Chapter 1: Introduction to Computer Hardware/Software

What are we studying in this chapter?

- **Introduction**
- **Computer Generations**
- **Bits, nibbles, bytes and words**
- **CPU**
- **Memory: Primary Memory, Secondary Memory**
- **Ports and Connections**
- **Input and output devices**
- **Computers in a network, Network hardware**
- **Software basics, software bytes**

1.1 Introduction

Computers are being used everywhere – in homes, offices, automobiles, schools and colleges, shops and so on. Now a days, majority of human beings from kids to old people whether literate or illiterate talk about *computers*, *computers,…*and *computers*. It is used in all the fields of engineering and non-engineering disciplines. So, it is necessary for us to know how to use the computer, how it works and how we can make use of the computer to do our activities by giving proper instructions. First, let us see the evolution of computers and then see *what is a computer and what are the features of computers.*

1.2 Computer generations

As the humans are evolved over millions of years and many generations, the computers also have evolved over the years and many generations. Each computer generation is marked by a major technological development both in size and speed. Now, let us see "What are various generations of computers and discuss the corresponding key features of computes in each generation?" The computer generations are classified as shown below:

- First generation computers (1940-1958)
- Second generation computers (1959-1964)
- Third generation computers (1965-1970)
- Fourth generation computers (1971-1990)
- Fifth generation computers (1990 and future)

1.2 💻 Introduction to Computer hardware/software

- **First generation computers:** The key features of first-generation computers are:
  - They used vacuum tubes for circuitry and magnetic drums for memory
  - They were big in size and each computer used to occupy the entire room space
  - They were very expensive and used to consume lot of electricity
  - The programmers used to write programs only in machine language to perform various operations.
  - They used to solve only one problem at a given time
  - Some of the examples are UNIVAC-I (**UNIV**ersal **A**utomatic **C**omputer), ENIAC (**E**lectronic **N**umerical **I**ntegrator **A**nd **C**alculator) etc.

- **Second generation computers:** The key features of second-generation computers are:
  - They used transistors for circuitry in place of vacuum tubes and magnetic core for primary memory and magnetic disks or magnetic tapes for secondary memory.
  - They were faster and smaller in size and are very cheap when compared to first generation computers
  - The programmers used to write programs in assembly language. Assemblers were built to convert from assembly language to machine language.
  - The high-level languages such as FORTRAN and COBOL were introduced using which various scientific and commercial problems can be solved

- **Third generation computers:** The key features of third-generation computers are:
  - They used integrated circuits where millions of transistors were placed on silicon chips.
  - They used very large memory to store the information using magnetic disks
  - They were much faster and much smaller in size and are very cheap when compared to second generation computers
  - Very easy to solve problems by writing programs using high level languages
  - They used operating systems to run many programs.
  - The keyboards and monitors were used to interact with the computers.
  - They consumed less energy and hence generated less heat and maintenance cost of computers were very less.

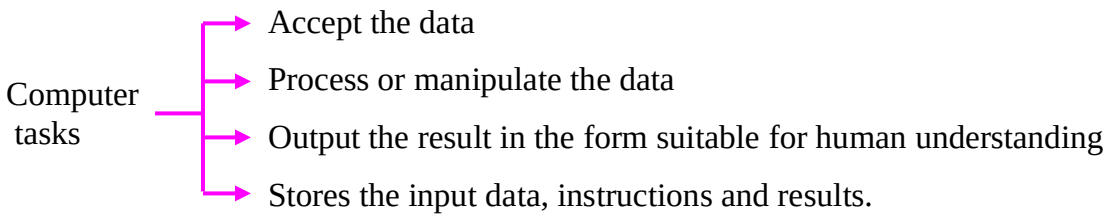
- **Fourth generation computers:** The key features of fourth-generation computers are:
  - They used large scale integrated circuits and microprocessors. Millions of components were placed on a small chip and hence size was very small.

- They had super computing power and execute billions and billions of instructions in one second.
- They used improved storage devices and storage capacity was in giga bytes and terra bytes when compared to earlier generations. IBM-PC and Apple's Macintosh were the examples of fourth generation computers
- They were very fast and smaller in size when compared to previous ones.
- As these computers were very powerful they were linked to form networks which eventually led to the development of the Internet. These generation of computers saw the development of Graphical User Interfaces, the mouse and hand held and smart devise such as smart phones.

♦ **Fifth generation computers:** The key features of fifth-generation computers are:
- They have very large storage capabilities, extremely high speed. The extreme high speed is possible using quantum computers. Google's D-Wave 2X quantum computer is 100 million times faster than today's machines.
- They are based on artificial intelligence that have the capability to think the way the humans think (machine learning)
- The use of parallel processing and superconductors makes the artificial intelligence a reality so that the computers can respond to natural languages such as English, Kannada etc.
- The main goal of artificial intelligence and machine learning is to do the activities that humans do and making the computers to interact with humans in the native language.
- Quantum computation and nanotechnology will radically change the face of computers in years to come.

1.2.1 What is a computer?

A computer is an electronic device that is used to accept the data, process the data and give the desired information so that people can understand. Thus, the main tasks that are carried out by any computer are:



1.4 💻 Introduction to Computer hardware/software

All the above said activities are performed based on the instructions that are stored in the memory of the computer. All these tasks are performed at tremendous speed with very high accuracy. For example, calculators, PCs, laptops, PDA (Personal digital assistant) etc., are all computers. **Note:** The word "computer" is derived from the word "compute".

1.2.2 Data/Information

We know that computer accepts the data. But, computers do not accept information. Now, let us see "*What is data? What is information?*"

Definition: The word *data* is the plural of *datum*. *Data* is a representation of facts or concepts in an organized manner. The data may be stored, communicated, interpreted, or processed from which we can draw conclusions. In short, we can say that *data* is a piece of information that is given. So, the data as such may not convey any meaning.

For example, numbers, names, marks etc are the data. Consider the two data items "Rama" and 10. Both these data does not convey any meaning to the reader.

Definition: *Information* is defined as a collection of data from which conclusions may be drawn. So, information is subset of *data* which when interpreted conveys meaning so that the people can understand. Hence, Information can be considered as a message that is received and understood. Information is obtained from data and it is the result of *processing*, *manipulating* and *organizing* the data which adds to the knowledge of a person.

For example, the output of the computer may be "Rama scored 10 marks". This sentence is an information. Because, it conveys meaning to the reader. The information may be instructions given to a person or commands issued to the computer.

1.2.3 ASCII

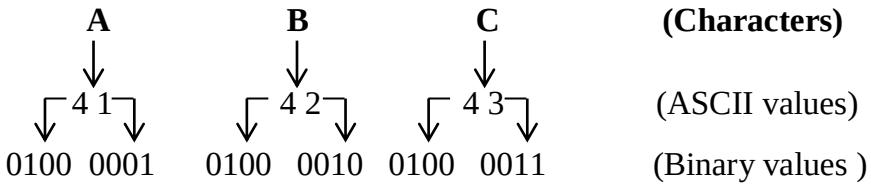
Now, let us see "*What is ASCII?*"

Definition: **ASCII** (pronounced as *as-key*) code is a 7-bit code to represent various characters such as alphabets, digits, punctuation marks and other symbols. **ASCII** stands for **A**merican **S**tandard **C**ode for **I**nformation **I**nterchange. Since, it is a 7-bit code, 2^7=128 symbols can be represented. The standard ASCII character set consists of 128 decimal numbers ranging from 0 through 127. The ASCII was established to achieve compatibility between various types of data processing equipments. Today,

ASCII character set is most commonly used in all types of computers. The ASCII Table is shown below.

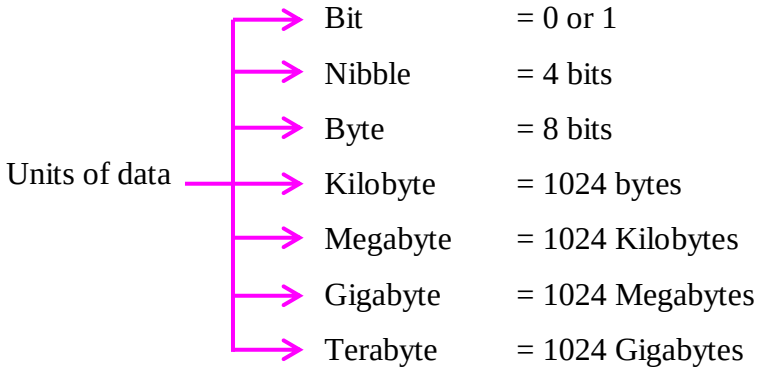
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2	SP	!	"	#	$	%	&	'	(	)	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{	|	}	~	DEL

Note: Observe from the table that character 'A' has an ASCII value 41 hexadecimal (65 in decimal), the character 'B' has an ASCII value 42 hexadecimal (66 in decimal) and character 'C' has an ASCII value 43. So, if we type the text "ABC" from the keyboard, to the computer, it appears as shown below:



1.3 Bits, bytes and words

All the quantities are measured in some units. For example, length may be measured in meters or feet. On similar lines, to measure computer memory, we require units. Now, let us see *"How does the data represented using 1's and 0's can be grouped or measured?"* The data represented can be grouped or measured using following units:



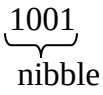
1.6 💻 Introduction to Computer hardware/software

First let us see "*What are bits?*"

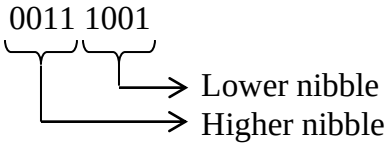
Definition: The data are represented by the state of electronic switches. A switch with ON state represents 1 and OFF state represents 0 (zero). These values which represent the state of a switch are called *bits*. In other words, the **bi**nary digi**t**s 1 and 0 are called *bits*. Thus, a *bit* is the smallest unit of data. Only bits are not used. The bits are grouped together to form a nibble or byte.

