1 INTRODUCTION TO COMPUTER SYSTEM

INTRODUCTION

Computer is an electronic device capable of carrying out sequence of instruction. It helps to accept (read) the input data, process it according to the instructions given and produces (write) processed output information that is results.

It combines five elements:

1. Hardware
2. Software
3. People
4. Procedures
5. Data / Information

Computer has ability of:

- Accept data
- Input, store & execute instructions
- Perform mathematical & logical operations
- Output results according to user requirements

Data: It is a collection of raw facts.

Processing: It is the work done by the computer to process input data and to produce output data.

Processing includes:

- Calculations: it includes equal +, –, %, *,
- Comparison: it includes equal to == <>
- Decision-making: it includes branching to a different path depending on condition.
- Logic: it includes the sequence or flow of step to be followed to get the desired result.

Information: It is the processed data that generates after data processing, which can be used to help people to make decisions. It refers the facts and figure or statistics that have meanings.

Average marks secured by each student from "data" from which class average is calculated that is the information generated.

CHARACTERISTICS OF THE COMPUTER

Speed

It is related to the volume of data processed per unit of time. A computer is very fast device. The speed of computer is measured in:

- Micro seconds10^{-6}
- Nano seconds 10^{-9}
- Pico seconds 10^{-6}

To classify speed of different computer's criteria of MIPS "Million instruction per seconds" is used. Data processors can compare computer-processing speed by the comparing the number of instructions a computer can perform in one second. Fastest computer can execute nearly one Billion instructions per second.

Accuracy

Errors are usually due to human rather then technological or may due to inaccurate data computer does not make mistakes. High accuracy for any computer system is desired. The degree of accuracy is depends on the design model of the computer. Computer is based on the principle of (GIGO) Garbage-In-Garbage-Out if wrong data inputted then a wrong output will be produced.

Reliability

It is exceptionally. It operates under the most advance conditions without showing any signs of fatigue. It is free tiedness, monotony lack of concentration etc. therefore it is used for type of jobs, which required great accuracy. It does not loose concentration.

Memory Capability

It has unlimited capacity to store the data and recall any amount of information because of its secondary storage.

Storage

Large no of the programs and data can be stored on the computer. We don't have to feed programs every time a job is to be done on the computer. Once the program for a job has been stored, it can be recalled and the computer can be asked to execute it.

Diligence

A computer can perform number of functions without suffering from tiredness, lack of concentration etc. if an instruction is given to perform 1 lakh calculation then the machine will perform all the calculation with same speed and accuracy.

Versatility

A computer can control a variety of jobs used in various fields. For performing any job a computer needs a set of program and its related data.

Drawbacks of Computers

- ✔ Lack of decision-making power
- ✔ No I.Q.
- ✔ No heuristics

The Evolution of Computers

1. **The Mark I Computer (1937–44)**

   Also known as automatic sequence controlled calculator, this was first fully automatic calculating machine design by Howard. The mark I was not an electronic computer, rather, it was an electromechanical computer. The size of this computer was huge and its design was very complex, but it was quite reliable. Its design was based on the techniques already developed for punched card machinery.

2. **The Atanasoff Computer (1939–42)**

   The electronic machine was developed by Dr. John Vincent Atanasoff to solve certain mathematical equation. It used 45 vacuum tubes for internal logic and capacitors for storage.

3. **The ENIAC (1943–46)**

   The Electronic Numeric Integrator And Calculator (ENIAC) was the first electronic computer. ENIAC was developed because of military need, and was used for many year to solve ballistic problems. It used about 19000 vacuum tube. The addition of two numbers was achieved in 200 microsecond and multiplication in 2000 microsecond.

4. **The EDVAC (1946–52)**

   The Electronic Discrete Variable Automatic Computer (EDVAC) was the effort to developed a stored instruction computer. The idea of storing both instruction and data in binary form, instead of decimal number system used by human being.

5. **The EDSAC (1947–49)**

   The Electronic Delay Storage Automatic Computer (EDSAC) . In this machine , addition was accomplished in 1500 microseconds and multiplication in 4000 microseconds.

6. **The UNIVAC I (1951)**

   The Universal Automatic Computer (UNIVAC) was the first digital computer. The commercially available digital computer which could be used for business and scientific applications, had arrived.

GENERATION OF COMPUTERS

First Generation (1942–1955)

We have already known about some of the early computer ENIAC, EDVAC etc. These machine and other of their time were made possible by the invention of "vacuum tube" which was a fragile glass device that could and amplify electronic signals. These vacuum-tube computers are referred to first generation computer.

Advantages

- ✔ Vacuum tubes were the only electronic components available during those days. Vacuum tube technology made possible the advent of electronic digital computers.
- ✔ These computers were the fastest calculating devices of their time.
- ✔ They could perform computations in milliseconds.

Disadvantages

- ✔ Too bulky is size and unreliable.
- ✔ Thousand of vacuum tubes that were used emitted large amount of heat and burnt out frequently.
- ✔ Air conditioning required.
- ✔ Non-portable.
- ✔ Commercial production was difficult and costly.
- ✔ Limited commercial use.

Second Generation (1955–1964)

The second generation of computer was marked by the use of transistors in place of vacuum tube. Invention of transistors at Bell Labs by John Bardeen William Shockley and walter Brattain Transistors had a number of advantages over the vacuum tube. As transistors were made from pieces of silicon , so they were more compact than vacuum tubes.

Advantages

- ✔ Smaller in size as compared to the first generation computer.
- ✔ More reliable.
- ✔ Less heat generated and less prone to hardware failures.
- ✔ Better portability.
- ✔ Wider commercial use.
- ✔ They were slightly faster.

Disadvantages

- ✔ Air condition required.
- ✔ Frequent maintenance required.
- ✔ Commercial production was difficult and costly.

Third Generation (1964–1975)

They were still more compacting faster and less expensive than previous generation along with the previous generation. This computer used "Integrated Circuit". These replacing the transistors, which was used in second-generation computer. It uses integrated circuits (IC). It integrates large no of circuit's elements into surface of silicon chips. These computers used IC for CPU computer. In the beginning they use magnetic core memory but later they used semi conductor memory. Semi conductor memory was LSI chip magnetic disk drums & tape was used as secondary memory. Cache memory was also incorporated.

Microprogramming, parallel programming, multi programming, multi user system (time sharing system) etc. were used. Multiprocessing too. The concept of virtual memory was also introduced. Example IBM 370 series, CDC 7600.

Advantages

- ✔ Smaller in size as compare to previous generation.
- ✔ More reliable.
- ✔ Lower heat generated
- ✔ Computational time was reduced from micro second to nano second

- ✔ Maintenance cost was low because hardware failure was rare
- ✔ Portable
- ✔ Used for commercial application
- ✔ Less power required than previous of generation of computer
- ✔ Commercial production was easier and inexpensive
- ✔ They were totally generally purpose machine.

Disadvantages

- ✔ Air condition was required in many cases
- ✔ Highly sophisticated technologies was required for the manufacturing of IC chip

Fourth Generation (1975 onwards)

It integrated up to 100 components on a single chip that is Medium Scale Integration (MSI)

LSI (Large Scale Integration)

It contain over 30,000 components on a single chip

VLSI (Very Large Scale Integration)

That is over 1 million components on a single chip.

Fourth generation use. VLSI chips for both CPU and memory. CPU consists of one or more microprocessors. The latest microprocessors can contain one million transistors. Cache memory is provided on CPU chip.

