following:**
    - **They have inadequate storage**
    - **They have low operating speed**
    - **Coordination of the various tasks and subsections may take time.**
- **General-purpose computers are typically more complex than special purpose computers.**

1.4.2.2 Special-Purpose Computers

A special purpose computer is one that is designed to solve a restricted class of problems. Such computers may even be designed and built to handle only one job. In such machines, the steps or operations that the computer follows may be built into the hardware. Most of the computers used for military purposes fall into this class. Other examples of special purpose computers include:

- **Computers designed specifically to solve navigational problems.**
- **Computers designed for tracking airplanes or missiles**

- **Computers used for process control applications in industries such as oil refinery, chemical manufacture, steel processing and power generation.**
- **Computers used as robots in factories like vehicle assembly plants and glass industries.**

Special-purpose computers are usually very efficient for the tasks for which they are specially designed. They are very much less complex than the general-purpose computers. The simplicity of the circuitry stems from the fact that provision is made only for limited facilities. They are typically cheaper than the general-purpose type since they involve fewer components and are less complex depending on the application.

1.4.3 Classification of Computers According to Capacity

In the past, the capacity of computers was measured in terms of physical size. Today, however, physical size is not a good measure of capacity because modern technology has made it possible to achieve compactness. A better measure of capacity today is the volume of work that a computer can handle. The volume of work that a given computer handles is closely tied to the cost and to the memory size of the computer. Here, both memory size and cost shall be used to rank (classify) computers into two main categories as follows:

- **Super computer / mainframes**
- **Microcomputers**

1.4.3.1 Supercomputers / Mainframes

A supercomputer is a computer at the frontline of contemporary processing capacity, particularly speed of

calculation. These are the equivalent of mainframe computers, often called number crunchers and are very expensive. They are used by many people for a variety of purposes. Supercomputers were introduced in the 1960s, designed initially and, for decades, primarily by Seymour Cray at Control Data Corporation (CDC), Cray Research and subsequent companies bearing his name or monogram. While the supercomputers of the 1970s used only a few processors, in the 1990s machines with thousands of processors began to appear and, by the end of the 20th century, massively parallel supercomputers with tens of thousands of "off-the-shelf" processors were the norm. As of June 2013, China's Tianhe-2 supercomputer is the fastest in the world at 33.86 petaFLOPS (33.86 * 10^{15} FLoating point Operations Per Second).

Supercomputers play an important role in the field of computational science, and are used for a wide range of computationally intensive tasks in various fields, including quantum mechanics, weather forecasting, climate research, oil and gas exploration, molecular modelling (computing the structures and properties of chemical compounds, biological macromolecules, polymers, and crystals), and physical simulations (such as simulations of the early moments of the universe, airplane and spacecraft aerodynamics, the detonation of nuclear weapons, and nuclear fusion).

1.4.3.2 Microcomputers

Microcomputers, also known as personal computers (PC's) are the cheapest class of computers. A few (one or more) microprocessor chips are the main data processing unit. They are cheaper and smaller, and can operate under normal office conditions. Examples of microcomputer manufacturers include IBM, APPLE, COMPAQ, Hewlett Packard (HP), Dell and Toshiba.

There are different types of Personal Computers which include:

- **Desktop Computers: Typically, this type of personal computer is placed on the desk. Such computers consist of separate system unit, VDU, Keyboard/Mouse and other peripherals such as UPS and speakers.**
- **Laptop Computers: These are small size types that are battery operated. The system unit, VDU, Keyboard/Mouse and other peripherals such as UPS (battery) and speakers are all integrated into a single box-like structure and can be operated in vehicles while on a journey.**
- **Notebook Computers: These are like laptop computers but smaller in size. Though small, the notebook computer comprises all the components of a full system. They are sometimes called Mini Laptops or Netbook.**
- **Palmtop Computers: The palmtop computer is far smaller in size. All the components are complete as in any of the above, but it is made smaller so that it can be held on the palm. Today's tablet computers (e.g. Apple iPad, Samsung Galaxy Tabs, Amazon Kindle) and modern smart phones (e.g. Apple iPhone, RIM BlackBerry) would fall in this category.**

Uses of the Personal Computer

A personal computer can perform the following functions:

- **It can be used to produce documents like memos, reports, letters and briefs.**
- **It can be used to calculate budgets and accounting tasks**
- **It can analyze numeric functions**
- **It can create illustrations**
- **It can be used for electronic mails**

- It can help in making schedules and planning projects
- It can assist in searching for specific information from lists or from reports.