Now, let us see "*What is a nibble? What is a byte?*"

Definition: A group of 4 bits (half byte) is called a *nibble* or ½ byte.
Ex 1: An integer 9 can be represented using a nibble as shown below:

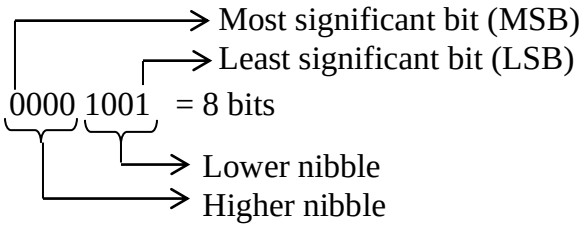


Ex2: The ASCII value of 9 is 39 (see the ASCII table in previous page) and it is expressed using 2 nibbles as shown below:

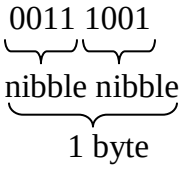


Definition: A group of 8 bits is called a *byte.* A byte is the smallest unit of data that can be manipulated by a computer at any given time. After accessing a byte, the bits or nibbles can be manipulated.

Ex 1: An integer 9 can be represented using one byte as shown below:



Ex 2: The ASCII value of 9 is 39 (see the ASCII table in previous page) and it is expressed using one byte (two nibbles) as shown below:



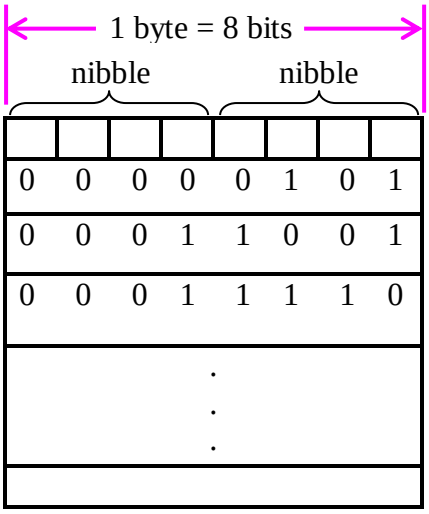
Note: 1 byte = 2 nibbles = 8 bits

Note: The computer cannot access a bit or a nibble. The smallest unit of information that can be accessed by a computer at any given time is one byte. After accessing a byte, the computer can manipulate the bits or nibbles.

Now, let us see how the data is stored in memory by taking a specific example. The memory is divided into various cells. A group of 8 cells we call a memory location. In each location one byte (8-bits) of data is stored. For example, let us take the following expression:

C = 5 + 25

This indicates that 5 and 25 are added and the result is stored in memory location identified by the variable C. Here, 5 and 25 are decimal numbers and computer cannot understand these two numbers. So, they are converted into binary form and are added as shown in the figure below:



Memory

decimal	Binary equivalent
5	0 0 0 0 0 1 0 1
+ 25	0 0 0 1 1 0 0 1
30	0 0 0 1 1 1 1 0

Note: Observe that the data is stored in memory in binary form only

Note: The data stored in the memory is normally measured using the number of bytes used to store the data. For example,

8 bits	= 1 byte	
1024 bytes	= 1KB	(Read as one kilobyte)
1024KB	= 1MB	(Read as one megabyte)
1024MB	= 1GB	(Read as one gigabyte)
1024GB	= 1TB	(Read as one terabyte)
1024TB	= 1PB	(Read as one petabyte)

So, 80GB = 80 x 1024MB
= 80 x 1024 x 1024 KB
= 80 x 1024 x 1024 x 1024 bytes
= 85899345920 bytes

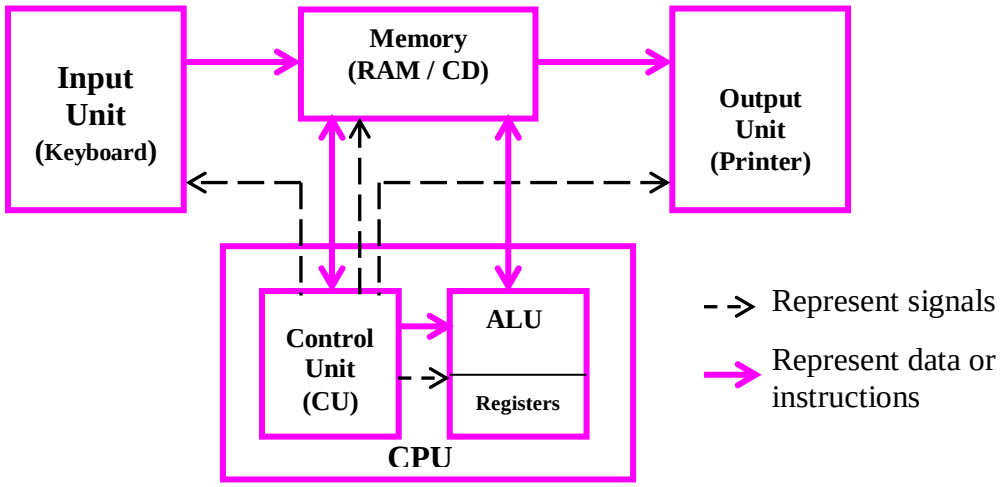
Now, let us see "*What is a word?*"

Definition: A **word** is the unit of information. Information may be stored, transmitted or operated on within a given computer. Most of the instructions are capable of manipulating the 1 word of data (2/4/8 bytes). The word is usually an even multiple of bytes such as 2/4/8/16 bytes etc. depending on the word length of the computer. For example,

- for a 16-bit machine, 1 word = 16-bits = 2 bytes
- For a 32-bit machine, 1 word = 32-bits = 4 bytes
- For a 64-bit machine, 1 word = 64-bits = 8 bytes

1.4 Basic structure of a computer

Now, let us *explain the basic structure of a computer.* The *basic structure of a computer* or *block diagram of a digital computer* and the communication between various units of computer system is shown in figure below:



Note: The dotted lines indicate the control signals issued by control unit and are single direction. The other lines may represent the *data* or the *instructions*. The lines connecting the various units represent a group of electrical wires. The group of wires is called a *bus*. Now, let us discuss each of the units one by one.

Input unit: The various input devices that are connected to the computer system are keyboard, mouse etc. Using the input unit, the user can perform the following activities:

- The user can enter the data or program (instructions) to the computer system or can move the cursor to the desired position and alter the data.
- It is the responsibility of the input unit to convert the data into a suitable form that can be understood by the computer.
- The converted data is stored in the memory in the form of 0's and 1's and then sent to central processing unit for further processing.

Output unit: The various output devices that are connected to the computer system are printer, monitor, speakers etc. Using the output unit, the processing unit can output the results or data to a form which is suitable for human understanding.

Memory unit: The memory unit can either be *primary memory* or *secondary memory*. The memory unit takes the data from input device and stores it until the computer is ready to process it. The memory unit is used:

- to store the instructions and data
- to store the intermediate results

If it is necessary, the data or results can be stored permanently using secondary storage devices such as hard disk drive or CD-ROM drive on storage media such as hard disk and CD.

Central Processing Unit (CPU) The **CPU** is the brain of the computer, which actually executes instructions one after the other to perform the desired actions. This activity is performed with the help of *control unit* and *Arithmetic logic unit*. The **control unit** controls and co-ordinates the activities of all the units by issuing proper commands. The **ALU** performs **A**rithmetic and **L**ogic operations.

Now, let us discuss CPU in detail.

1.4.1 Central Processing Unit

The computers accept the data, process the data and transform data into information. In this section, let us see "*Explain the various components of CPU?*"

The **CPU** is the brain of the computer, which actually executes instructions one after the other. The execution of each instruction is performed with the help of four basic parts:

- Control unit (CU)
- Arithmetic and logic unit (ALU)
- Registers
- Clock

Control Unit (CU): The **control unit** controls and co-ordinates the activities of all the units by issuing proper commands. The various responsibilities of the control unit are:

- Issue proper signals to the memory to obtain (Fetch) the next instruction to be executed.
- Identify what action to be taken (this is called decoding the instruction).
- Execute the instruction.
- Issue proper commands to the ALU, memory, input and output units to get the job done. These commands or signals are indicated by *dotted lines* (in the figure).

Arithmetic Logic Unit (ALU): This is the place where actual arithmetic and logical operations are carried out under the supervision of control unit. As the name indicates, **ALU** performs **A**rithmetic **L**ogic operations and hence the name **ALU(A**rithmetic **L**ogic **U**nit). The control unit issues signals to ALU to perform the following activities:

- The ALU performs four basic arithmetic operations such as addition, subtraction division and multiplications
- ALU also performs logical operations

The final results are sent to memory unit and then to the output unit if required by user.

Arithmetic Logic Unit (ALU): This is the place where actual arithmetic and logical operations are carried out under the supervision of control unit. As the name indicates, **ALU** performs **A**rithmetic **L**ogic operations and hence the name **ALU** (**A**rithmetic **L**ogic **U**nit). The control unit issues signals to ALU to perform the following activities:

- The ALU performs four basic arithmetic operations such as addition, subtraction division and multiplication
- ALU also performs logical operations

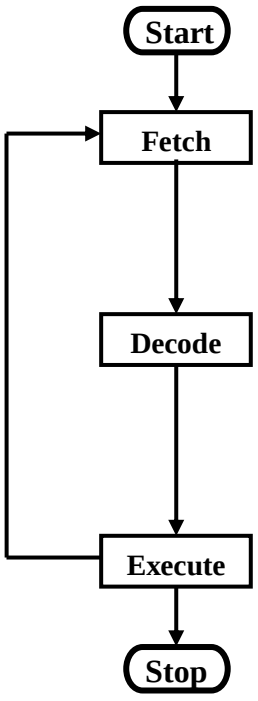
The final results are sent to memory unit and then to the output unit if required by user.

Registers: A register is nothing but a small high-speed memory inside the CPU which is used to store the results temporarily. The results obtained after performing arithmetic and logical operations are stored in these registers. There will be number of registers inside the CPU. These registers may hold the *data* or the *instruction* or the *address* of the instruction. There is one register called **Program Counter (PC)** which contains the address of the next instruction to be executed. Using this register, the CPU fetches/obtains the next instruction to be executed and executes the instruction.

Clock: *Clock* is a microchip inside the CPU that generates high and low signals called **pulses** which determines the speed at which the CPU executes the instructions. Every computer contains an internal **clock** that regulates the rate at which instructions are executed and synchronizes all other computer components. **Clock** speeds are expressed in terms of megahertz (MHz) or gigahertz (GHz).