Advantages

- ✔ Smaller in size because of high density of component in a small chip.
- ✔ Very reliable
- ✔ Heat generation very low
- ✔ Less powerful air conditioning
- ✔ Faster computation
- ✔ Minimum maintenance is required
- ✔ In expensive among all generation
- ✔ They were totally generally purpose machines

Disadvantages

- ✔ Highly sophisticated technology required for the manufacturing of LSI chip

Fifth Generation (yet to come)

These computers required genuine IQ that is ability to reason logically and contain real knowledge. The structure will be parallel. They will be able to perform multiple tasks simultaneously. Data and knowledge are processed they shall follow 'kips' than dips/lips.

KIPS	(Knowledge Information Processing System)
DIPS	(Data Information Processing System)
LIPS	(Logical Information Processing System)

Japan has started research using prolog languages as operating system.

Classification of Computer

Microcomputer

Two major events led to the introduction of the first microcomputer: the development of the first microprocessor in 1969, and the creation of the first general purpose microprocessor chip in 1971. Microcomputers had a major economic impact due to their small size and drastically improved cost-effectiveness. These powerful, yet easy to operate, machines have been called home computers, laptop computers, notepad computers, personal computers (PC), palmtop computers, and micros. The microcomputer can perform jobs once only handled by the largest computers. A typical microcomputer memory unit stores between 32 and 1,000 million characters of data. Since these computers are slower than their mainframe and minicomputer counterparts, processing speeds are measured in Megahertz (MHz) or Gigahertz (GHz) instead of MIPS. Generally, the larger the processing speed, the faster the computer is. A computer running at 2.53 GHz (2,530 MHz) is faster than one running at 450 MHz.

Minicomputer

These are powerful than microcomputer and supports several uses. They use large RAM and backing storage capacity and are quick in data processing *e.g.*, PDP ii VAX 7500. The minicomputer was introduced in 1963. It is a scaled-down version of the mainframe computer. It is capable of performing many jobs that old mainframe computers used to accomplish. Due to its cost-effectiveness, the minicomputer's most important contribution has been the introduction of distributed computing, where a number of small computers can be used instead of one giant central computer.

Mainframe Computer

These computers are very powerful large general-purpose compute. They are faster and more powerful than minicomputer. The word length 48,60 or 64 bits. They are used for complex calculation where large amount of data is to be processed. They are useful in research organization banks, airline, reservations where large capacity of these computer have increased *e.g.*, IBM4381, ICL39 series.

The first general-purpose computer to be developed, in the 1940s, was the mainframe. Mainframe computers can be used by many people at the same time, and can handle very complex problems or large volume jobs. Mainframes are generally used for data processing for business and scientific applications. For example, major banks can process bills for credit card holders all over the world with the aid of a mainframe. Large research projects also use the mainframe's vast memory, storage capacity, and super-fast processing speed to conduct tests and co-ordinate operations.

Mainframes can have memory capacities measured in billions of characters and more. The largest mainframes can process over 3,000 Million Instructions Per Second (MIPS).

Super Computer

They are much faster and more powerful than main computer. There processing speed lines in a range of 400 MIPS to 10000 MIPS. Word length is 64 to 96 bytes. These computers are design to maximize the number of FLOPS and it is usually more than one GFLOPS. They contain number of CPU, which operate in parallel to make it faster. They are used for whether for casting, rocketing in aero dynamics, atomic nuclear and plasma physics *e.g.*, CRAY-XMA. India has designed super computer indigenously called PARAM 10000.

Broad Classification

Analog Computer

Analog computers measure, analyses, and control events that are continuous, such as the flow of electrical current along power lines.

Uses : In scientific and engineering fields use then most because quantity is very constantly.

Digital Computer

Digital computers, on the other hand, control events that are discrete or finite *e.g.*, process in automobile, control fuel, braking system.

Hybrid Computer

This machine combines feature of both analog and digital computer *e.g.*, ECG machine. Analog machine measure patient temperatures then they are converted to numbers and supply digital computer.

□□□

2 COMPUTER ORGANIZATION

ARCHITECTURE OF COMPUTER SYSTEM

The internal architecture design of computers differs from one system model to another. However, basic organization remains same for all computer systems. A block diagram of the basic computer organization is shown in figure.

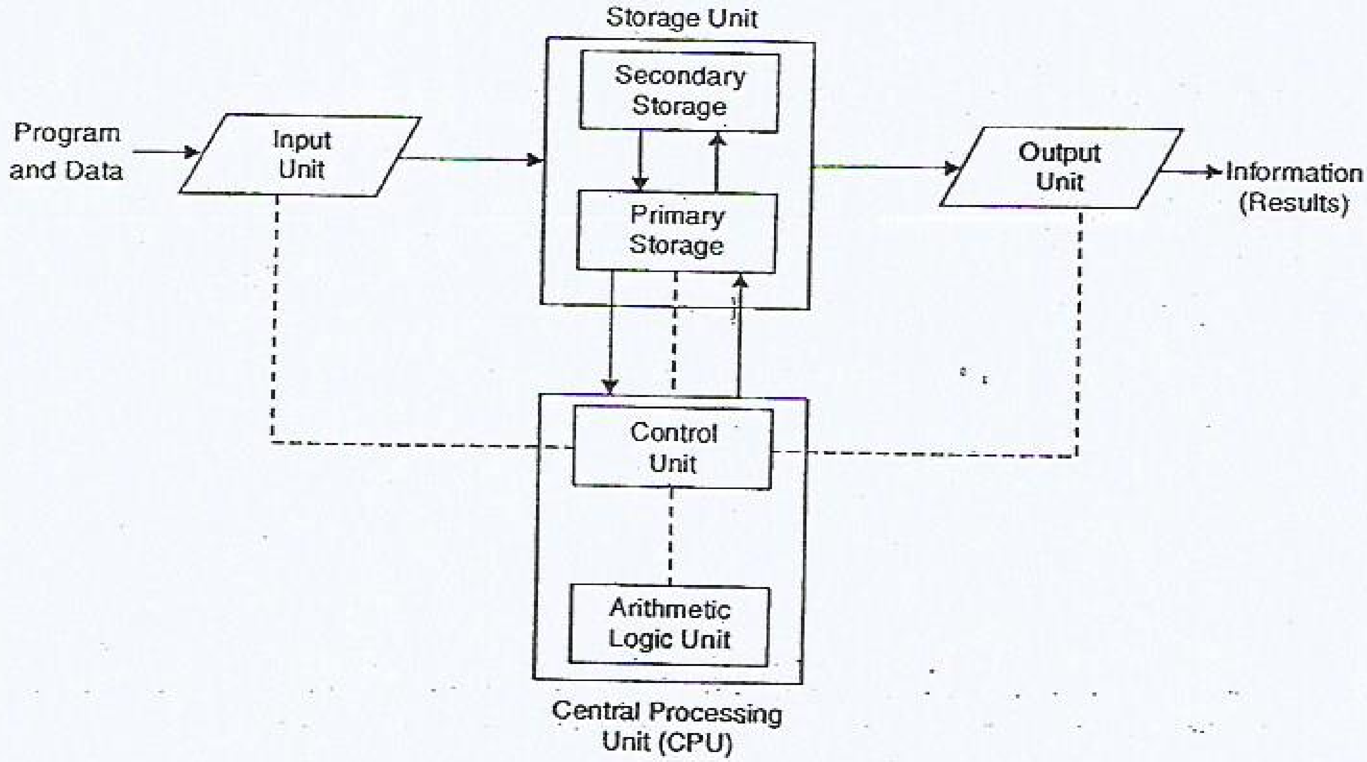


Figure 2.1. Basic organization of a computer system.