Advantages of the Personal Computer

- The personal computer is versatile; it can be used in any establishment.
- It has faster speed for processing data
- It can deal with several data at a time
- It can attend to several users at the same time, thereby being able to process several jobs at a time
- It is capable of storing huge data
- Operating the personal computer gives less fatigue
- It is possible to network personal computers, that is, linking of two or more computers.

Disadvantages of the Personal Computer

- The personal computer is costly to maintain
- It is very fragile and complex to handle
- It requires special skill to operate
- It can lead to unemployment, especially in less developed countries
- Some computers cannot function properly without the aid of a cooling system, e.g. air conditioners or fans in some locations.

1.5 BASIC COMPUTER OPERATIONS

A computer performs basically five major operations or functions irrespective of their size and make. These are

1. It accepts data or instructions by way of input
2. It stores data

3.   It can process data as required by the user,
4.   It gives results in the form of output, and
5.   It controls all operations inside a computer.

These operations are illustrated in Fig. 1.2 and discussed below.
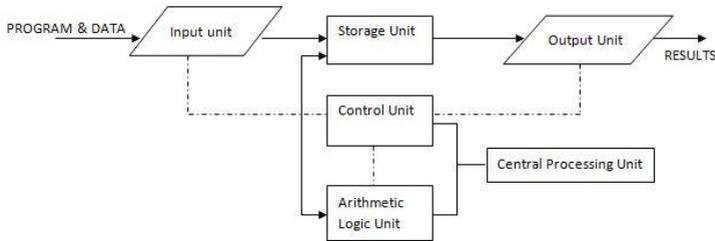
Fig. 1.2 Basic Operations of a Computer

1. Input: This is the process of entering data and programs into the computer system. It should be known that the computer is an electronic machine like any other machine which takes as inputs, raw data, and performs some processing to give out processed data. Therefore, the input unit takes data from us to the computer in an organized manner for processing.

2. Storage: The process of saving data and instructions permanently is known as storage. Data has to be fed into the system before the actual processing starts. It is because the processing speed of Central Processing Unit (CPU) is so fast that the data has to be provided to CPU with the same speed. Therefore the data is first stored in the storage unit for faster access and processing. This storage unit or the primary storage of the computer system is designed to do the above functionality. It provides space for storing data and instructions. The storage unit performs the following two major functions:

- All data and instructions are stored here before and after processing.
- Intermediate results of processing are also stored here.

3. Processing: The task of performing operations like arithmetic and logical operations is called processing. The Central Processing Unit (CPU) takes data and instructions from the storage unit and makes all sorts of calculations based on the instructions given and the type of data provided. It is then sent back to the storage unit.

4. Output: This is the process of producing results from the data for getting useful information. Similarly the output produced by the computer after processing must also be kept somewhere inside the computer before being given to users in human readable form. Again the output is also stored inside the computer for further processing.

5. Control: This is the manner in which instructions are executed and the other four operations above are performed. Co-ordination of all operations like input, processing and output are performed by control unit. It takes care of step by step processing of all operations inside the computer.

1.6 CENTRAL PROCESSING UNIT

The Central processing Unit (CPU) is made up of the Arithmetic Logical Unit (ALU) and the Control Unit (CU); it is the brain of any computer system. Similar to the function of the human brain, the CPU takes all major decisions, makes all sorts of calculations and directs different parts of the computer functions by activating and controlling the operations.

1.6.1 Arithmetic Logical Unit (ALU)

After data is entered through the input device, it is stored in the primary storage unit. The actual processing of the data and instruction are performed by Arithmetic Logical Unit. The major operations performed by the ALU are addition, subtraction, multiplication, division, logic and comparison. Data is transferred to ALU from storage unit when required. After processing the output is returned back to storage unit for further processing or getting stored.

1.6.2 Control Unit (CU)

The next component of CPU is the Control Unit, which acts like the supervisor seeing that things are done in proper fashion. The control unit determines the sequence in which computer programs and instructions are executed. Things like processing of programs stored in the main memory, interpretation of the instructions and issuing of signals for other units of the computer to execute them. It also acts as a switch board operator when several users access the computer simultaneously. Thereby it coordinates the activities of computer's peripheral equipment as they perform the input and output.

1.7 MEMORY SYSTEM IN A COMPUTER

There are two kinds of computer memory: primary and secondary. Primary memory is accessible directly by the processing unit. Random Access Memory (RAM) is an example of primary memory. As soon as the computer is switched off the contents of the primary memory is lost. Data can be stored and retrieved much faster with primary memory compared to secondary memory. Secondary memories such as floppy disks, magnetic (hard) disk, etc., are located outside the CPU. Primary memory is relatively more expensive than secondary memory. Because of

this, the size of primary memory is typically less than that of secondary memory.

