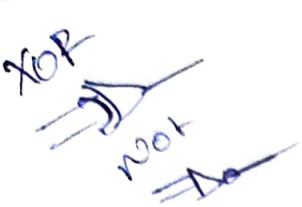
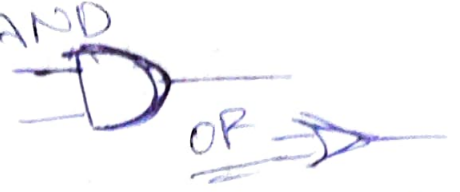


1 chapter
Data Storage



* **Bits:** Inside today's computers information is encoded as patterns of 0s and 1's. These digits are called bits

Some times pattern of bits are used to represent numeric values; sometimes they represent characters in an alphabet and punctuation marks, sometimes they represent images and sounds

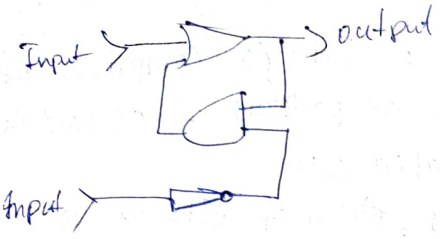
* **Boolean operations:** Operations that manipulate true/false values are called Boole — OR

Boolean operators: AND, OR and XOR

* **Gates:** A device that produces the output of a Boolean operation when given the operation's input value is called a gate
- Gates can be constructed from a variety of technologies such as gears, relays and optic devices

* **Flip-Flop:** A flip-flop is a circuit that produces an output value 0 or 1, which remains constant until a pulse from another circuit causes it to shift to the other value.

A simple flip-flop circuit



* First, it demonstrates how device can be constructed from gates, a process known as digital circuit design

- Second, the concept of a flip-flop provides an ex of abstraction and the use of abstract tools.

- The third purpose of introducting the flip-flop is that it is one means of storing a bit within a modern computer

- Very large-scale integration (VLSI): Thus, many flip-flops, constructed as very small electrical circuits, can be used inside a computer as a means of recording info that is encoded as patterns of 0s and 1s. Indeed, technology known as (VLSI), which allows millions of electrical components to be constructed on a wafer (called a chip), is used to create a miniature devices containing millions of flip-flop along with their controlling circuitry.

- In turn, these chips are used as abstract tools in the construction of computer systems. In fact, in some cases VLSI is used to create an entire computer system on single chip.

* Stream: A long string of bits is often called a stream.

* Hexadecimal notation:
To simplify the representation of such bit patterns, [101101010011] we usually use a short hand notation called hexadecimal notation.

4 bits / 2^4 = 16

Bit pattern	Hexadecimal notation
0000	0
0001	1
0010	2
0011	3
0100	4
0101	5
0110	6
0111	7
1000	8
1001	9
1010	A
1011	B
1100	C
1101	D
1110	E
1111	F

* Main memory:

For purpose of storing data, a computer contains a large collection of circuits (such as flip-flops) each capable of storing a single bit.

This bit reservoir is known as the machine's **main memory**.

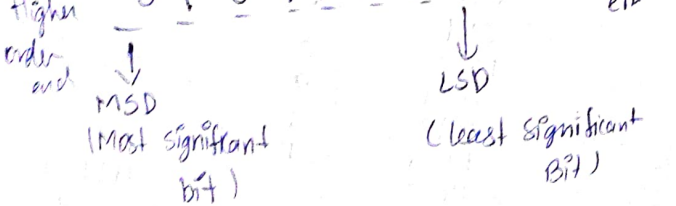
* Memory Organization:

- **Cell:** A computer's main memory is organized in manageable units called **cells**.

  cell size = 8 bits

- **Byte:** A string of eight bits called byte

- Thus typical memory cell has capacity of 1 byte
- Small computers used in such household activity have few hundred cells
- large computers have few ~~thousand~~ billions of cells



- **Address:** To identify individual cells in a computer's main memory, each cell is assigned a unique "name" called its address.

- Other circuits can get data from the memory by electronically asking for the contents of a certain address called a read operation.

- they can record information in the memory by requesting that a certain bit pattern be placed in the cell at a particular address called a write operation.

* RAM

It is the internal memory of the CPU for storing data, program, and program result. It is read/write memory which stores data until machine is working. As soon as the machine is switched off, data is erased.

-> Static RAM (SRAM)

The word static indicates that the memory retains its contents as long as power is being supplied. However, data is lost when the power gets down due to volatile nature.

SRAM chips uses a matrix of 6-transistors no-capacitor

→ Dynamic Ram (DRAM):

DRAM, unlike SRAM, must be continually refreshed in order to maintain the data. This is done by placing the memory on a refresh circuit that rewrites the data several hundred times per second.

→ Mass storage:

Due to the volatility and limited size of a computer's main memory, most computers have additional memory devices called mass storage (or secondary storage).

Ex; magnetic disks, CDs, DVDs, magnetic tapes, and flash drivers.

Advantages: less volatility, large storage capacities, low cost, and in many cases, the ability to remove the storage medium from the machine.