1.4.2 Instruction cycle

The CPU executes the instruction using instruction cycle. Now, let us see "What is instruction cycle?" The instruction cycle also known as *fetch-decode-execute* cycle is the process by which every computer retrieves the instruction from memory, determines what action to be taken and carries out those actions. These actions are pictorially represented as shown below:



Fetch: Fetch (get) the instruction into *instruction register* from the memory using *program counter*. The *instruction register* and *program counter* will be inside CPU.

Decode: The instruction in *instruction register* is decoded to perform what action has to be done. The decoding is done using *control unit*. The *control unit* gives various signals to other units such as ALU to perform the action.

Execute: Once the necessary signals are issued the action is performed by various units inside the CPU. This is called execution. For example, if the instruction involves arithmetic or logic operations, the ALU is utilized to carry out these operations.

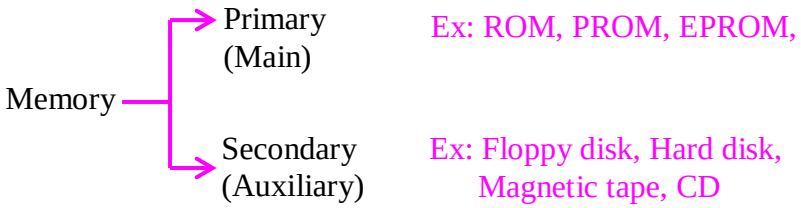
All the above actions *fetch*, *decode* and *execute* are repeatedly performed till the end of the program. The results obtained are stored in memory.

1.5 Memory

In this section, let us see "*What is memory? What are the different types of memories?*" *Memory* is defined as the place or locations where the data and/or the instructions are stored. The programs or the data can be permanently stored in memory so that they can be used whenever there is need. Memory is required for the following purposes:

- Memory is required to store both the data and instructions given to the computer to achieve a specific task
- Memory is also required to store the partial or complete results during processing.

The memory can be broadly classified as shown below:



Primary memory: The memory that is accessible directly by the CPU of a computer is called *primary memory.* This memory is part of the main computer system which is plugged into mother board along with CPU. Hence, this memory is also called *internal memory* or *main memory*. This primary memory allows the CPU to store and retrieve data quickly. The primary memory is very fast when compared to the secondary memory devices such as disks, CDs and tape drives.

Advantages

- The data or instructions can be accessed at very high speed
- To execute a program, at first step, all the instructions or data used by CPU have to be loaded into main memory. Otherwise, we cannot execute the programs

Disadvantages

- The primary memory is volatile i.e, the data or instructions stored in the primary memory will be lost as soon as the computer is turned off and the data cannot be retrieved back.
- Very expensive
- Huge amount of data cannot be stored

Secondary memory: The memory that is not part of computer's main memory and which is not directly connected to CPU but connected to the motherboard through ports and connectors is called *secondary memory.* In secondary memory we can store all the data and programs permanently and can be moved from one place to other place and can be connected to another computer. The *secondary memory* is also called *auxiliary memory*. The data and instructions are loaded from secondary memory to main memory so that the CPU can process the data.

Ex: floppy diskette, hard disk, CD-ROMs, magnetic tapes, Blu-ray disk, flash memory (Pen drive) etc.

Advantages

- The data stored in the secondary memory are not lost even when the computer is turned off
- Huge amount of data can be stored
- Very cheap
- They are portable i.e., they can be easily moved from one place to another place.

Disadvantages

- Accessing of data is very slow

Now, let us see *"What is the difference between primary memory and secondary memory?"* The difference between primary memory and secondary memory is shown below:

Primary memory	Secondary memory
1. Primary memory is internal memory and is also called main memory	1. It is an external memory of the computer
2. Accessing the data is much faster	2. Accessing the data is slower
3. Expensive	3. Less expensive
4. Temporary storage	4. Permanent storage
5. Huge amount of data cannot be stored	5. Huge amount of data can be stored
6. Semiconductor memory	6. Magnetic memory

Now, let us discuss primary memory and secondary memory in detail.

1.5.1 Primary memory

Now, let us see "*What is main/primary ,memory? What are the different types of primary memory?*"

Definition: The memory that is accessible directly by the CPU of a computer is called *primary memory*. This memory is part of the main computer system which is plugged into mother board along with CPU. Hence, this memory is also called *internal memory* or *main memory*. This primary memory allows the CPU to store and retrieve data quickly. The primary memory is divided into following types:

- Random Access Memory (RAM): SRAM and DRAM
- Read Only Memory (ROM, PROM, EPROM, EEPROM)
- Cache memory (L1, L2 and L3)
- CPU registers

1.5.1.1 Read only Memory (ROM)

Definition: **ROM** stands for **R**ead **O**nly **M**emory. ROM signifies the permanent memory that can only be read and but not written. The contents of ROM are written during manufacturing time of ROM. Accessing of data/instructions are much faster in ROM when compared to RAM. ROMs are non-volatile as the contents of ROM are not lost when the computer is switched off.

ROM normally contains a special program called Basic Input Output System (BIOS). BIOS is a small program that initiates the startup process when the computer is switched on. The BIOS will transfer the control to the operating system.

Now, let us see "What are the different types of ROMs" The ROMs are classified as shown below:

- Programmable Read Only Memory (**PROM**)
- Erasable Programmable Read Only Memory (**EPROM**)
- Electrically Erasable Programmable Read Only Memory (**EEPROM**)

PROM: It is a special type of ROM where the data/instructions are not written into ROM during manufacturing. Initially it is blank. Later, using a device called PROM programmer we can burn the data/instructions into ROM. But, once written contents cannot be changed.

EPROM: It is a special type of ROM where the data/instructions are not written into ROM during manufacturing. Once the data/instructions are written into this ROM, is can be rewritten only once. An EPROM is erased by exposing it to ultra-violet radiation.

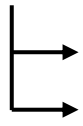
EEPROM: It is a special type of ROM where the data/instructions are not written into ROM during manufacturing. Once the data/instructions are written into this ROM, is can be rewritten multiple times. The Pen-drive that we use is EEPROM. Since the can be written any number of times, EEPROMs are more like RAMs rather than ROMs.

1.5.1.2 Random Access Memory (RAM)

RAM stands for **R**andom **A**ccess **M**emory. The programs that have to be executed should be copied from secondary memory to primary memory. The CPU then fetches the instructions from the main memory executes the program. The results obtained will be temporarily stored in RAM. In multi-programming environments where more than one application is running or in a multi-user environment, RAM contains several programs and each program will be executed one by one in fixed time intervals. RAM is an integrated chip(IC) consisting of densely packed transistors where data or instructions are stored and the entire RAM IC can be fixed on a slot in the mother board.

The data can be stored in the RAM as long as the computer is on. Once the computer is off, the data stored in the RAM will be lost and cannot be retrieved. So, RAM is a *volatile memory*.

Now, let us see "*What is volatile?*" The memory that loses its contents when the computer is turned off is called *volatile* memory. They hold the data temporarily as long as the computer is on. Since the CPU can read the data from the memory or can write the data into the memory, these memories are called *read-write* memories. The RAM is classified as shown below:



- Static RAM (SRAM)
- Dynamic RAM (DRAM)

Static Ram(SRAM): Here, multiple transistors are used for each cell to store one bit of information. So, an integrated chip of SRAM will take more space. SRAM is faster, bigger and more expensive than DRAM.

Dynamic Ram(DRAM): DRAM uses one transistor for each cell to store one bit of information. So, an integrated chip of DRAM will take less space. In DRAM capacitors holds the charge but because of internal leakage, DRAM cells have to be refreshed every two or three milli-seconds.

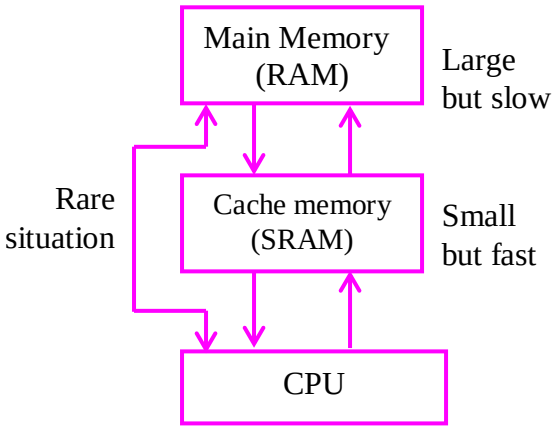
Now, let us see "*What is the disadvantage of RAM (normally DRAM)?*" and "*How to overcome this disadvantage?*" The CPU is the place where computations are performed. So, after the data is obtained from RAM, the arithmetic or logic operations are performed by the CPU and the result is transferred back to RAM. Moving data between RAM and CPU is one of the most time-consuming operations a CPU must perform. This is because, accessing the data/instruction from RAM is very slow when compared to the speed of the CPU. This time-consuming movement of data between the CPU and RAM can be overcome using *cache memory*.

1.5.1.3 Cache memory

In this section, let us see "What is cache memory?" A *cache memory* is an intermediate storage placed between the CPU and the main memory. It is very fast and small memory. This is used to store the frequently used data and instructions during processing. Since cache memory is several times faster than main memory, it is very expensive and hence size of the cache memory is small.

Working: The figure shows how cache memory works with CPU and RAM. When a program is running and if CPU wants to read the data or next instruction, the following actions are taken.

- CPU first checks whether the data or instruction is available in the cache memory
- If data or instruction is not there in cache memory, the CPU reads data or instruction from main memory into the registers of CPU. The same copy is copied into cache memory
- If same data or instruction is needed for the next time, CPU reads from the fast cache memory instead of slow RAM.