Computer memory is used to store two things:

i) Instructions to execute a program

ii) Data: when the computer is doing any job, the data that have to be processed are stored in the primary memory. This data may come from an input device like keyboard or from a secondary storage device like a floppy disk.

As the program or the set of instructions is kept in primary memory, the computer is able to follow instantly the set of instructions. For example, when you book ticket from railway reservation counter, the computer has to follow the same steps: take the request, check the availability of seats, calculate fare, and wait for money to be paid, store the reservation and get the ticket printed out. The program containing these steps is kept in memory of the computer and is followed for each request.

But inside the computer, the steps followed are quite different from what we see on the monitor (screen). In the computer's memory, both programs and data are stored in the binary form. The binary system has only two values 0 and 1. These are called bits (BInary digiTS). A large number of integrated circuits inside the computer can be considered as switches, which can be made ON or OFF. If a switch is ON, it is considered 1 and if it is OFF, it is 0. A number of switches in different states will give you a message like this: 110101....10. Therefore, the computer takes input in the form of 0s and 1s, and gives output in the form of 0s and 1s only. However, there are encoding techniques such as ASCII and EBCDIC that allows conversion of alphabets, numbers and symbols into binary and vice versa.

The primary memory in the computer is in the form of ICs (Integrated Circuits). These circuits are called Random Access Memory (RAM). Each of RAM's locations stores one byte of information. One byte is equal to 8 bits. This can be either 0 or 1. The primary or internal storage section is made up of several small storage locations called cells. Each of these cells can store a fixed number of bits called word length. Each cell has a unique number assigned to it called the address of the cell and it is used to identify the cells. The memory is like a large cabinet containing as many drawers as there are addresses. Each drawer contains a word and the address is written on the outside.

1.7.1 Capacity of Computer Memory

Each cell of memory contains one character or 1 byte of data. Therefore, the capacity of memory is defined in terms of byte or words. Thus, 64 kilobyte (KB) memory is capable of storing 64 * 1024 = 32,768 bytes. 1 kilobyte equals 1024 (2^{10}) bytes. A memory size ranges from few kilobytes in microcomputers to several thousand kilobytes in super computers. Today's personal computers have primary memory capacity ranging 256MB to several Gigabytes.

1.7.2 Primary Memory

There are different types of primary memory on the computer. These are discussed below. Note that the RAM and ROM are also known together as main memory.

1. Random Access Memory (RAM): This primary storage is referred to as random access memory (RAM) because it is possible to randomly select and use any location of this memory to directly store and retrieve data. It takes same time to any address of the memory as the first address. It is also

called read/write memory. The storage of data and instructions inside this primary storage is volatile. The data disappears from RAM as soon as power to the computer is switched off. In other words, the memories lose their content on failure of power supply.

2. Read Only Memory (ROM): There is another primary memory in the computer. Again it is an IC that forms the ROM. The storage of program and data in the ROM is non-volatile. The ROM stores some standard processing programs supplied by the manufacturers to operate the personal computer. The ROM can only be read by the CPU but it cannot be unnecessarily changed. The basic input/output system (BIOS) program is stored in the ROM that examines and initializes the basic devices attached to the PC when it is switch on. The ROM does not lose their content on failure of power supply. Two special types of ROM exist apart from the standard ROM namely PROM and EPROM.

a. Programmable Read Only Memory (PROM): It is known that it is not possible to modify or erase programs stored in a standard ROM. However, it is possible to program the contents of an empty PROM chip (IC). Once the programs are written, it cannot be changed and remain intact even if power is switched off. This means that programs or instructions written in PROM or ROM cannot be erased.

b. EPROM: This stands for Erasable Programmable Read Only Memory, which overcomes the problem of PROM & standard ROM. EPROM chips can be programmed time and again by erasing the information stored earlier in it. Information stored in EPROM can be erased and reprogrammed using a special programming facility.

When the EPROM is in use, information can only be read. This is the most common form of ROM today.

3. **Cache Memory:** The speed of CPU is extremely high compared to the access time of other primary memories. Therefore the performance of the CPU decreases due to the slow speed of main memory. To decrease the mismatch in operating speeds, a small memory chip is attached between CPU and Main memory whose access time is very close to the processing speed of CPU. It is called CACHE memory. CACHE memories are accessed much faster, than conventional RAM. It is used to store programs or data currently being executed or temporary data frequently used by the CPU. So each memory makes main memory to be faster and larger than it really is. It is also very expensive to have bigger size of cache memory and its size is normally kept small; currently in Megabytes.