There are three major unit of an electronic digital computer. The units are Input unit, Output unit, and Central Processing Unit (CPU).

Input Unit

The function of input unit is to read necessary data into computer system or we can say that unit links the external environment with the computer system. Some of the more common input devices are keyboards, punched card and punched paper tape reader, magnetic tape reader etc. Data and instructions are taken in various forms depending on input devices used. For example, data through keyboard entered similar to typing, and this differs from the way in which data is entered through a punched card. However regardless of the forms in which they accept input, all input devices transformed that into binary codes that the primary memory of the computer is designed to accept. This transformation is accomplished by units called input interfaces.

The following functions are performed by an input unit:

- ✔ It must provide a computer with data and instructions that are necessary to perform the task.
- ✔ It converts that input data in computer acceptable form.
- ✔ It supplies the converted instructions and data to the computer system for further processing.

Output Unit

The function of an output unit is just reverse of that of input unit. Output unit supplies the computed data to the outside world. Thus it links computer with the external environment. As computer work with the binary code, the results produced are also in the binary form. Hence before supplying the results to the outside word, it must be converted to human readable form. This task is accomplished by units are called output interfaces. Common output devices are CRT displays, card-punching machines, and magnetic tape drives etc.

The following functions are performed by an output unit:

- ✔ It accepts the results produced by the computer.
- ✔ It converts these coded results to human readable form.
- ✔ It supplies the converted results to the outside world.

Central Processing Unit (CPU)

The CPU is the brain of a computer system. In a computer system all major calculations and comparisons are made inside the CPU and the CPU is also responsible for activating and controlling the operations of other units of a computer system. The two basic components of a CPU are the control unit and the arithmetic logic unit.

Arithmetic Logic Unit (ALU)

The Arithmetic Logic Unit (ALU) of a computer system is the place where the actual execution of the instructions takes place during the processing operation. In general, all calculations (addition, subtraction, multiplication and division etc.) are performed and all comparisons (decisions) are made in the ALU. The data stored in the primary storage are transferred to ALU where processing takes place and intermediate and final results are transferred to storage. The control unit tells the ALU which operation to perform and then see that necessary data to be supplied.

Control Unit

The control unit sequences the operation of the computer, controlling the actions of all other units. That interprets the instructions and then directs the rest of the machines in its operation. Although, it does not perform any actual processing on the data, the control unit acts as a central nervous system coordinates the entire computer system.

Storage Unit

The input data, final and intermediate results produced by the computer must be kept somewhere inside the computer system. The storage unit of a computer is designed to fulfill to all these needs.

The functions of storage unit are to hold:

- ✔ All the data to be processed and the instructions required for processing.
- ✔ Intermediate results of processing.
- ✔ Final results of processing before these results are released to an output device.

The storage unit of all computers is comprised of the following two types of storage:

Primary Storage

The primary storage, also known as main memory, is used to hold pieces of program instructions and data, intermediate results of processing, and recently produced results of processing. However, the primary storage can hold information only while the computer system is on. As soon as the computer system is switched off or reset, the information held in the primary storage disappears. The primary storage of modern computer systems is made up of semiconductor devices.

Secondary storage

The secondary storage is used to take care of limitations of the primary storage; it is used to supplement the limited storage capacity and the volatile characteristic of primary storage. Secondary storage is much cheaper than primary storage and it can retain information even when computer system is switched off or reset. The most commonly used storage medium is the magnetic disk.

□□□

The Personal Computer (PC) an overview: -

The development of the PC followed closely upon the development of integrated circuits. The introduction of very large and ultra large scale integration (VLSI and ULSI) which contain the equivalent of a million more transistors pared the way for the manufacture of complete circuits of complex circuits and complete system on a single silicon chip.

The very first PC was the XT (extended Technology). IBM launched it at 12 August 1981. This was followed a few years later by the AT (Advanced Technology) type using the 80286, 80386, 80486 and the current Pentium processors.

Although a computer system may have several external devices connected to it, such as a printer, mouse, and a keyboard; the essential elements of PC's are the same, namely a system unit, a keyboard and a visual display unit (VDU) or monitor. The keyboard provides the user with access to the system unit while the monitor provides a visual display of the textual output of the keyboard as well as information and messages from the system unit.

The system unit contains all the devices and circuitry necessary for the operation of the computer. The component parts of a system unit are:

- The power supply.
- Disk drives.
- Expansion or adapter cards.
- The motherboard.
- Connectors or ports, which provide access to external device (printer, mouse, VDU, modem, etc).

The Power Supply: -

The power supply converts a.c. mains into the necessary stabilised d.c. Voltages for the motherboard, the keyboard and other elements within the system unit including the expansion cards. The power supply provides +5V, -5V, +12V and –12V together.

Disk drives: -

Disk drives are mass data storage devices. A typical computer employs two types of disk drives: a floppy and a hard disk drive. Data is stored on concentric tracks that have been formatted onto plastic (floppy) or metal (hard) disk, which has been load with a very thin layer of electromagnetic material.

Expansion (adapter) cards: -

Expansion or adopter cards are printed circuit boards, which support essential elements such as disk drives and video displays or add-on facilities such as CD-ROM, and Modems by simply inserting them into expansion slots on the motherboard.

The motherboard: -

The motherboard also knows as system board, is the core (main part) of the computer system. It contains all the necessary components (including the processor and other support and logic chips) and circuitry for the operation of the system. A motherboard known by the type of processor it supports, hence a 386 (using an 80386 processor), a 486 or a Pentium motherboard. Modern motherboards are designed to support more than one type of processor, making it possible to upgrade the system without changing the motherboard.

Central Processing Unit (CPU): -

The brain of any computer system is the CPU. It controls the functioning of the other units and processes data. The CPU is sometimes called processor, or in the PC field called "microprocessor". Microprocessor: - is a single integrated circuit (I.C) that contains all the electronics necessary to follow instructions stored in internal memory. In other words, it contains all the electronics needed to execute a program. The microprocessor calculates (adds, multiplies. and so on). Performs logical operations (compares number, makes decisions). and controls the transfer of data (moves information among devices), and timing. The capacity or size of a microprocessor chip is determined by the number of data bits it can handle.

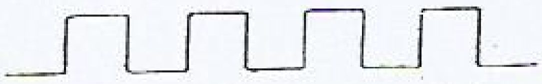
Microprocessor also differ in the speed with which they execute instructions CPU speed is indicated by the frequency of the system clock in megahertz (MHz, million of cycles per second). While the bit width or size determines the quantity of information that may be transferred in any one cycle. The speed determines the number of such transactions per second.

Table bellow lists the CPU's in common use by manufactures, the year were launched, their size and speed.

CPU	Launched	Bit size	Speed (MHz)
8088	1979	8	5
8086	1978	16	8
80286	1982	16	8-12
80386	1985	32	16-25
80486	1989	32	25-100
Pentium	1993	32	60-200

CPU control signals: —the number and type of control signals depends on the microprocessor used and the design of the system. Control signals are normally active low, i.e. active when at logic 0. Active low signals are signified by a bar (-). The main control signals of a CPU are as follows.

The clock pulse signal: - A clock pulse is essential requirement for the operation of the processor.



