DATA STRUCTURES AND ALGORITHMS

Data Organization in Computing

By
Samuel K. Opoku
Computer Science Department
Kumasi Technical University

Video Link

https://blackboard.skopoku.org/tutorials/data_rep.mp4

2

CONTENT

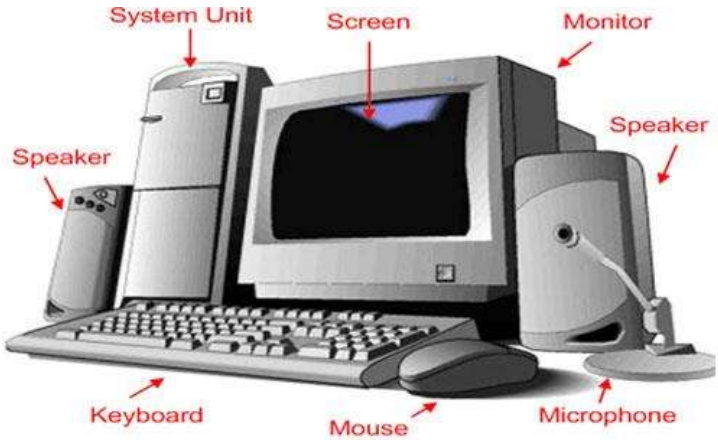
- Computer Architecture and Operations
- Data Representation and Processing
- Internal Representation of Data