- On-line: means that the device or info is connected and readily available to the machine without human intervention.
- Offline: means that human intervention is required before the device or info can be accessed by the machine - perhaps because the device must be turn on, or the medium holding the info must be inserted into some mechanism.

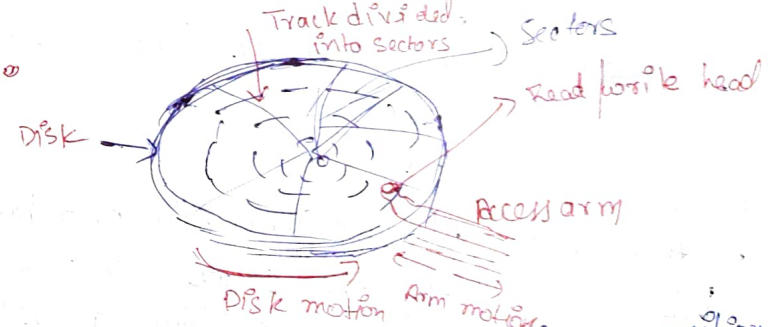
[The term on-line and off-line are often used to describe devices that can be either attached to or detached from a machine]

→ Magnetic System:

↳ common ex is magnetic disk

- Magnetic disk: is a storage device that uses a magnetization process to write, rewrite and access data. It is covered with a magnetic coating and stores data in the form of tracks, spots and sectors.

  Ex; Hard disks, zip disks and floppy disks.



- Tracks: A drive head, while in one position can read or write a circular ring, or band called a track.
- Cylinder: Each time the read/write heads are repositioned, a new set of tracks - which is called a cylinder.

- **Sectors:**

Since a track can contain more info than we would normally want to manipulate at any one time, each track is divided into small arcs called sectors on which info is recorded as a continuous string of bits

1 sector = 512 bytes to a few KB

① ~~T~~Bits within a sector on a track near the outer edge of the disk are less compactly stored than ~~t~~ those tracks near center.

② ~~Outer~~ Outer tracks are longer than the inner ones.

- **Zoned-bit recording:**
  - A method of recording data on a hard disk drive whereby the sectors per track on the drive are not consistent across the platter.
  - In general, tracks closest to the center have fewer sectors than tracks towards the outside of the platter where the tracks are larger and can fit more sectors.
  - Platter rotates at a constant angular velocity the clock speed, or clock rate, changes as the read/write head moves from one zone to an

- **Formatting disks:**

Several measurements are used to evaluate a disk system performance:

- **Seek time:**

the time required to move the read/write heads from one track to another.

- **Rotation delay or latency time:** half the time required for the disk to make a complete rotation, which is the average amount of time required for the desired data to rotate around to the read/write head once the head has been positions over the desired track.
- **Access time:** the sum of seek time and rotation delay
- **Transfer rate:** the rate at which data can be transferred to or from the disk.

- Magnetic tape :

An older form of mass storage using magnetic technology is called magnetic tape.

- In these systems, info is recorded on the magnetic coating of a thin plastic tape that is wound on a reel for storage.
- To access the data, the tape is mounted in a device called tape drive that typically can read, write and rewind the tape under control of the computer.
- Tape drives range in size from small cartridge units called streaming tape units, which use tape similar in appearance to that in stereo system to older, large reel-to-reel units.

* Optical Systems :

Another class of mass storage systems applies optical technology.

Ex = (CD) compact disk

- Compact disk

- These disks are 12 cms in diameter and consists of reflective material covered with a clear protective coating.
- Info is recorded on them by creating variations in their reflective surfaces.
- This info can then be retrieved by means of a laser beam that detects irregularities on the reflective surface of the CD as it spin.
- (Compact disk - digital audio) CD-DA
- CD technology was originally applied to audio recordings using a recording format known as CD-DA, and the CDs used today for computer data storage use essentially the same format.
- In particular, info on these CDs is stored on a single track that spirals around the CD like a groove in an old-fashioned record.

- This track divides into units called sectors
  capacity = 2KB
  1/75 of a second of music in the case of audio recording

Remain T-B
pg-32

- Tradition CDs have capacities in the range of 600 to 700 MB.

* DVDs (Digital Versatile Disks):

- Constructed from multiple, semi-transparent layers, that serves as distinct surfaces when viewed by a precisely focused laser.
- Provides storage capacities of several GB.
- Such disks are capable of storing lengthy multimedia presentations, including entire motion pictures

* Blu-ray Disks (BDs):

- Blu-ray technology, which uses a laser in the blue-violet spectrum of light (instead of red).
- Able to focus its laser beam with very fine precision
- As BDs provides over 5 times capacity of a DVD

* Flash Drives

Flash memory:

- Flash memory technology has the potential of alleviating this drawback.
- Bits are stored by sending electronic signals directly to the storage medium where they cause electrons to be trapped in tiny chambers of silicon dioxide, thus altering the characteristics of small electronic circuits
- these chambers are able to hold their captive e⊖s for many years
- It is suitable for off-line storage of data.

Flash memory devices called flash drives.