The clock pulse wave form

The clock controls signal syschronises the movement of the data around the various elements of the system and determines the speed of operation. without which the system comes to a halt.

Read (RD) and Write (WR): - the CPU determines the direction of data transfer to or from the microprocessor chip. This is carried out by the read and write control lines. In a READ operation when the CPU is receiving data from memory, the READ line is active allowing data to be transferred to the CPU. In a WRITE operation when the CPU is sending data to memory, the WRITE line is active enabling data transfer from the CPU to memory.

Interrupts: -when peripheral devices such as a printer, a keyboard or a modem needs attention, hardware interrupt signal, INTR (interrupt request), is send to the CPU. When such a signal is received, the main program is interrupted temporarily to allow the CPU to deal with the request. After servicing the peripheral device, the CPU returns to the original program at the point where it left it.

In general CPU consists of the following parts:

1- Arithmetic and Logic unit (ALU).

Execution of the most operations within a computer (such as arithmetic or logic operations) takes place in the ALU.

For example, suppose there are two numbers located in the main memory are to be added.

Solution: they are brought into the ALU, where the actual additions is carried out, the sum may then be stored in the memory and from there to an output devices.

Similarly any other arithmetic or logic operations can be performed in a similar way.

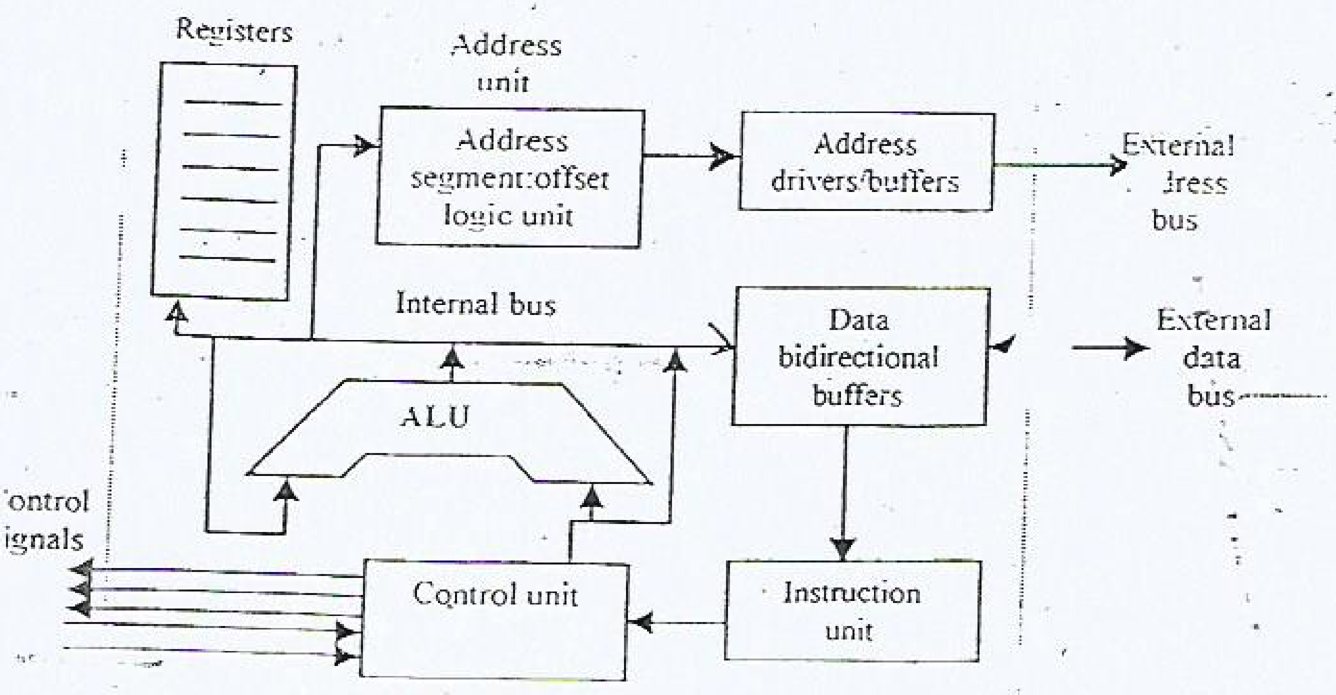
In general not all operands in a going computation reside in the main memory, since processors normally contain a number of high-speed storage elements. called "registers", which may be used for temporary storage of often used operands.

2- **Internal bus structure:** - this allows the various parts to communicate with each other. Driver/buffer interface circuits accomplish communication between the internal bus and the external word. one for the address lines and another for the data lines.

3- **A number of registers:** - the processor provides a number of registers to be used as temporary stores of digital information.

4- **Instruction unit:** - the instruction unit receives and stores each instruction, decodes it and informs the control unit of all the necessary steps to execute the instruction.

5- **Address logic unit:** - the address of a memory location is made available to the processor in the segment: offset format. Before an address is sent to the address bus, it must be translated into an absolute (physical) address. The address logic unit carries out this task.

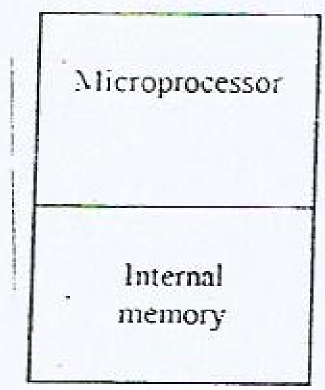


3- Control unit: the operation of the previous described units (I/O unit & ALU) must be coordinated in some organized way, which is the task of a control unit. Although the control unit doesn't performed any actual processing on the data, it acts as a central nervous system used to sent control signal to other unit.

The control unit provides the timing and control signals necessary to synchronise the internal operation of the CPU as well as the computer system as a whole.

Notes:

- The microprocessor and internal memory work together closely. The microprocessor fetches instructions one at a time from memory and completes each instruction before beginning the next. Each instruction is a simple. quite understandable operation such as adding the number in one memory address to the number in another address, moving a character from the keyboard to memory, or deciding where to find the next instruction. Because each instruction is extremely simple. completing a useful task takes hundreds, thousands, or even millions of primitive operations. Obviously, a microprocessor must work quickly to be useful.



CPU

-Both the program (instructions) and the data (letters, decimal numbers, and so on) are stored in memory as numbers. For example, the letter A is normally stored as the number 65.

- A microprocessor is capable of performing a limited number of primary operations, called an **instruction set**. Program for one microprocessor will not run on another unless they are rewritten or translated from one instruction set to the other. Depending on how a

program ... ...
much work as written it from scratch.

- Four factors have a major influence on a microprocessor's power:
  1- How many bits of data are processed in one operation? This depends on :
     1- The number of bits processed internally in each operation.
     2- The number of bits transferred between the microprocessor and internal memory at once.

  2- How many operations are in the instruction set and how useful are they?

  3- How long does it take to complete an instruction? This depends on :
     1- The clock rate, which paces all the operations in the CPU. Clock rates are measured in megahertz (MHz) or millions of cycles per second. For example: a Pentium 2 has 433 MHz clock rate, which means there are 433,000,000 ticks (called clocks) of the CPU's clock each second.
     2- The number of clocks per instruction. For example, it takes IBM PC 4 clocks to transfer one byte from internal memory to the microprocessor, and it can take more than 100 clocks to multiply two 16-bit numbers together.

  4- how much internal memory can the processor manage? Each word of internal memory needs its own unique address, so the length of each memory address limits the maximum amount of internal memory. For example, 16-bit memory address, which limit the internal memory to 64 KB, and 20-bit memory address, which limit the internal memory to 1megabyte (1MB).