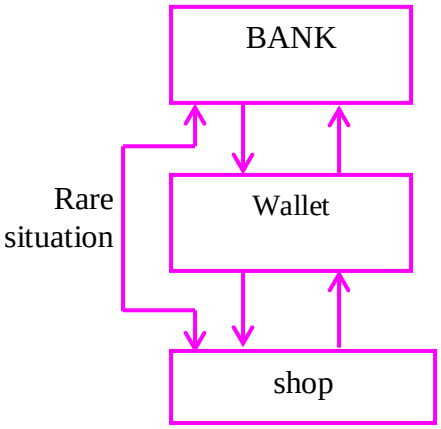


Thus, using cache memory we can save the time needed to load the data from RAM thereby increasing the efficiency of CPU and execute the program as fast as possible. Cache memory can be further categorized based on where cache is placed as shown below:

- L1 (Level 1) cache: It is known as primary cache. It is built into the CPU chip and works at the speed of CPU. It can work with L2 (Level 2) cache for much faster data accessing.
- L2 (Level 2) cache: It is known as secondary cache. It is built into motherboard outside the CPU. Usually, this is in between RAM and L1 cache. It is larger when compared to L1 cache and but smaller than RAM. So, it is slower than L1 cache and faster than RAM.
- L3 (Level 3) cache: It is not used normally. But, if used it can further speed up the execution process.

Note: The cache memory operation can be explained using the following analogy. If you observe, the two figures are similar.

We keep cash in wallet since it is faster and easier to access. When we want to spend, we directly get cash from the wallet. In case, if cash is not there, we go to bank, get the money and spend. At the same time, we keep some money in the wallet so that we can spend later. Here, cash = data, BANK = RAM, shop owner = CPU, Wallet = Cache memory



1.5.1.4 Registers

In this section, let us see "What is a register?" Registers are temporary storage locations inside the CPU that hold data, addresses and instructions. Using these registers the CPU does all its work. The registers are integrated into the CPU and hence they are fastest memory of the computer. The size of each register is word length of the computer. For example,

	16-bit machine	32-bit machine	64-bit machine
Size of register	16-bit	32-bit	64-bit

In Intel 32-bit microprocessor the registers are named as EAX, EBX, ECS, EDX etc.

1.5.2 Secondary Memory

All the disadvantages of primary memory are overcome using secondary memory. Now, let us see *"What is secondary memory?"*

Definition: The memory that is not part of computer's main memory and which is not directly connected to CPU but connected to the motherboard through ports and connectors is called *secondary memory.* In secondary memory we can store all the data and programs permanently and can be moved from one place to other place and can be connected to another computer. The *secondary memory* is also called *auxiliary memory*. The data and instructions are loaded from secondary memory to main memory so that the CPU can process the data.

Ex: floppy diskette, hard disk, CD-ROMs, magnetic tapes, Blu-ray disk, flash memory (Pen drive) etc.

Now, let us see "What are the different types of storage devices?" The different types of storage devices are shown below:

- Floppy disk (1.2MB and 1.44MB) which is obsolete
- Hard disk (500GB to 4TB)
- Magnetic Tape (20TB)
- Optical discs: CD-ROM, DVD-ROM, Blue-Ray Disk
- Flash memory

1.5.3.1 Diskettes (Floppy disks)

In this section, let us see *"What is a floppy disk or diskette?"*

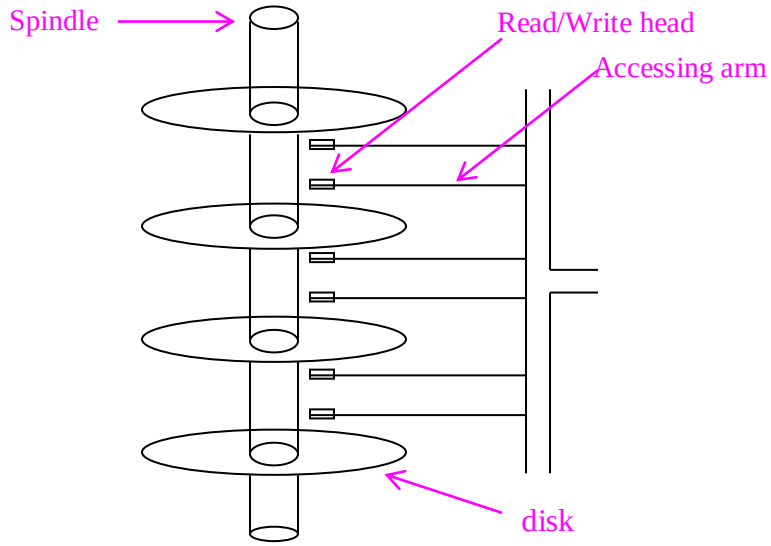
Definition: Floppy disks were invented by IBM as a means of loading programs into a computer. A *floppy disk* also called *diskette* is a storage media consisting of thin flexible circular plastic material which is coated with a magnetic material and enclosed in a rigid rectangular plastic case. Floppy disks once were useful to transfer the data between computers and store the files. Because of very small capacity of a floppy diskette, now they are not used and have become obsolete.

1.5.3.2 Hard disks

All the disadvantages of floppy disk can be overcome using another important storage media called *hard disk.* In this section, let us see *"What is a hard disk?"*

Definition: *Hard disk* is the primary storage unit of the computer. Huge amount of data can be stored and accessed in few milliseconds. The hard disk consists of more number of disks arranged in cylindrical pattern one above the other on a common *spindle*. Each disk is made up of aluminum alloy coated with magnetic material and both the surfaces are used for storing the data.

The read/write heads are attached to a single access mechanism so that they cannot move independently. All read/write heads are moved together to position the heads on the required track. But, at any point of time one of the read/write heads is used to access the data. The read/write head can move horizontally inwards and outwards using the access mechanism while accessing the data. The entire unit is enclosed in a sealed chamber. The hard disks available today have 500GB to few terrabytes of storage. Most entry-level consumer PCs now come with hard disks of at least 500GB.



The present day hard disks in most PCs spin at speed of 3,600 to 10,000 rpm. The speed at which the disk spins determines the performance of the hard disk. This high rotational speed allows more data to be recorded on the disk surface. Hence, they are capable of storing more data and accessing of data is much faster. The outer surface of the first disk is not used for storing the data.

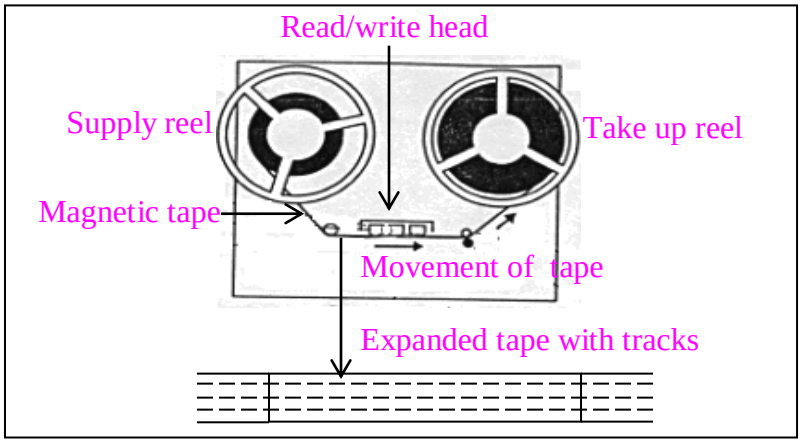
1.5.3.3 Tape drive (Magnetic tape)

In this section, let us see *"What is a magnetic tape?" What are the advantages and disadvantages of magnetic tape?"*

1.20 💻 Introduction to Computer hardware/software

Definition: A *magnetic tape* is a secondary storage device consisting of a thin tape coated with magnetic material. This tape is used for recording both analog and digital data. Magnetic tape recording uses the same principle as it is used in audio-cassette recorder. But, the difference is that a computer tape drive writes digital data in terms of 0's and 1's instead of analog data (continuous signals). It is a serial access medium and hence data cannot be quickly located. The magnetic tapes are used to store bulk data for backup purposes. [Backup => store a copy of file, directory etc on a separate storage device from the original. This helps to retrieve the data if original data is lost.] The magnetic tapes of 200TB capacities are expected to launch in near future.

Working The magnetic tapes are similar to the cassettes that we use in our tape-recorder. The various components of tape along with usage of tape for reading and writing is shown below:



Supply reel, Take-up reel: Each magnetic tape uses two reels namely *supply reel* and *take-up reel*. The tape moves from supply reel to take-up reel.

Magnetic tape: The tape is coated with a magnetic material. The width of the tape ranges from 0.38 cm to 1.27 cm. The tape is mounted on a spindle for rotating during reading and writing. A magnetic tape can be have 7 to 9 tracks. The tape is divided into parallel tracks running length-wise as shown in the expanded portion of tape.

Read/write head: The magnetic tape passes directly over the read/write head making contact with the head as shown in figure. As the tape passes under read/write head, the data can be read into main memory or data can be written onto the tape. Normally, 8 bits of data is read from the tape or 8-bits of data is written onto the tape.

A tape drive is a sequential device. So, if the tape head is positioned at first record, then to read n^{th} record, it is necessary to read all the records from 1 to n-1. The tape is in motion only during read and write operation. Now, let us see *"What are the advantages and disadvantages of magnetic tapes?"*

1.5.2.4 Optical storage devices

We have seen the working of magnetic storage devices. In recent years, optical disks are being used widely. In this section, let us see *"What are optical storage devices?"*

Definition: An *optical storage medium* is a flat, circular storage medium which is coated with a thin metal on which bits may be stored in the form of highly reflexive area or less reflexive area. The data can be stored in the optical storage media or read from the optical storage media. The devices which perform read or write operation on optical storage media are called *optical storage devices.* The data is read when the disc is illuminated with laser beam. These devices are capable of storing enormous amount of data in a limited amount of space. Since laser technology is used to read the data or write the data, these devices are also called *laser disks* or *optical laser disks*.

For example, CD-ROM, DVD etc. are optical storage devices.

Now, let us see *"What are various types of optical storage devices?"* The various optical storage devices are classified as shown below:

- CD
- DVD
- Blue-Ray Disc

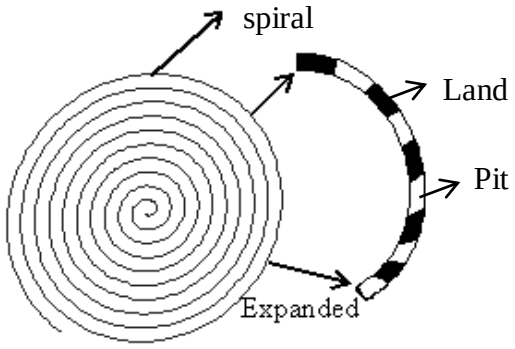
4.4.1 CD-ROM

In this secction, let us see *"What is a CD-ROM? How the data can be stored or retrieved?"* and *"What are the advantages and disadvantages?"*

Definition: A ***CD-ROM*** stands for **C**ompact **D**isk-**R**ead **O**nly **M**emory. A CD is a small optical disk on which data such as music, text or graphic images can be stored digitally (in the form of 0's and 1's). The CDs are shiny, silver color metal disk of 12cm in diameter. The CDs have become the standard medium for distributing large quantities of information. These are originally developed for storing digital audio.