Applications of flash technology:

1. SD (Secure digital) memory cards:
   These provide up to two GBs of storage and are packaged in a plastic rugged wafer about the size a postage stamp

2. SDHC (High capacity) memory cards
Can provide upto 32 GBs

3. SDXC (Extended capacity) memory cards
may exceed a TB

∴ Thus, they are ideal for digital cameras, smartphones, music players, car navigation systems and a host of other electronic applications.

* Files storage and Retrieval:

- Files: Info stored in a mass storage system is conceptually grouped into large units called files.

Ex: They may contain a program, a music recording, a photograph etc.

- Physical record:

A block of data conforming to the specific characteristics of a storage device is called a physical record.

Thus, a large file stored in mass storage will typically consist of many physical records.

Logical records:

- a file often has natural divisions determined by the info represented.
- For ex: a file containing info regarding a company's employees would consist of multiple units, each consisting of the info about one employee. Or a file containing a text document would consists of paragraphs or pages.

These naturally occurring block of data are called logical records.

* logical records often consist of smaller units called fields.

- Key field:

A field in a record that holds unique data which identifies that record from all the other records in the file or database.

Acc num, product code and customer name are typical key fields.

- The value held in a key field is called key

- Buffer:

~~An area of~~ a region of a physical memory storage used to temporarily store data while it is being moved from one place to another. However, a buffer may be used when moving database b/w processes within a computer.

* ANSI — American National Standards Institute
* ASCII — American standard code for Information Interchange
* ISO — International Organization for standardization
* UNIcode — A universal character encoding standard that assigns a code to every character and symbol in every language in the world. Since no other encoding standard supports all languages, unicode is the only encoding standard that ~~you~~ ensures that you can retrieve or combine data using any combination of langu—

- Text file: A file consisting of a long sequence of symbols encoded using ASCII or unicode is often called a text file
- Text editor: It is imp to distinguish b/w simple text files that are manipulated by utility programs called text editors
- Word processors -

- Representing images:

Pixel: One means of representing an image is to interpret the image as a collection of dots, each of which is called a pixel, short for "picture element".

Bitmap: The appearance of each pixel is then encoded and the entire image is represented as a collection of these encoded pixels. Such a collection is called a bit map.

- Brightness component which is called the pixel's ~~Luminance~~ Luminance, it is essentially ~~=~~ the sum of the red, green, and the blue components.

which is essentialy the sum of the red, green and blue components

The other two components called the blue chrominance and the red chrominance, are determined by computing the difference b/w the pixel's luminance and the amt of blue or red ~~tang~~ light in the pixel.

Disadvantage:

1. Image as bit maps is that an image cannot be rescaled easily to any arbitrary size.

- The base ten and binary system

a. Base ten system

3 7 5]— Representation
hun/ten/ones]→ Position's quantity

b. Base two system

1 0 1 1]— Represen
Eight/Four/Two/ones]→ Position's quantity

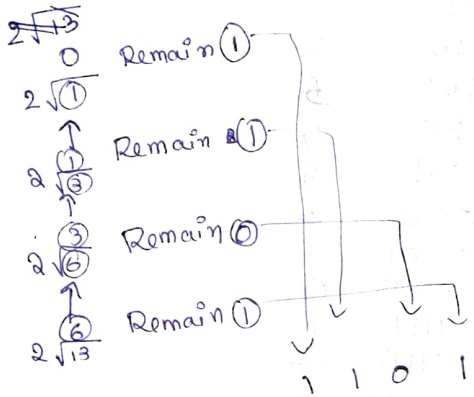
- Decoding the Binary representation

1 0 0 1 0 1]→ Pattern Bit
$2^5 2^4 2^3 2^2 2^1 2^0$

$= 2^5 \times 1 + 2^4 * 0 + 2^3 \times 0 + 2^2 \times 1 + 2^1 \times 0 + 1^0 \times$

$= 37$ total

- Obtain Binary rep of 13



Calculate until the quotient becomes zero

- Two's complement notation ⇒

→ This system uses a fixed number of bits to represent each of the values in the system.
→ Each value is represented by a pattern of 32 bits.

a. Using patterns of length three

Bit pattern	Value represented
011	3
010	2
001	1
000	0
111	-1
110	-2
101	-3
100	-4

b. Using patterns of length four

Bit pattern	value
0111	7
0110	6
0101	5
0100	4
0011	3
0010	2
0001	1
0000	0
1111	-1
1110	-2
1101	-3
1100	-4
1011	-5
1010	-6
1001	-7
1000	-8

* The left most bit is sign bit

negative values = 1 (represented by pattern)

positive values = 0

Complement

The complement of the pattern is the pattern obtained by changing all the 0s to 1s and all 1s to 0s

Ex: 0110 and 1001 are complement

Excess notation:

Another method of representing integer values is excess notation.

length four

Bit	Value
1111	7
1110	6
1101	5
1100	4
1011	3
1010	2
1001	1
1000	0
0111	-1
0110	-2
0101	-3
0100	-4
0011	-5
0010	-6
0001	-7
0000	8

length three

111	3
110	2
101	1
100	0
011	-1
010	-2
001	-3
000	-4

* Excess system based on patterns
  of length five would be called excess
  16 notation

* Storing Fractions :-
  To store fractions we use floating point
  notation