A microprocessor's raw processing power is not important for most applications. It is much better to have a system in which the software makes full use of a processor with rather limited capabilities than to have poor software running on a potentially powerful processor.

Memory unit: -

A computer cannot work without some form of memory to store programs and routines. computer memory is used to store the sequence of instructions and related data, which form what is known as a program or routines. It is also used to store temporary or permanent data used or created by the program. In running a program, the CPU will need to access the memory store at a rate of between 10 and 50 MHz.

In theory any type of memory including a disk drive can be use to provide a large enough storage to store the program instructions and associated data. However, because disk drives are extremely slow compared with the speed of the CPU, using them as the main processor memory would make the whole system very slow. The processor will spend more of its time waiting to access the disk drive than carrying out program instructions. For this reason the main memory where application and other support programs are loaded must have a speed comparable with that of the CPU itself. This means a memory store in the form of integrated circuits or chips. A small amount of memory is also provided by the processor itself.

The internal memory of a PC is composed of integrated circuits. In today's technology these circuits are printed on small chips cut from thin slices (or wafers) of large silicon crystals. Hence the circuits are called memory chips. From the outside, a memory chip looks just about like any other integrated circuit chip.

On the surface of each memory chip are a large number of electronic cells. Each cell can be either full or empty. We can therefore represent its possible states with just the two digits in the *binary number system*, 0 and 1, and data coded into these binary digits can be represented by the state of the electronic cells. Each of the binary digits is called a bit, which is short for binary digit. A grouping of eight bits is a byte. A byte is being able to store one character.

It is easiest to build and use memory chips that have a round number of storage cells in binary arithmetic (which in decimal numbers means 2, 4, 8, etc.). A more convenient unit of measure for memory uses the symbol K. in the metric system K is the symbol for 1,000; but when memory is being measured, K is the symbol for 2^{10}, or 1024. Thus a memory chip with 2^{14} (16384) binary storage cells is called a 16K-bit chip. Similarly, the terms kilobit and kilobyte (KB) mean 1024 bits or bytes, of memory. To designate the memory capacities of mainframe computers and hard disk drives, we need even larger units. A megabyte

3

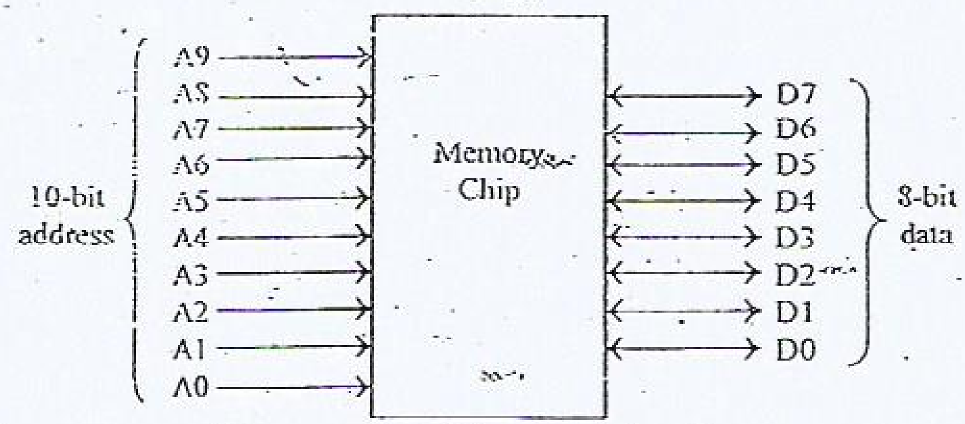
(MB) is 1024 KB, 2^{20} bytes, or roughly a million characters of storage. A gigabyte is 1024 megabytes, roughly a billion characters.

Memory contains a number of semiconductor storage cells, each capable of storing a one-bit of information. Memory has two main properties that determine their application: storage capacity or size and access time or speed.

A memory chip contains a number of locations, each of that stores one or more bits of data known as its bit width. The storage capacity of a memory chip is the product of the number of locations and the bit width. For example, a chip with 512 locations and a 2-bit data width, has a memory size of 512 * 2=1024 bits.

The standard unit of data is a byte (8 bits), the above storage capacity is normally given as 1024/8 =128 bytes.

The number of locations may be obtained from the address width of the chip. For example, a chip with 10 address lines has 2^{10} =1024 or 1K locations. Given an 8-bit data width, a 10-bit address chip has a memory size of $2^{10} * 8 = 1024 * 8 = 1K * 1byte = 1Kbyte$ or 1KB.



Memory chip address and data lines.

A single chip is usually insufficient to provide the memory requirements of a computer. A number of chips are therefore connected in parallel to form what is known as a memory bank. Figure below shows a memory bank consisting of eight 1-bit chips. Each chip has 18 address lines (A0------A17). The total storage capacity of the bank may be calculated as follows:

Capacity of one chip $2^{18} * 1 = 256 K * 1 = 256$ K bits

Total size of the memory bank = 256 K bits * 8 = 256 K byte

The second properties of memory chips is access time, access time is the speed with which allocation within the memory chip may be made available to the data bus. It is defined as the time interval between the instant that an address is sent to the memory chip and the instant that the data stored into the location appears on the data bus. Access time is given in nanoseconds (ns) and varies form 25ns to the relatively slow 200ns.

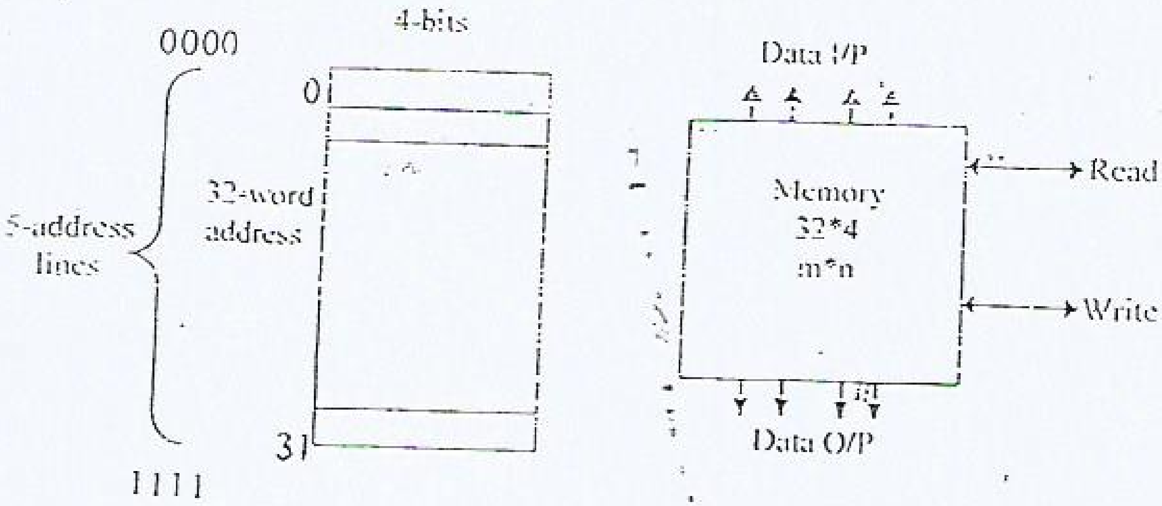
Notes: -

- The large computers (main frames) have word-sizes that are usually in the 32-to-64-bit range.
- Minicomputers have word sizes from 8-to-32-bit range.
- Microcomputers have word sizes from 4-to-32-bit range.