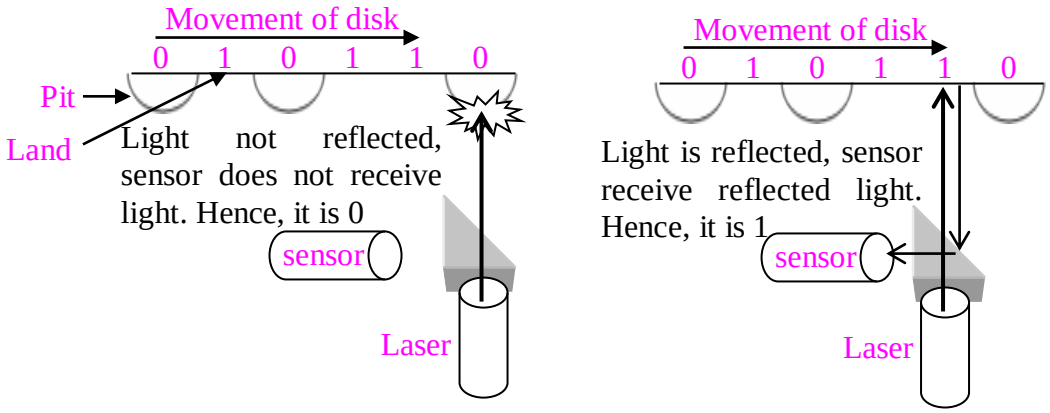
Like hard disk, a CD can store huge amounts of digital information permanently. The data stored on CD cannot be erased and it is the most reliable storage medium. The capacity of CD range from 650MB to 700MB and is inexpensive to manufacture. Unlike hard disks the CDs are very portable and handy to use.

Construction and working: A CD is made up of plastic polycarbonate which is coated with highly reflective material. A CD-ROM has only one track that spirals from the centre to the outside edge as shown below. This single track is divided into number of sectors of equal length. The data is stored in these sectors. Now, let us see, *"How the data is stored in CD?"*



The data in the form of 1's and 0's is written into the disk in a spiral manner. Laser beam is used to burn the CD's forming pits and lands. A pit is a *hollow* or *depression* caused during burning.

When light falls on the pit, light is not reflected. The non-burnt area is called *land* and represent 1. Thus, the data is stored in a CD in the form of *pits* and *lands* as shown in the figure. Now, let us see, *"How the data is read from CD?"* The laser beam is passed over the spiral as shown below:



We know that data is stored in the form of pits and lands. During reading, when the laser beam falls on the pit, the light will be scattered as shown in the figure and sensor will not receive the reflected light. This is interpreted as 0. As the disk rotates (look at the next figure), when the laser beam falls on the land, the light is reflected and is passed to the sensor using a prism. The sensor receives this reflected light and it is interpreted as bit 1. The sensor passes this information of 1's and 0's to the CPU. Thus, a CD reads the information from CPU.

4.4.2 DVD-ROM

Today's new PCs comes with built in DVD-ROMs. Now, let us see *"What is DVD-ROM?"* and *"What are the advantages"*

Definition: A ***DVD-ROM*** stands for **D**igital **V**ideo **D**isk - **R**ead **O**nly **M**emory. The DVDs looks like CDs and can play ordinary CD-ROM discs. A DVD is a small optical disk, high density medium capable of storing a full-length movie on a single disk. This high density is achieved:

- By using both sides of the disk
- By using special data-compression technology
- By using extremely small tracks to store the data.

Advantages

- Storage capacity is more when compared to CDs (around 9.4GB)
- The DVD player can play an ordinary CD.

Flash memory

In this section, let us see *"What is flash memory?"* Flash memory is an electronic non-volatile computer storage medium that can be electrically erased and reprogrammed any number of times. This is a form of EEPROM and it is available in various forms such as *pen drive, sloid state disk (SSD)* and *magnetic card* (SD Card) and they are normally used to transfer the data between personal computers and other digital devices. The flash memories are portable (can be carried or used transfer data between any digital device), small in size, they need little power and are quite reliable. The capacities of flash memories are increasing exponentially and their cost is being decrease drastically.

The memory stick also called *pen drive* is mostly commonly used in the industry and when connected to the computer via USB port, it is detected as a separate drive and files can be transferred to/from pen drive very easily.

1.6 Ports and connectors

In this section, let us see "What is a port? what is a connector?" A port is a connection point between a computer and a device (external/internal). The port has either holes or slots so that devices can be connected. Internal ports may connect devices such as hard disk drives, CD-ROM/DVD drives. External ports may connect modems, printers, mouse, keyboard and other devices.

A device that physically connects two things together is called a connector. A male connector has pins that are exposed and female connectors have holes in which the male connector can be inserted.

Now, let us see *"What are various ports that are supported by a computer?"* The various ports that are normally available are shown below:

- Universal Serial Bus (USB): The USB port has virtually replaced all serial and parallel ports in the motherboard. Majority of computers and laptops nowadays will have minimum four USB ports that can support scanners, printers, mouse etc.
- Serial port: These ports were used by keyboard, mouse, modems and were offered in 9 and 25 pin configurations. Data passes through this port one bit at a time.
- Parallel port: Earlier parallel ports were used to transfer the data to a printer. Here, the data is transferred parallelly.
- Video graphics Adapter (VGA): It is a 15-pin port and allows transfer of analog video to the monitor.
- Rj45 port: This port is used by Ethernet network.
- PS/2 port: Earlier keyboard and mouse was connected to the computer using this port. These ports are not supported in any of the laptops. In laptops using USB ports, the mouse is connected.
- High Definition Multimedia Interface (HDMI): The audio and video signals are transferred using this port. Using this port, we can connect a projector or a TV so that output is visible on the screen.

1.7 Input Devices

Now, let us see "*What are input devices? What are the various input devices?*"

Definition: *Input devices* are the external devices that are connected to CPU using which the data or the commands can be entered into computer. The various input devices that we discuss in this section are shown below:

- ♦ Keyboard
- ♦ Pointing devices
- ♦ Scanners

1.7.1 Keyboard

In this section, let us see "*What is a keyboard?*"

Definition: A keyboard is the primary input device used in all computers. Keyboard has a group of switches (also called key-switches) resembling the keys on an ordinary typewriter machine. Associated with every key-switch is a character. Using the keyboard, the user can type data or the commands.

Normally, the keyboard has around 101 keys. The keyboard includes keys that allows us to type letters, numbers and various symbols such as *, / , [, %, etc., that allow user to enter data and instructions into the computer.

The most commonly used keyboard layout is called QWERTY layout. Since the first six alphabetical characters from left to right are Q, W, E, R, T, Y (towards right of Tab key)it is called QWERTY layout. Even though keyboards come in different sizes and styles, their layout (arrangement of the keys on the board) almost remains same.

1.7.2 Pointing devices

Now, we shall see "*What is a pointing device? What are the various types of pointing devices?*"

Definition: A *Pointing device* is an input device using which we can control the movement of a pointer on the screen. Using this device we can communicate with the computer by pointing to any desired location on the screen, we can select the items on the screen, and give instruction by clicking appropriate buttons which are part of the device. Some of the pointing devices are listed below:

Definition: A *mouse* is a hand-held pointing device. It acts as an input device. It is designed to fit comfortably under the palm. This input device is used to control the movement of the pointer on the screen. It is also used to make selection from the screen by pointing to the item. It will be connected to the computer using the cable.

By appearance it looks like a *mouse*, the cable similar to the tail of the mouse and hence the name "mouse".

The top of the mouse usually has 2 buttons (left and right) along with a scroll wheel as shown in the figure and a rubber ball situated at the bottom of the mouse. Now, we shall see "*How to use the mouse?*"

How to use? Mouse is used to point to any required position on the screen. We can hold the mouse in our palm and slide the mouse without lifting it along a hard flat surface. To move the cursor to the desired position, various actions to be done are shown below:

1. Move the mouse forward – the pointer also moves up

1. Move the mouse to left and right– the pointer also moves left and right

- Move the mouse diagonally up and down – the pointer moves diagonally up and down

In general, whichever direction the mouse moves on the hard surface, correspondingly, the mouse-pointer also moves on the screen. So, the movement of pointer on the screen is directly proportional to the movement of mouse on the flat surface. Once the mouse is placed at the desired location on the screen, the various actions usually performed are:

How mouse works? When the mouse is moved, the ball which is situated at the bottom of the mouse rolls. This movement is converted into electrical signals and is transferred to CPU. These movements are interpreted and displayed on the screen in the form of an arrow. This arrow is called mouse pointer. So, moving the mouse in any direction on the hard flat surface, moves the arrow in that direction on the screen. Thus, mouse can point to any area on the screen.

1.7.3 Data scanning devices (Optical input devices)

Now, let us see "*What are optical input devices?*"

Definition: The device which uses light as a source of input for detecting or recognizing various things is called an *optical input device* (or *optical reader* or *data scanners*). Using these devices, information such as characters, marks, codes etc, can be read. This information is converted into digital data and sent to computer for further processing. The various types of *optical input devices* (or *optical readers*) are classified as shown below:

- The hand-held scanners (bar code readers) are used to identify items in shops and malls.
- The modern scanners have the *optical character recognition* (OCR) facility by which a document (or its image) can be scanned and can extract the text as a stream of characters.
- Banks use Magnetic Ink Character Recognition (MICR) scanners to read the codes printed on bank cheques.