4. **Registers:** The CPU processes data and instructions with high speed; there is also movement of data between various units of the computer. It is necessary to transfer the processed data with high speed. So the CPU (esp. ALU) uses a number of special memory units called registers. They are not part of the main memory but they temporarily store data or information used by the ALU and passes it on as directed by the control unit.

1.7.3 SECONDARY STORAGE

The operating speed of primary memory or main memory should be as fast as possible to cope with the CPU speed. These high-speed storage devices are very expensive and hence the cost per bit of storage is also very high. Again, the storage capacity of the main memory is also very limited. However, it is necessary to

store hundreds of millions of bytes of data for the CPU to process. Therefore, additional memory is required in all computer systems. This memory is called auxiliary memory or secondary storage.

In this type of memory, the cost per bit of storage is low. However, the operating speed is slower than that of the primary storage. Huge volume of data are stored here on permanent basis and transferred to the primary storage as and when required. Most widely used secondary storage devices are magnetic tapes and magnetic disk.

1. Magnetic Tape: Magnetic tapes were used for large computers like the mainframe computers where large volume of data is stored for a longer time. They also come in the form of cassettes. The cost of storing data in tapes is inexpensive. Tapes consist of magnetic materials that store data permanently. It can be 12.5mm to 25mm wide plastic film-type and 500 meter to 1200 meter long which is coated with magnetic material. The deck is connected to the central processor and information is fed into or read from the tape through the processor. It similar to cassette tape recorder. Note that the magnetic tapes are no longer in use.

Fig. 1.3 Magnetic Tape

Advantages of Magnetic Tape:

- **Compact:** A 10-inch diameter reel of tape is 2400 feet long and is able to hold 800, 1600 or 6250 characters in each inch of its length. The maximum capacity of such tape is 180 million characters. Thus data are stored much more compactly on tape.
- **Economical:** The cost of storing characters is very less as compared to other storage devices.
- **Fast:** Copying of data is easier and fast.
- **Long term Storage and Re-usability:** Magnetic tapes can be used for long term storage and a tape can be used repeatedly without loss of data.

2. **Magnetic Disk:** You might have seen the gramophone record, which is circular like a disk and coated with magnetic material. Magnetic disks used in computer are made on the same principle. It rotates with very high speed inside the computer drive. Data is stored on both the surface of the disk. Magnetic disks are most popular for direct access storage device. Each disk consists of a number of invisible concentric circles called tracks. Information is recorded on tracks of a disk surface in the form of tiny magnetic spots. The presence of a magnetic spot represents one bit and its absence represents zero bit. The information stored in a disk can be read many times without affecting the stored data. The reading operation is non-destructive. But if you want to write a new data, then the existing data is erased from the disk and new data is recorded. This is the technology used in today's hard drives.

3. **Floppy Disk:** It's similar to magnetic disk discussed above. They are 5.25 inches in diameter. They come in single or double density and recorded on one or both surface of the

diskette. The capacity of a 5.25-inch floppy is 1.2 megabytes; whereas for 3.5inch floppy it is 1.44 megabytes. It is cheaper than any other storage devices and is portable. The floppy is a low cost device particularly suitable for personal computer system. The use of floppy discs has diminished over time and going into extinction.
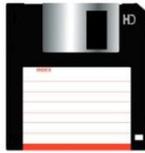
Fig. 1.4 Floppy Disk

4. **Optical Disk:** With every new application and software there is greater demand for memory capacity. It is the necessity to store large volume of data that has led to the development of optical disk storage medium. Optical disks can be divided into the following categories:

   a. **Compact Disk Read Only Memory (CD-ROM):** CD-ROM disks are made of reflective metals. CD-ROM is written during the process of manufacturing by high power laser beam. Here the storage density is very high, storage cost is very low and access time is relatively fast. Each disk is approximately 4 1/2 inches in diameter and can hold over 600MB of data. As the CD-ROM can be read only we cannot write or make changes into the data contained in it.

   b. **Write Once, Read Many (WORM):** The inconveniences that we cannot write anything in to a CD-ROM is avoided in WORM. A WORM allows the

user to write data permanently on to the disk. Once the data is written it can never be erased without physically damaging the disk. Here data can be recorded from keyboard, video scanner, OCR equipment and other devices. The advantage of WORM is that it can store vast amount of data amounting to gigabytes (10^9 or 2^{30} bytes). Any document in a WORM can be accessed very fast, say less than 30 seconds.

c. Erasable Optical Disk: These are optical disks where data can be written, erased and re-written. This also applies a laser beam to write and re-write the data. These disks may be used as alternatives to traditional disks. Erasable optical disks are based on a technology known as Magnetic Optical (MO). To write a data bit on to the erasable optical disk the MO drive's laser beam heats a tiny, precisely defined point on disk's surface and magnetizes it.