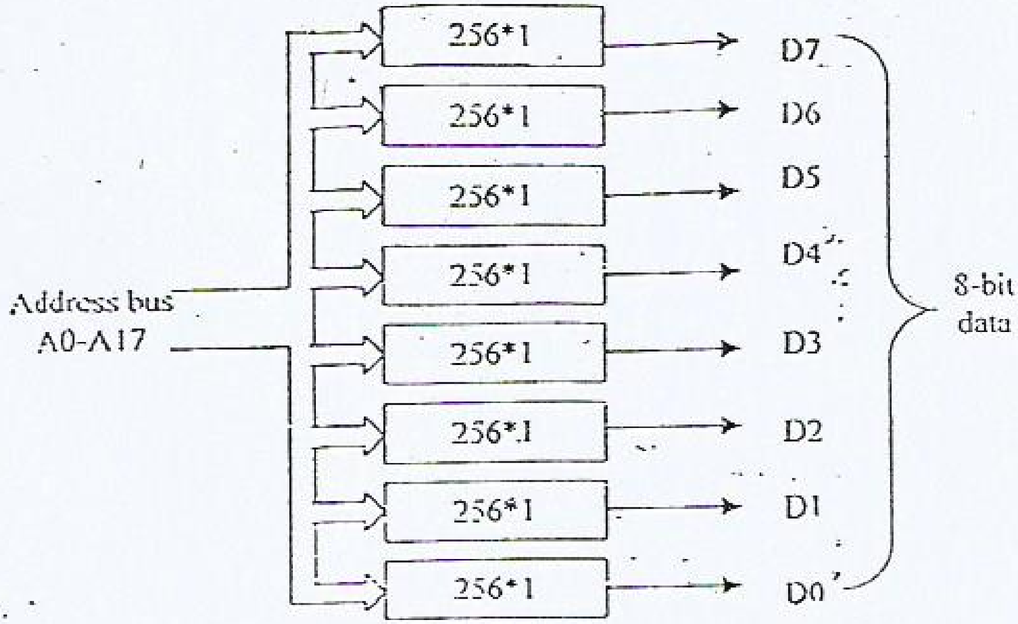
In general a computer with a larger word size. can execute programs of instructions at a faster rate because more data and more instruction are stuffed into one word. The larger word sizes however, mean more lines making up the data bus, and therefor more interconnections between the CPU and memory and I/O devices.

A more conventional unit of measure for memory is that to use kilobyte (KB), means 1024 bytes (2^{10}). A large unit called a Mega Byte (MB), means $2^{10}\ KB = 2^{20}\ Byte\ = 1.048\ .576\ Byte$

A more important than a computer's word size is the amount of memory the computer has (i.e.) the memory capacity. A memory capacity is a way-of specifying how many bits can be stored in a particular memory device or complete memory system.



This memory stores 32. 4-bit words.



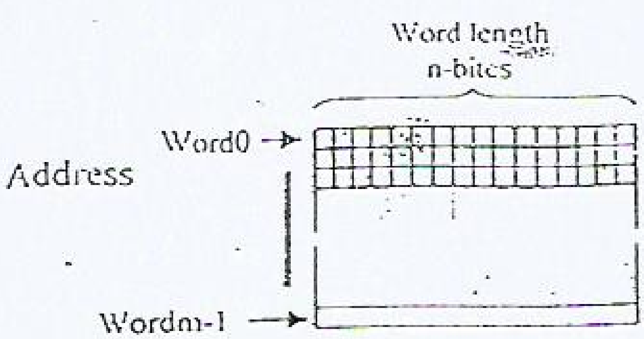
Memory bank

The computer's word size can be expressed in bytes as well as in bits. For example, a word size of 8-bit is also a word size of one byte, a word size of 16-bit is a word size of two byte. Computers are often described in terms of their word, such as an 8-bit computer, a 16-bit computer and so on. For example, a 16-bit computer is one in which the instruction data are stored in memory as 16-bit units, and processed by the CPU in 16-bit units.

The word size also indicates the size of the data bus, which carries data between the CPU and memory and between the CPU and I/O devices.

To access the memory, to store or retrieve a single word of information, it is necessary to have a unique address. The word address is the number that identifies the location of a word in a memory.

Each word stored in a memory device has a unique address. Addresses are always expressed as binary number, although hexadecimal and decimal numbers are often used for convenience.



Main memory organization

-The word size is 4-bit, therefore these are 4 data I/P lines and 4 data O/P lines

-This memory has 32 different words and therefore has 32 different addresses (storage location) from 00000 to 11111. Thus we need a 5 address I/P lines.

Memory capacity = number of memory storage location * size of each word
=(number of word) * (number of bits per word)
=m (word) * n (bits)
=m * n bits

- The capacity of memory depends on two parameters the number of words (m) and the number of bits per word (n).
- Every bit added to the length of address will double the number of words in the memory.

The increase in the number of bits per word requires that an increase., the length of data I/P and data O/P lines.

EX: - a certain memory chip is specified as 2K * 8

1- How many words can be stored on this chip?
2- What is the word size?
3- How many total bits can this chip store?

Sol: -

1- 2K = 2 * 1024 = 2048 words
2- The word size is 8-bits (1 byte)
3- Capacity = 2048 * 8 = 16,383 bits = 16 KB.

EX: - a certain memory chip is specified as 2K * 16

1-How many words can be stored on this chip?
2-What is the word size?
3-How many total bits can this chip store?

Sol: -

1-2K = 2 * 1024 = 2048 words
2-The word size is 16-bits (2 byte)
3-Capacity = 2048 * 16 = 32,768 bits = 32 KB.

EX: -

Which memory stores the most number of bits
2M * 8 memory or 2M * 16 memory?

Sol: - 1- capacity = (2 * 1024 * 1024) * 8 = 16,777,216 bits
2- capacity = (2 * 1024 * 1024) * 16 = 33,554,432 bits

EX: -

Which memory stores the most number of bits
4M * 8 memory or 2M * 16 memory?

Sol: - 1- capacity = (4 * 1024 * 1024) * 8 = 33,554,432 bits
2- capacity = (2 * 1024 * 1024) * 16 = 33,554,432 bits

EX: - A certain memory has a capacity of 4K * 8
1- How many data I/P & data O/P lines?
2- How many words address line?
3- What is its capacity in byte?

Sol: - 1- 8 each line so data I/P lines = data O/P lines = 8
2- 4 * 1024 = 4096 words thus there are 4096 memory addresses
Therefore that is required a 12-bit address line
3- the capacity = (4 * 1024) * 8 = 32,768 bit
=32,768/8 = 4096 byte.

EX: - A certain memory has a capacity of 4K * 16
1-How many data I/P & data O/P lines?
2-How many words address line?
3-What is its capacity in byte?

Sol: - 1- 16 each line so data I/P lines = data O/P lines = 16
2- 4 * 1024 = 4096 words thus there are 4096 memory addresses
Therefore that is required a 12-bit address line
3- the capacity = (4 * 1024) * 16 = 65,536 bit
=65.536/8 = 8192 byte.

Note: -

A computer with memory less than 16 KB is limited to trivial application. because it can execute only small program. A personal computer must have at least 640-KB memory.