1..8 Output Devices

Now, let us see "*What are output devices? What are the various output devices?*"

Definition: *Output devices* are *electromechanical* devices that accept the data from the computer and convert the data into a form that can be understood by people. In other words, the devices that are used to convert machine-readable information (which is in terms of 0's and 1's) into human-readable form are called *output devices.* The various classifications of output devices are shown below:

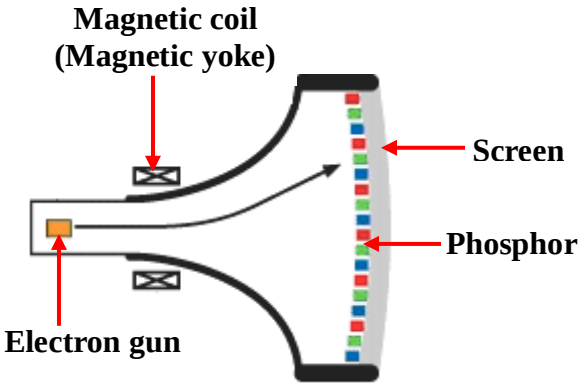
- Monitors
- Impact printers
- Non-impact printers
- Plotters

1.8.1 Monitors

Now, let us discuss how monitors are used as output devices. In this section, let us see "*What are monitors? What are the various types of monitors?*"

Definition: *Monitors* are the display devices that are housed in a plastic or metal case using which we can display text or image. The monitors are used to view the input typed using the keyboard. The output can be displayed and viewed on the computer screen (TV like display unit).

CRT (**C**athode **R**ay **T**ube) monitor: It is a TV like display attached to the computer on which the output is displayed. The CRT monitors are normally used in majority of the desktop computers. The major components of the CRT monitor are shown in the figure below:



- **Electron gun**: An electron gun situated at the back of the monitor generates a beam of electrons. The guns emit electrons to create images on the screen.
- **Deflection plates**: The electron beam is directed towards the screen in front of the tube using the deflection plates called *magnetic coil* or *magnetic core*.
- **Phosphor**: The back of the screen is coated with phosphor, a chemical that glows when it is hit by the electron beam.
- **Screen**: Since the light emitted by the phosphor fades very rapidly, it is necessary to redraw the picture repeatedly on the screen thus displaying a continuous image.

LCD monitors (**L**iquid **C**rystal **D**isplay): They are flat panel monitors and are widely used. The LCD monitors use a special kind of liquid crystals to display the images on the screen. The liquid crystals that are used in these monitors are normally transparent. But, when charged with electricity they become opaque. The advantages and disadvantage of flat panel monitors are shown below:

1.8.2 Printers

A *printer* is an output device that is used to print the text and graphical information on a physical medium (that we can touch and carry) such as paper. The document printed on the paper or a transparency film is more permanent form of output than we see the results on the monitor. So, printed information on paper or transparency film is called *hardcopy*. The printers are mainly classified into two categories:

- Impact printers
- Non-impact printers

1.8.2.1 Impact printers

In this section, let us see "*What is an impact printer? What are the advantages and disadvantages of impact printers?*"

Definition: An *impact printer* is a printer which works by physically striking an ink ribbon that is placed on the paper. Normally, the print head acts like a hammer which has pins that take the shape of the letter or character and when the hammer strikes the ribbon, the shape of the character is transferred through the inked ribbon onto the paper creating a printer character.

For example, *dot matrix printer*, *daisy-wheel printer*, line printer etc are some of the impact printers.

Dot Matrix Printer: A *dot matrix printer* is a character printer that prints all types of characters and images as a pattern of dots by striking against the ink ribbon. The hammer is a *print head* consisting of pins representing the character or the image. When these pins hit the ink ribbon, the character or the image appears on the paper. Usually the printers are bi-directional i.e., first the printing will be done from left to right and then from right to left. Since they can print only dots, they are called *dot matrix* printers. Thus, any character can be printed using the pattern (matrix) of dots.

Line printer: A *line printer* is a special type of impact printer. The working is similar to that of dot matrix printer. But, it uses a special print head so that entire line of text is printed at once. Even though resolution is very less (i.e., quality of characters printed on the paper is not that good), but are incredibly faster and can print 2500 to 3000 lines of text per minute.

Daisy wheel printer: This printer is obsolete and uses a different technology to strike the ribbon. It uses a wheel with separate characters distributed along its outer edge. The characters are thus pre-formed and not generated. To print a character, the wheel is rotated so that the desired character directly faces the ribbon.

1.8.2.2 Non-impact printers

In this section, let us see "*What is a non-impact printer? What are the advantages and disadvantages of non-impact printers?*"

Definition: A *non-impact printer* is a printer which forms the characters and images without making the physical contact between the printing mechanism and the medium on which printing takes place (normally paper). The striking mechanism and ribbons are not used here. Instead, the ink is sprayed against the paper in the form of a character. In some type of machines, the heat and pressure are used to fuse the toner powder in the shape of a character as used in Xerox (photocopying) machines.

For example, *ink jet printers, laser printers* etc., are some of the *non-impact printers.*

Inkjet printer: *Inkjet printer* is a printer that forms all types of characters and images by spraying small drops of ink on the paper. All inkjet printers are non-impact printers. Single colored inkjet printers can only print black text and images whereas the color inkjet printers can print even the quality photos. Here, the character is produced by spraying small droplets of ink through tiny nozzles. Instead of print head, the inkjet printer uses ink nozzles. Since many nozzles are used for spraying the ink, the quality is far superior compared to that of dot matrix printer. Each nozzle consists of 50 to several hundred small holes for spraying the ink. The ink can be sprayed either by heat or pressure.

The commonly used ink jet printers are HP, Epson and Canon printers. The most ink-jet printers support from 360 to 1440 dots per inch (dpi) thus producing high quality characters. For printing, it is better to use the high quality paper. If low quality paper is used, the paper may absorb the ink and the ink may spread till it dries out. Color ink jet printers are used for taking the printouts with varying colored text, graphical images and quality photos.

Laser Printer: A *laser printer* is a high-speed, high-quality *non-impact* printer which operates in a manner similar to a photocopying machine using laser technology. It uses laser beam to form the images. These images are transferred to paper *electrostatically*.

Working: The laser printer has its own CPU and memory built into the printer. The CPU inside the printer is used to interpret the data received from the computer and to control the laser. The laser printer works as shown below:

1. The laser beam moves across the drum using the mirror and creates an image on the drum in the form of electrical charge.
2. The electrical charge on the drum attracts the toner particles (The toner consists of tiny particles of oppositely charged dry powder type of ink and hence are attracted towards drum) thus forming the complete image to be printed on the paper
3. The paper is taken from the tray and is passed across the drum so that the image created on the drum is transferred on to the paper

4. The toner particles transferred on to the paper are imprinted or fused permanently on the paper using heat and pressure. For this purpose, heaters and rollers are used.
5. The paper comes out of the printer. Thus, one full page is printed once. For this reason, the laser printers are also called *page printers.*
6. After the document is printed, the electrical charge is removed from the drum, the drum is rotated and cleaned using rubber blade and the excess toner is collected. Now, the drum is ready for the next page printing.

Now, let us *"Compare dot matrix printer and laser printer"* The differences between dot matrix printer and laser printer are shown below:

Dot matrix Printer	Laser Printer
1. Impact printer	1. Non impact printer
2. Slow printing and very cheap	2. Fast printing and costlier
3. Carbon copies can be taken	3. Carbon copies can not be taken
4. Produce noise	4. Do not produce much noise
5. Print quality is not good	5. Print quality is very good
6. Prints only text	6. Prints both text and graphics
7. Print using standard font (normal, italic or bold)	7. Print using fonts of varying thickness and forms

1.8.2.3 Plotters

Now, let us see "What is a plotter?"

Definition: A plotter is a large-format inkjet printer used to produce high quality color graphics using electrostatic charges. They are generally used in technical drawing and **C**omputer **A**ided **D**esign (CAD) applications. They are used in engineering applications where precision is mandatory. The plotters are very large and huge and their length will be in few meters.

Printing a drawing size 4ft x 6ft

These large-format printers are not called large printers, but are generally called plotters only. They are used to work on very large paper sizes while maintaining high resolution. Plotters are more expensive than printers. The plotters are also used to print very large banners with high quality resolution as shown in figure.

Printing a banner of size 4ft x 10ft

1.9 Computers in a network

Even though stand-alone computers are powerful and can be used for variety of purposes, they have their own limitations. Some of the limitations of stand-alone computers are:

- Difficult to share the data between the computers
- Difficult to share the peripheral devices

These disadvantages were overcome using *Networks*. Now, let us see *"What is a network?"*

Definition: A *network* or *computer network* is an arrangement where two or more computers are connected to share the data as well as resources such as hardware and software using communication media. Computers in a network may be connected using telephone lines, cables, satellite links etc. Wireless networks also allow computers to exchange information by radio signals. Computer network help the user to communicate, exchange information and share various devices that are connected to the network.

For example, computers can receive e-mail, send files and instant messages to each other.

1.9.1 Common types of networks

The computer networks are classified into two main categories:

- LAN (Local Area Network)
- WAN (Wide Area Network)
- Metropolitan Area Network (MAN)
- Campus Area Network (CAN)
- Personal Area Network(PAN)

LAN: *LAN* stands for Local Area Network. A *local area network* is defined as a computer network which operates over short distances at very high speed. As the name indicates, the connection of two or more computers is limited locally to a single organization. The computers are connected via cables for transmitting the data. The range between two end points is limited from few meters to less than 2 kilo meters.

WAN: *WAN* stands for Wide Area Network. A *Wide Area Network* is a defined as a computer network which operates over long distances covering larger geographical area. It consists of two or more LANs. The scope of WAN is not limited to a building or campus like LAN. But, spans over wide area. So, communication is mostly through telephone or satellite links. Big organizations and companies, which have several branches all over the world, are the ones who greatly benefit from WANs. The WANs enable them to easily communicate with their partners or employees around the globe.

MAN: Metropolitan Area Network is a type of network whose power is in between an intermediate between LAN and WAN. A MAN is employed for interconnecting computers in the same city.

PAN: Personal Area Network is smallest network of all which covers a distance of few meters. It connects small devices like cellphones and laptops with infrared or Bluetooth technology.