5. Flash drives: A USB flash drive is a data storage device that includes flash memory with an integrated Universal Serial Bus (USB) interface. USB flash drives are typically removable and rewritable, and physically much smaller than an optical disc. As at September 2013, drives of up to one-terabyte (TB) drive are available. Some allow up to 100,000 write/erase cycles, depending on the exact type of memory chip used, and a 10-year shelf storage time. USB drives with USB 2.0 support can store more data and transfer data faster than much larger optical disc drives like CD-RW or DVD-RW drives and can be read by many other systems such as the Xbox 360, PlayStation 3, DVD players

and in a number of handheld devices such as smart phones and tablet computers. A flash drive has no moving parts to be driven; therefore, it's not a true drive. The term *drive* persists because computers read and write flash drive data using the same system commands as for a mechanical disk drive, with the storage appearing to the computer operating system and user interface as just another drive. USB flash drives draw power from the computer via the USB connection.

1.8 INPUT/OUTPUT DEVICES

A computer is only useful when it is able to communicate with the external environment. When you work with the computer you feed your data and instructions through some devices to the computer. These devices are called Input devices. Similarly, computer, after processing gives output through other devices called output devices. We will discuss various types of I/O devices that are used for different types of applications. They are also known as peripheral devices because they surround the CPU and enable communication between computer and its environment.

1.8.1 Input Devices

Input devices are necessary to convert our information or data in to a form which can be understood by the computer. A good input device should provide timely, accurate and useful data to the main memory (RAM) of the computer for processing. The following are the most useful input devices.

i.   Keyboard: - This is the standard input device attached to all computers. The layout of keyboard is just like the traditional QWERTY typewriter. It also contains some extra

command keys and function keys. It contains a total of 101 to 104 keys. A typical keyboard used in a computer is shown in Fig. 1.5. You have to press correct combination of keys to input data. The computer can recognize the electrical signals corresponding to the correct key combination and processing is done accordingly.

Fig. 1.5: Keyboard

ii. **Mouse: -** Mouse is an input device that is used with your personal computer. It rolls on a small ball and has two or three buttons on the top. When you roll the mouse across a flat surface, the mouse cursor/pointer on the screen moves in the same direction. The cursor moves very fast with mouse giving you more freedom to work in any direction. It is easier and faster to get tasks done using the mouse especially with today's Graphical User Interface (GUI) based operating systems.

iii. **Scanner:** The keyboard can input only text through keys provided in it. If we want to input a picture the keyboard cannot do that. Scanner is an optical device that can input any graphics and display it back. The common optical scanner devices are Magnetic Ink Character Recognition (MICR), Optical Mark Reader (OMR) and Optical Character Reader (OCR).

a. **Magnetic Ink Character Recognition (MICR): -** This is widely used by banks to process large volumes of cheques

and drafts. Cheques are put inside the MICR. As they enter the reading unit the cheques pass through the magnetic field which causes the read head to recognize the character of the cheques.

b. **Optical Mark Reader (OMR):** This technique is used when students have appeared in objective type tests and they had to mark their answer by darkening a square or circular space by pencil. These answer sheets are directly fed to a computer for grading where OMR is used.

c. **Optical Character Recognition (OCR):** - This technique involves the direct reading of any printed character. Suppose you have a set of hand written characters on a piece of paper; you put it inside the scanner of the computer. This pattern is compared with a site of patterns stored inside the computer. Whichever pattern is matched is called a character read. Patterns that cannot be identified are rejected. OCRs are expensive though better the MICR.

1.8.2 Output Devices

i. **Visual Display Unit:** The most popular output device is the Visual Display Unit (VDU). It is also called the monitor. A keyboard is used to input data and a monitor is used to display the input data and to receive messages from the computer. A monitor has its own box which is separated from the main computer system and is connected to the computer by cable on desktop computers. In some systems like laptops and notebooks, it is compact with the system unit. It can be coloured or monochrome.

ii. **Printer:** It is an important output device which can be used to get a printed copy of the processed text or result on

paper. There are different types of printers that are designed for different types of applications. Depending on their speed and approach of printing, printers are classified as impact and non-impact printers. Impact printers use the familiar typewriter approach of hammering a typeface against the paper and inked ribbon. Dot-matrix printers are of this type. Non-impact printers do not hit or impact a ribbon to print they use electro-static chemicals and ink-jet technologies. Laser printers and Ink-jet printers are of this type. This type of printers can produce coloured prints and elaborate graphics.

Fig. 1.6 Laser Printer