- Instruction and data can be written into or read out under control of the CPU. It is essential to be able to access any word location within the main memory as quickly as possible.
- Memories, where any location can be reached by specifying its address, are called random-access memories.
- The time required to access one word (from memory) is called the memory access time. That is (i.e.) access time is measure of memory devices operating speed.

The two major types of memory are: -

1- Read-Write or Random Access Memory (RAM): it has variable content, also it is generally used to store the variable data. RAM can also be used to store frequently changed programs and other information.

RAM allows the computer to store information quickly for later reference, so that (in most personal computer). RAM hold: -

- The active part of the operating system, the fundamental programs that control the operation of the computer.
- The application program being executed (for example word processing program).
- Data used by the application program (for example a letter being written with the word processing program).
- A representation of the data being presented on the video display.
- Any thing else that is likely to change frequently (for example the time of the day in the computers a clock).

In RAM, the stored information will be lost when computer power supply is removed (even a short interruption). That is, RAM, is volatile memory.

When program instruction, reads the data in memory address it gets a copy of the data. This is called a nondestructive read, because the contents of the memory address are not changed. Sending data to a RAM memory address is called a destructive write because the new data erases whatever was there before.

RAM can be divided as: -

Static RAM: - semiconductor memory devices in which the stored data will remain permanently stored as long as power is supplied, Dynamic RAM: - semiconductor memory devices in which the stored data will not remain permanently stored even with power applied, unless the data are periodically rewritten into the memory.

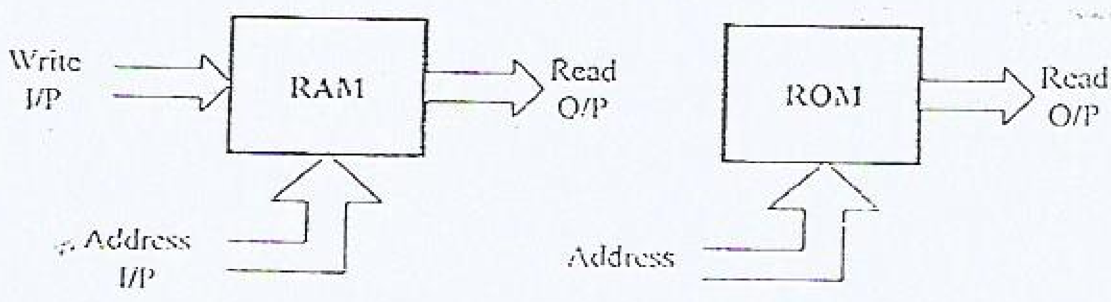
2- Read Only Memory (ROM): - is a nonvolatile memory. when the computer off, the contents of ROM are not change. That is the main different between RAM & ROM. Sending a data to a ROM memory address is ineffective, because the contents of ROM are permanent and can not be changed.

ROM can be divided as: -
1-PROM (Programmable ROM): -ROM that can be electrically programmed by the user. It can not be erased & programmed. PROM can be programmed once with special circuitry.

2- EPROM (Erasable Programmable ROM): - ROM that can be electrically programmed by the user. It can be erased (usually with ultraviolet light) and reprogrammed as often as required.

3- MPROM (Mask-Programmed ROM): - ROM that can only be programmed at the factory.

4- EEPROM (Electrically Erasable Programmable ROM): - ROM that can be erased with ultraviolet light and reprogrammed as often as required.



NOTE: -

The two major type of internal memory is RAM and ROM.
-RAM (Random Access Memory or Read/Write Memory): -is a memory that information can be both read from and written into it. When the power turn off from the computer, the information will be lost.

-ROM (Read Only Memory): -is a memory manufactured to store permanently a fixed set of information. For example, in all personal computer ROM stores the instructions that tell the microprocessor what to do when the power is turn on.

7 INPUT OUTPUT DEVICES

INTRODUCTION

The input-output devices provide the means of communication between the computer and the outer world. They are also known as peripheral devices because they surround the CPU. Input devices are used to enter data into primary storage and output devices accept results from the primary storage.

Input Devices

Input devices are generally of two types—Online and Offline.

1. Offline: An operation that does not directly involve a computer.
2. Online: An operation that directly involves a computer.

Offline Input Devices

Input devices that allow data entry operations without the direct involvement of a computer are called off-line input devices. Generally the input entered through these device is recorded on some media first and then processed by a computer later. For example key-to- punch, floppy etc. Most of lines are now obsolete and are very rarely used.

Online Input Devices

Online input devices provide direct and interactive communication between the user and the computer. The data from that device is sent directly to the computer with no need for intermediate media. These devices are economical when the volume of data is low and irregular. For example Keyboard, Mouse, Punched cards, Light Pen, joystick etc.

Punched Cards

There are two types of punched cards, one has 80-columns and the other has 96-columns.

The 80-columns card: It is 19.3 cm in length, 9.5 cm in width, and 0.018 cm in thickness. The card is divided from left to right into 80 vertical columns numbered 1 to 80. It is again divided into 12 rows numbered 12,11,0,1,2,3,4,5,6,7,8 and 9 from top to bottom. So a single card can represent a maximum of 80 characters. The digits 0 to 9 are represented by punching just one hole in the corresponding row position. The alphabets A to Z are represented by a combination of two holes in two of the row positions. The top three rows—2,11 and 0 is zone-punching positions and the rows 0 to 9 are numeric punching positions. A logical combination of zone and numeric punches is required to represent alphabets. The coding system used in this is known as Hollerith code.

The 96-columns card: It is only one third the size of an 80-columns card. The 96-columns are separated into three 32-columns sections or ties. The upper portion of the card, which is not used for punching holes, is used as a print area. These cards have round holes instead of rectangular holes of 80-columns cards. The standard 6-bit BCD code is used instead of Hollerith code for recording data on a 96-columns card. Each of the 96-columns has six punch positions. The upper two are zone positions and the remaining four are numeric positions. The presence of a hole in a punch position indicates a 1-bit.

Keyboard

The keyboard is a universally used online input device. This input method is similar to typing on a typewriter.

The following keys are present on most keyboards.

Alphabets A–Z

Numeric digits 0–9

Delimiters ',',';','.',':'etc.

Operators '+','*','–','/' etc.

Special keys CTRL, ALT, ENTER, SHIFT etc.

Function keys F1, F2, and F3 etc.

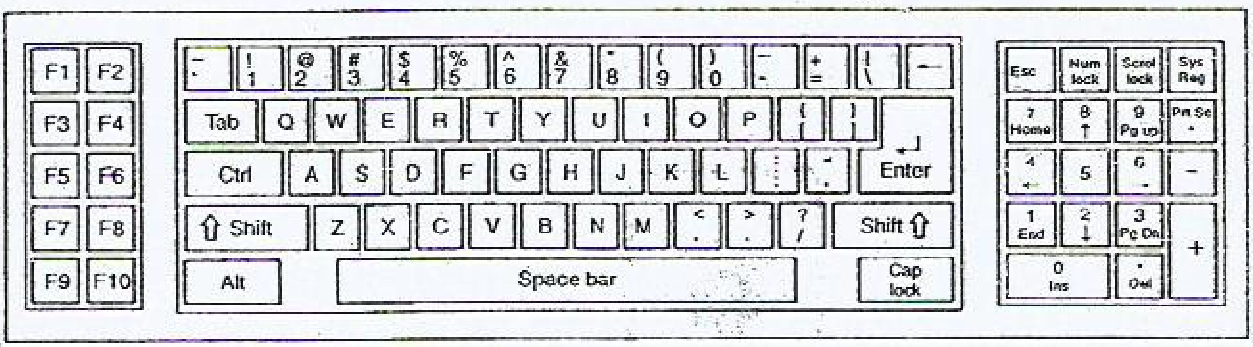


Figure 7.1. The layout of keys on a QWERTY keyboard.