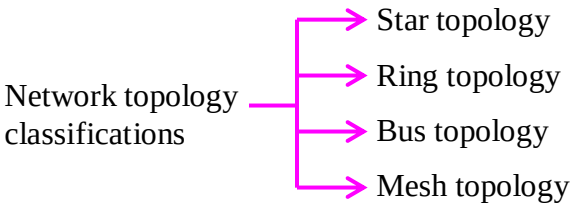
Now, let us see *"What is the difference between LAN and WAN?"* The difference between LAN and WAN is shown below:

LAN	WAN
1. Interconnection of computers within a building or campus	1. Interconnection of computers across countries and continents
2. Data transfer rate is much higher.	2. Data transfer rate is relatively less
3. Fast data transfer because of short distance	3. Slow data transfer because of large distance
4. Error rate is very less	4. Error rate is slightly higher
5. Reliable and less expensive	5. Reliable and more expensive
6. Communication is through cables	6. Communication is through telephone or satellite links
7. It is a private property	7. It is not private property

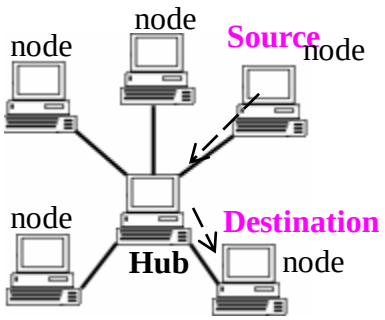
1.9.2 Network topologies

In this section, we concentrate on network topologies. So, first we shall see *"What is a topology?"*

Definition: Based on the way various devices are connected, the layout changes. A *network topology* is defined as the layout of the cables, computers and other devices that are connected in a computer network. Now, let us see *"What are the various types of network topologies?"* The network topologies are classified as shown below:



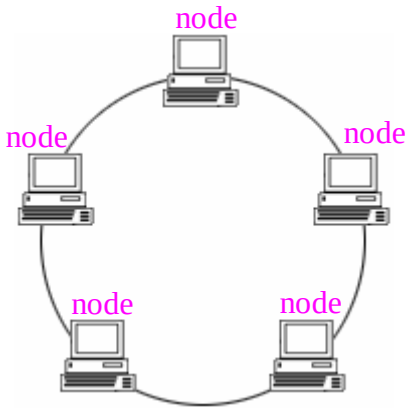
STAR topology: A network setup in which each of the computers and network devices are connected directly to a central computer (server) is called *star topology*. The central computer is also called *hub*. The communication in star topology takes place only through hub. So, the data originated from source machine passes through the central hub before reaching the destination as shown in figure.



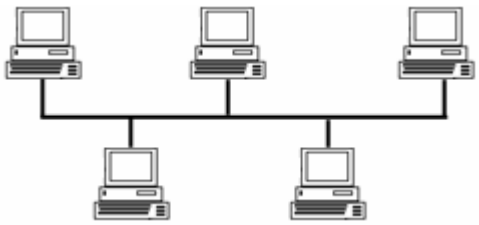
Any device which is connected to the central hub is called a *node*. In this topology, the central hub manages and controls all functions of the network. Observe that even if one node (computer or any device) fails, only that node can be removed and the network will continue to work. But, if hub fails, then whole network fails and communication from one node to another node is not possible.

Ring topology: A network setup in which all the computers and network devices are connected in the form of ring is called *ring topology*. As the name indicates, first node is connected to second, second with third and so on, and finally, last node is connected to the first node, thus completing a ring. A node receives data from one of its two adjacent nodes. Here, there is no central hub as in case of star topology.

At any point of time, only one node sends the data and so no collision of data occurs. When the data from source node is moving in the network, each node checks whether the data is sent to that node. If data is for that node, it utilizes it. Otherwise, it merely passes it to the next node.

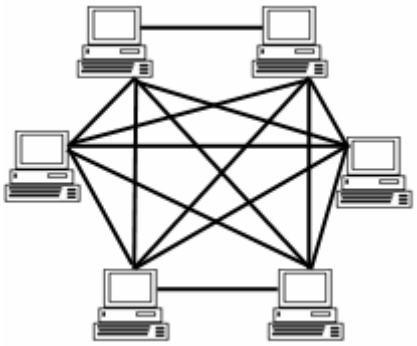


BUS toplogy: A network setup in which each of the computers and network devices are connected to a single cable (also called bus) is called *bus topology.* Here, all the nodes and peripheral devices are connected to a common bus and thus share communication media.



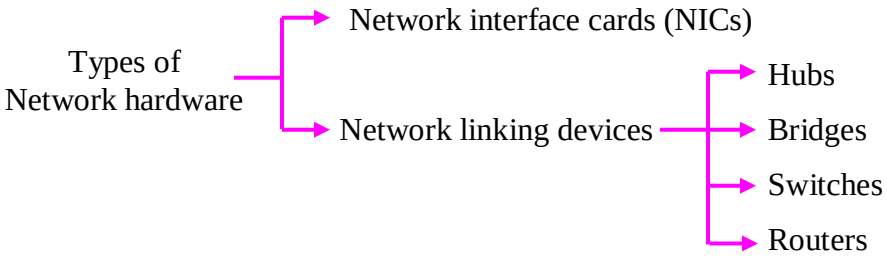
MESH toplogy: A network setup in which each of the computers and network devices are interconnected with one another in the form of a mesh is called *mesh topology.* Here, every node has a connection to every other node in the network.

Since every node has a connection to every other node in the network, this type of topology is not commonly used for most computer networks as it is difficult and expensive. In a mesh topology if any cable or node fails, there are many other ways for two nodes to communicate. So, the network is more reliable when compared to the rest of the network topologies. But, mesh networks are expensive to install because they use a lot of cabling.

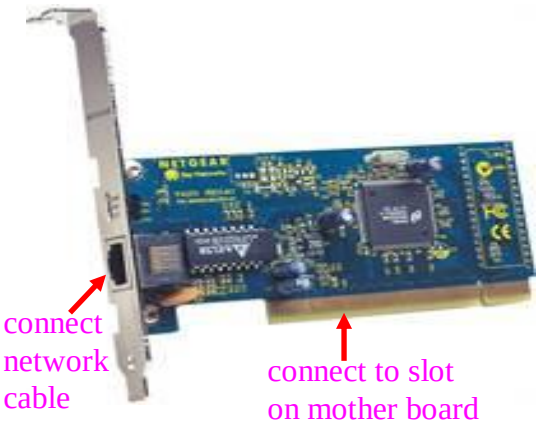


1.9.3 Network Hardware (Network devices)

The data that is transmitted between PCs should be channeled properly so that data reaches destination. The various hardware used in computer network during data transfer are shown below:



Network Interface Cards (NICs): This card is the first contact between the machine and the network. This is a printed circuit board which is fixed into one of the free slots on the motherboard. The network card provides a port to which network cable can be connected. The cable in turn can be connected to various devices in the network so that they can communicate.



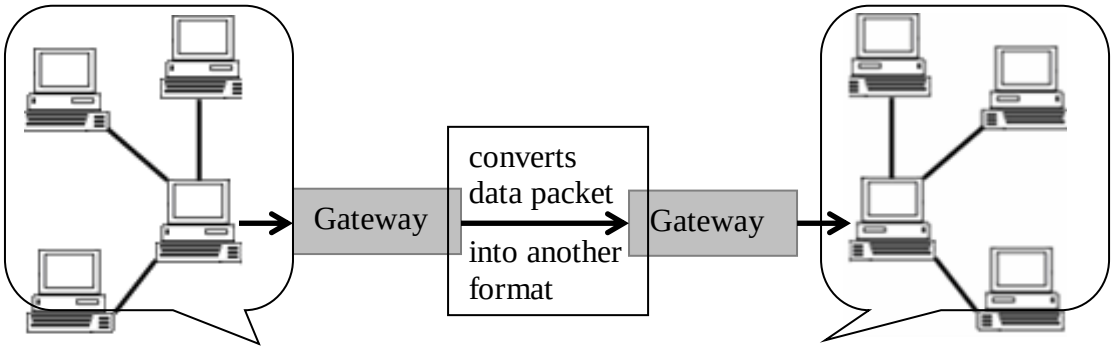
Network linking devices: These are the external devices that are used to link various devices in the network. The various linking devices which are widely used are:

- **Hub:** It is a small box to which various devices are connected so that they can communicate each other.

  The hub performs the following activities:
  - Gathers signals from various network devices
  - Amplify the signals and send them to connecting devices
  - Amplification of signals ensures that the devices on the network receive correct information.
- **Bridges:** A device that connects two or more LANs is a bridge. The bridge regenerates the signal (or amplify the signals) and then forwards the data from one LAN to another LAN. Amplification ensures that devices on the network receive reliable information.

- **Switch:** A switch similar to a hub, where all the devices are connected at one point so that they can communicate. But, a switch is more intelligent. The switch performs the following activities:
  - Inspect the data as and when they arrive
  - Determine the source address and destination address of the data
  - Based on the destination address, forward the data to the appropriate device.

  So, it is the responsibility of the switch to send data to the appropriate device. It has to select one device among many devices connected to it. So, it is called a switch.
- **Routers:** A router is a device that is used to connect two or more networks. The responsibility of the router is to select the best path for fast data transmission. As and when the network traffic changes, the router can redirect the data into less congested routes. Routers are expensive and difficult to maintain.
- **Gateway:** A gateway is a device which is used to connect two or more dissimilar networks that use different communication protocols. So, the protocols of one network have to be converted into protocols of another network. This job is done by gateway. The figure below shows a gateway which accepts data from "type A network" and sends data to "type B network".



Type A network Type B network

So, gateways are also called protocol converters. Mainly gateway does the following activities:

- Accept the data which is formatted for one network
- Convert the data into format of another network

This conversion can be implemented in hardware and/or software.

1.9.4 What is Internet?

Internet is the word which is commonly used. So, let us see *"What is the Internet?"*

Definition: Internet is defined as interconnection of two or more networks. In short, Internet is network of networks. This network interconnects thousands of computer networks worldwide. Internet links various organizations, academic institutions and millions of people to share the information and resources. Any information can be accessed using the Internet.

Using this internet, any computer in one part of the world can communicate with computer in another part of the world. These connections allow the user to:

- Exchange messages
- Communicate in real time (chatting)
- Share data and programs
- Access to unlimited information

1.10 Software basics

Now, let us see "What is software? Explain different types of software?"