Keyboards are usually available in two kinds of layout that is QWERTY and DVORAK keyboard. Most typewriters and computer keyboards are QWERTY keyboards.

The speed of input entry on a keyboard is limited by the typing speed of the operator.

Mouse

Mouse is a pointing device, which controls the position of the cursor or pointer on the screen. The mouse is a palm-size device with a ball built into the bottom. It has one or more buttons and attached to the computer by a cable.

When the user rolls the mouse across the flat surface, such as a desk, the ball on its undersides rotates. This causes the cursor to move in a corresponding direction on the monitor. If the user rolls the mouse to the left the pointer on the screen also moves to the left same as in the right side. The cursor can be moved in any of the four directions. The cursor is positioned on an appropriate object on the screen and the button on the mouse is clicked to select the object. The mouse interface is also called point and click interface.

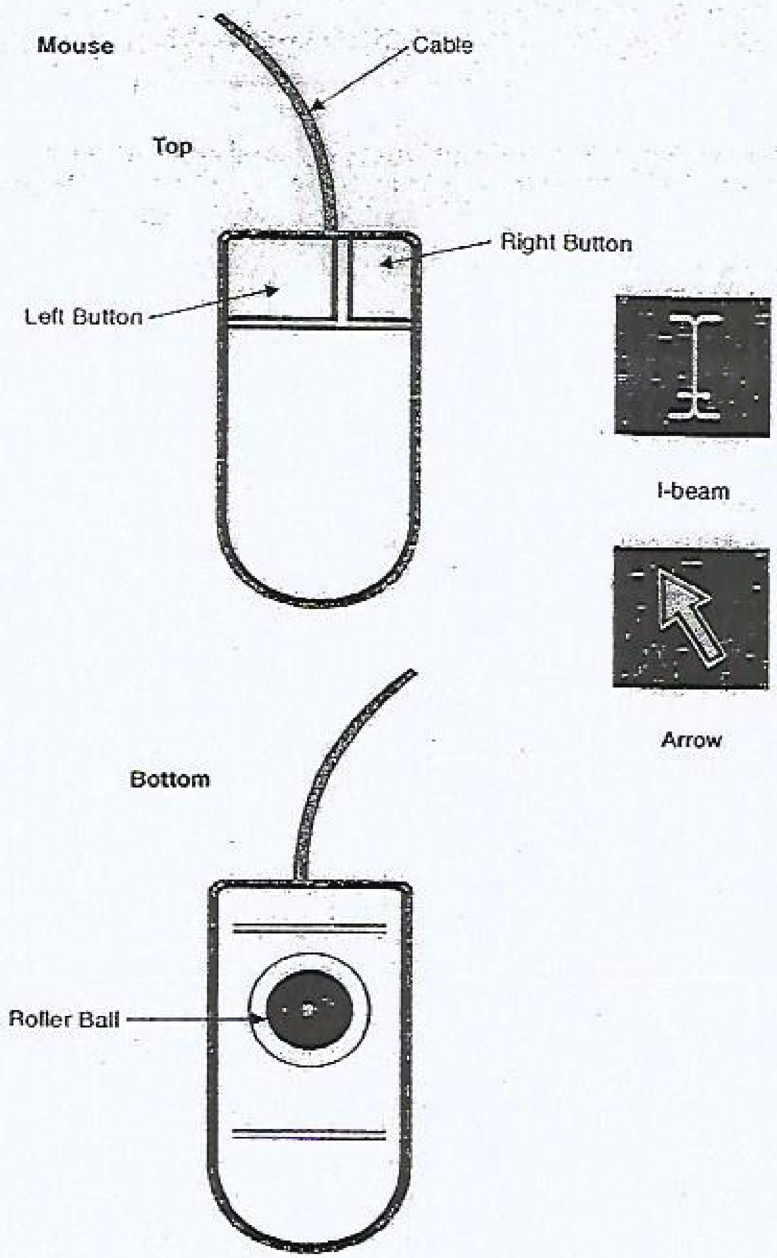


Figure 7.2.

1. Mechanical Mouse

The first mouse was a mechanical design based on a small ball that protruded through its bottom and rotated as the mouse was pushed along a surface. Switches inside the mouse detected the movement and relayed the direction of the ball's rotation to the host computer.

The mechanical mouse works on just about any surface. In general, the rotating ball has a coarse texture and is made from a rubbery compound that grips even on smooth surfaces.

The mechanical mouse requires that you move it across a surface of some kind, but all too many desks do not have enough free space to give the mouse a good run. Also, the mechanical mouse tends to pick up dirt and lint that can impede its proper operation. Clean it regularly even if you think your desktop is spotless.

2. **Optical Mouse**

The alternative technology to the mechanical mouse is the optical mouse. Instead of a rotating ball, the optical mouse uses a light beam to detect movement across a surface. No moving parts means that the optical mouse has less of a chance to get dirty or break.

Joystick

A joystick is a pointing device which works on the same principle as a trackball. Joystick has a gearshift-like lever that is used to move the pointer on the screen. The lever can be pushed in any direction. When released it returns to its original position. On most joysticks, a button on top is used to select option. Joysticks are commonly used to play games.

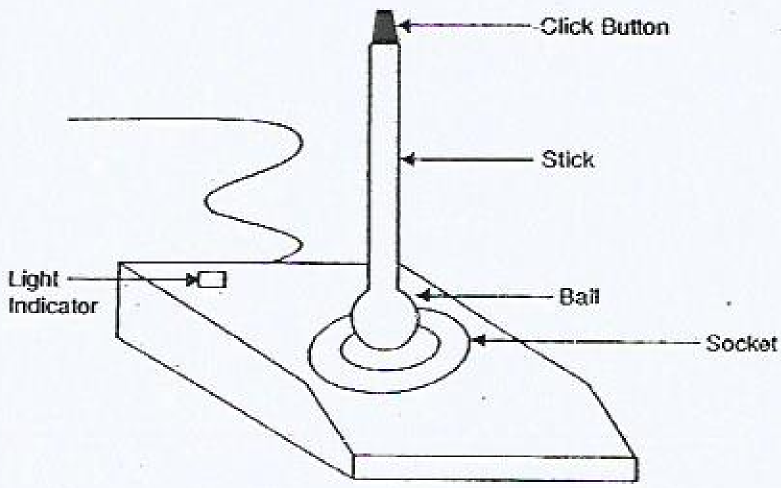


Figure 7.3. A joystick.

Touch Pad

The touch pad is a stationary pointing device that many people find less tiring to use than a mouse or a track ball. The movement of a finger across a small touch surface is translated into cursor movement on the computer screen. The touch sensitivity surface may be just 1.5–2 inch square, so the finger does not have to move much. Its size makes it most suitable for the laptops.

Light Pen

A light pen is also a pointing device having the size and shape similar to that of a regular pen. The pen consists of a photocell placed in a small tube. As the user moves the tip of the pen over the screen surface, the light pen senses the light coming from the screen area and sends a signal to the computer. Thus to identify a specific location, the light pen is very useful. But the light pen provides no information when held over a blank part of the screen because it is a passive device with a sensor only.

The light pen is primarily used for graphics work. A typical use of light pens is in drawing lines of various thickness.