Definition: A program is a set of instructions given to the computer to perform some specific task or activity. The set of programs that does the specific jobs are called software. Using the software, computer can receive the data, store the data, manipulate the data and output the data in the correct order and format. Software can be classified as shown below:

- System software — **Ex:** Operating system
- Application software — **Ex:** MS WORD

1.10.1 System software

Now, let us see "What is system software ?"

Definition: The programs that support the operation of a computer are called system software. The system programs help the user to utilize the various resources of computer system effectively and efficiently. Thus, System software is a term referring to any software whose main purpose is to help the user to use the computer system efficiently.

Ex:	**Operating system**	:	It is a set of programs that help the user to interact with various resources in the computer system. Ex: UNIX, LINUX, WINDOWS XP, WINDOWS VISTA etc.

1.10.2 Application software

Now, let us see "What is application software?"

Definition: The set of programs which are developed only for the specific need from the customer are called *application software*. For example, ticket reservation software, payroll software, Software for library management, student information management, business software etc. serve only the specific interests and hence are grouped under application software. Even the various software such as MS-WORD, Microsoft Excel are also grouped under application software as they have been designed and used for specific purpose.

Now, we shall see *the difference between system software and application software.*

System software	Application software
1. Collection of programs that help the user to interact with hardware components efficiently	1. Collection of programs written for a specific application such as payroll, banking, reservation of tickets etc.
2. System software control and manage the hardware	2. Application software uses the services of the system software to interact with hardware components
3. To write system software the programmer needs to understand the architecture and hardware details and hence are machine dependent	3. To write the application software the programmer need not worry about the architecture and hardware details and hence are machine independent
4. Programmer should be more familiar with architecture, instruction formats, addressing modes and so on	4. Programmer should be more familiar with programming languages, data structures and clear knowledge of the problem domain

5. Directly interact with hardware	5. Interact with hardware using system calls provided by various system software. Application software will not interact with hardware directly.
6. Development of system software is complex task	6. Development of application software is relatively easier.
7. Examples: Compiler, Assembler, Operating system, Assembler etc.	7. Examples: Payroll, ticket reservation, Banking software, MS-WORD, EXCEL etc.

1.11 Introduction to computer languages (Programming languages)

We communicate with people using languages like Kannada, English, Telugu etc. On similar lines, we can communicate with computers by giving instructions. These instructions that are given to the computer to do a specific job is called program.

Now, let us see "What is a program? What is a programming language?"

Definition: The set of instructions given to the computer to achieve a specific task is called a **program**. The process of providing instructions to solve a specific problem is called **programming**. The person who writes the program is called ***programmer.*** The language that is used to write a program is called ***programming language***.

Now, let us see "What are the different types of languages?" The types of programming languages are shown below:

- Machine level language
- Assembly level language
- High level language

1.11.1 Machine level language

Machine language is nothing but the set of instructions given to the computer in the form of 0's and 1's. So, machine language consists of only 0's and 1's. This is the language of computers. Various operations can be performed by the combinations of

0's and 1's. For performing addition of two numbers, the machine instruction may be 0100 0111.

Observe that it is very difficult for the human beings to write the program and to understand the program. But, it is easy for the computer to understand machine language.

Advantages

1. Program execution is very fast
2. Since the computer can understand and execute machine language, the translator (such as compiler which we will discuss later) is not required.

Disadvantages

1. Difficult to remember the machine instructions
2. Difficult to read and write machine language programs
3. Very difficult to debug, correct and modify
4. They are machine dependent and are not portable i.e., the program written for one machine cannot be used for another machine with a different processor.
5. It is unstructured language.

To overcome the difficulties in using and understanding machine language, the assembly language was invented.

1.11.2 Assembly level language

Instead of using 0's and 1's to represent an instruction, in assembly level language we use symbolic names. For example, consider the instruction

```
ADD   R1, R2
```

This instruction tells that the values in R1 and R2 are to be added. Thus, instead of writing the instructions in 0's and 1's as in machine language, we can write the program using symbolic names such as ADD, SUB, MUL, DIV. Hence, it is also called ***symbolic language*.** These symbolic names are called ***mnemonics***. This makes the assembly language easier to read and understand when compared with machine language.

Definition: Assembly language also called symbolic language is nothing but the set of instructions given to the computer in the form symbolic names. The symbolic names are also called *mnemonics*.

For example, ADD R1, R2 adds the values of R1 and R2
MUL R1, R2 multiplies the values of R1 and R2

Here, instead of representing the instructions in the form of 0's and 1's, the instructions are represented in the form of symbolic names. In the two instructions, ADD and MUL are symbolic names (also called mnemonics).

Observe that compared to machine language, it is easy for the human beings to write the program and to understand the program.

Advantages of assembly language

1. Easy to remember the symbolic names and there is no need to remember the machine code
2. Program understanding, correction/modification is relatively easier when compared to machine language.
3. The language is simple and easy to write when compared to machine language.
4. The execution speed of a program written in assembly language is same as that of the equivalent program written in machine language.

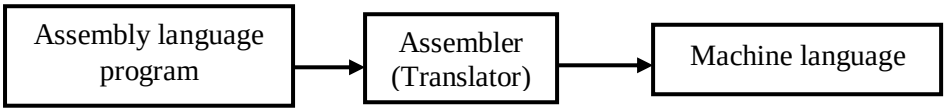
Disadvantages of assembly language

1. There is a need for translating the assembly language program into machine language program.
2. Like machine language, the assembly language is also machine dependent.
3. It is unstructured language.
4. Difficult to understand and debug when compared to high level languages.

Even though it is relatively easier to write assembly language programs, computer cannot understand assembly language. Computer can understand only machine language. As there are translators who can convert Hindi to Kannada, there is a translator which converts assembly language to machine language. This translator is called *assembler*. Now, we briefly discuss about *assembler*.

1.11.3 Assembler

An assembler is translator which translates the program written in an assembly language into the machine language so that the computer can understand. For each assembly language instruction, the corresponding machine instruction is generated by the assembler. Thus, the output of the assembler is the machine language. The translation process is shown in the figure below:



Ex: ADD R1, R2 **Ex:** MASM **Ex:** 0100 0100

Note: We know that writing programs using low level language such as machine language and assembly language is difficult. We all know English language. Why can't we write a program in English like language? Fortunately, we can write programs in English like language. Such a language is called **high level language**. One such high level language is **C language**. In the coming chapters, we concentrate more on C language.

1.11.4 High level languages

A *high level language* is one, which is written using symbols and words just like English language. The high level languages enable the programmer to write machine independent code. The main advantage of high-level languages over low-level languages is that they are easy to read, write, and maintain. The programs written in high-level languages are portable i.e., they can be moved and run on computers of different manufacturers. For example, consider the C program to add two numbers:

```
void main()
{
        int  a = 10,  b = 20, sum;      /* Value of a = 10, b = 20 */
        sum = a + b;                    /* Add value of a, b and store in sum */
        printf("Sum = %d", sum);        /* print the result on the screen*/
}
```

Note that we can easily read and understand this program since all the instructions are similar to English like language.

Advantages of high-level languages

1. Easy to understand
2. Easy to read, write, and modify
3. The code is very compact and self-explanatory
4. They are machine independent programs

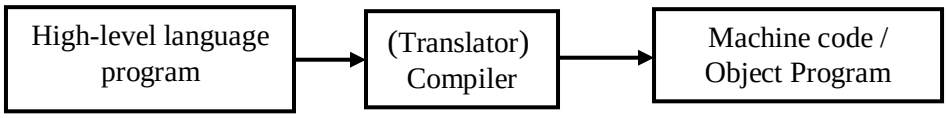
Disadvantages of high-level languages

1. Takes more time to execute when compared with machine language or assembly language.
2. Translator is required to convert the programs written in high-level language to machine language.

Note: The computer cannot understand high level language. So, all programs written in a high-level language must be translated into machine language by the translators. One such translator is called compiler.

1.11.5 Compiler

The compiler is a translator which accepts the program written in high-level language and produces the corresponding machine language program (machine code). The compilation process is shown in figure below.



The compiler after accepting the program written in high level language, checks for the errors introduced by the programmer. If there are any errors, the compiler displays the appropriate error messages along with the line numbers. The programmer can correct these errors and repeat the compilation process. This process is repeated as long as there are no errors in the program. If there are no errors, the compiler produces the machine code.

For example, some of the compilers that are commonly used are: Turbo C, Visual C/C++ etc.

Exercises

1. What are various generations of computers and discuss the corresponding key features of computes in each generation
2. What is data? What is information?
3. What is ASCII?"
4. What are bits?" What is a nibble? What is a byte?" What is a word?
5. explain the basic structure of a computer.
6. Explain the various components of CPU?
7. What is instruction cycle?
8. What is memory? What are the different types of memories?
9. What is the difference between primary memory and secondary memory?
10. What is main/primary memory? What are the different types of primary memories?
11. What are the different types of ROMs?
12. What is volatile memory and non-volatile memory?
13. What is the disadvantage of RAM (normally DRAM)? and How to overcome this disadvantage?
14. What is cache memory? What is a register?
15. What is secondary memory?
16. What are the different types of storage devices?
17. What is a floppy disk or diskette? What is a hard disk?
18. What is a magnetic tape? What are the advantages and disadvantages of magnetic tapes?
19. What are optical storage devices?
20. What are various types of optical storage devices?
21. What is a CD-ROM? How the data can be stored or retrieved?
22. What are the advantages and disadvantages?
23. What is DVD-ROM? What are its advantages?
24. What is flash memory?
25. What is a port? what is a connector? What are various ports that are supported by a computer?
26. What are input devices? What are the various input devices?
27. What is a pointing device? What are the various types of pointing devices?
28. What are optical input devices?
29. What are output devices? What are the various output devices?
30. What are monitors? What are the various types of monitors?
31. What is an impact printer? What are the advantages and disadvantages of impact printers?

1.46 💻 Introduction to Computer hardware/software

32. What is a non-impact printer? What are the advantages and disadvantages of non-impact printers?
33. Compare dot matrix printer and laser printer
34. What is a plotter?
35. What is a network? What are the different types of networks?
36. What is the difference between LAN and WAN?
37. What is a topology? What are the various types of network topologies
38. What is the Internet?
39. What is software? Explain different types of software?
40. What is system software? What is application software?
41. List the differences between system software and application software.
42. What is a program? What is a programming language?