8

Computer Architecture and Operations

9

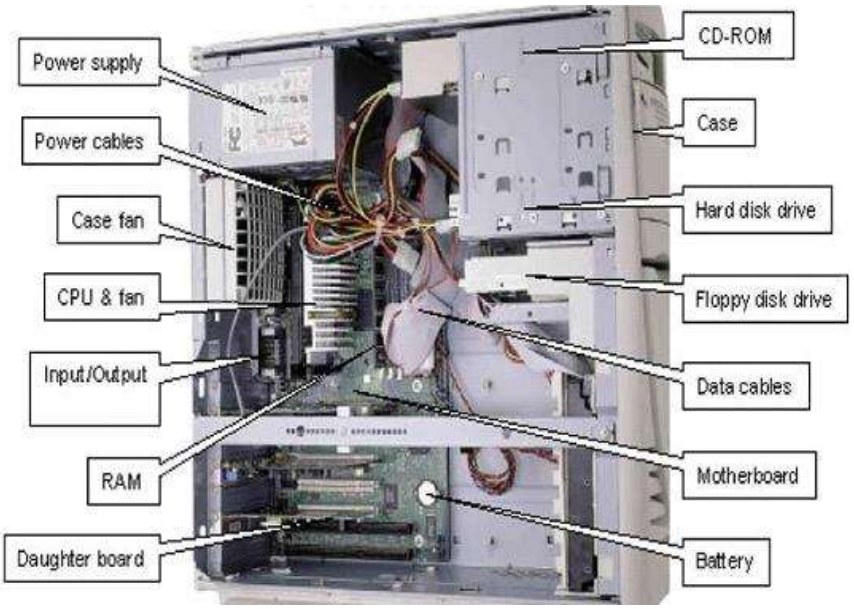
COMPONENTS OF A COMPUTER SYSTEM



Recall the functions of the various components of the computer system

10

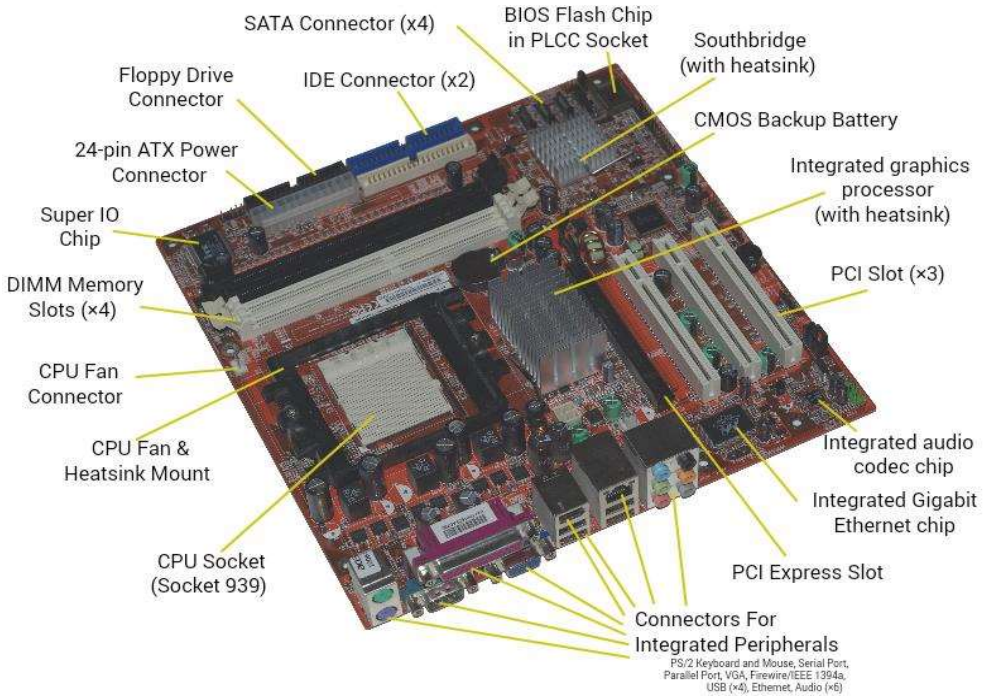
COMPONENTS OF A SYSTEM UNIT



Recall the functions of the various components

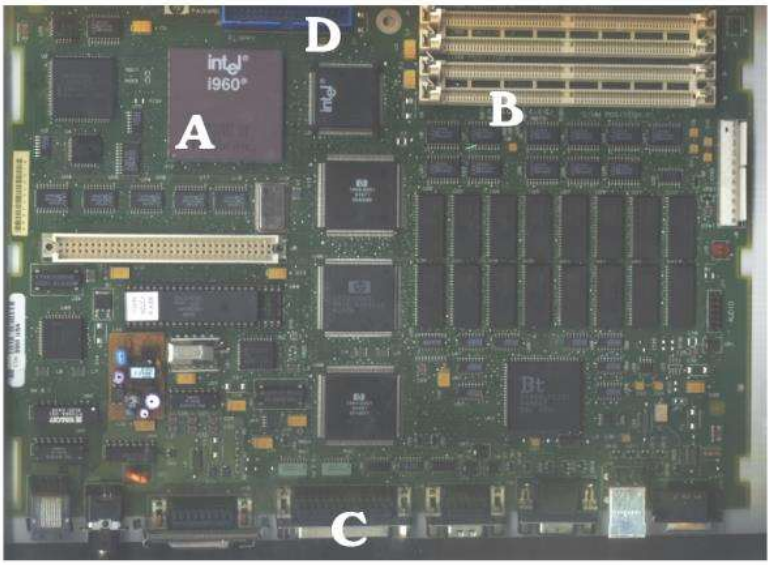
11

COMPONENTS OF THE MOTHER BOARD



12

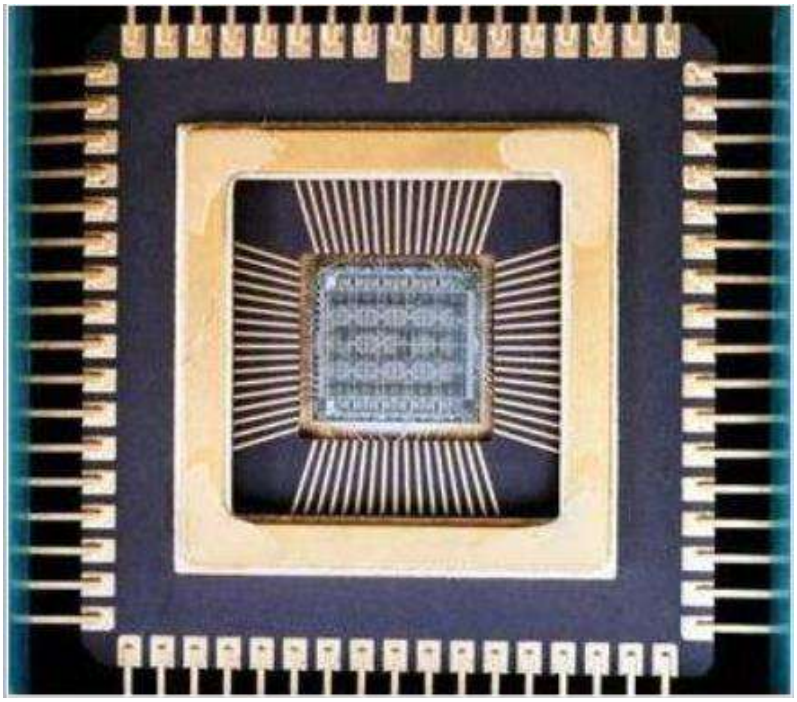
COMPONENTS OF THE MOTHER BOARD



Components on the motherboard are usually made as **Integrated Circuit (IC)** which consists of thousands or millions of transistors, capacitors and resistors acting as computer memory, amplifiers, oscillators, counters and timers

13

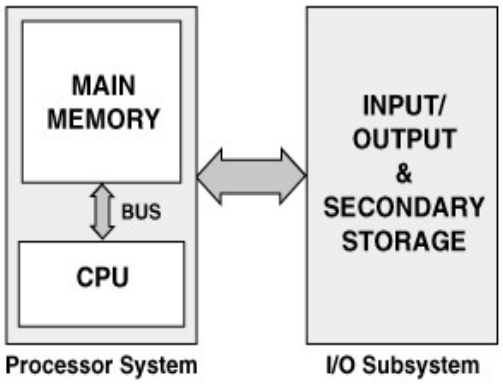
- A typical inner architecture of an Integrated Circuit



16

SCHEMATIC REPRESENTATION

- based on stored program design
- processor system
  - CPU
  - memory
- input/output system
  - input/output devices
  - secondary storage



18

CENTRAL PROCESSING UNIT

- CPU is a complex set of electronic circuitry that controls and executes stored program instructions

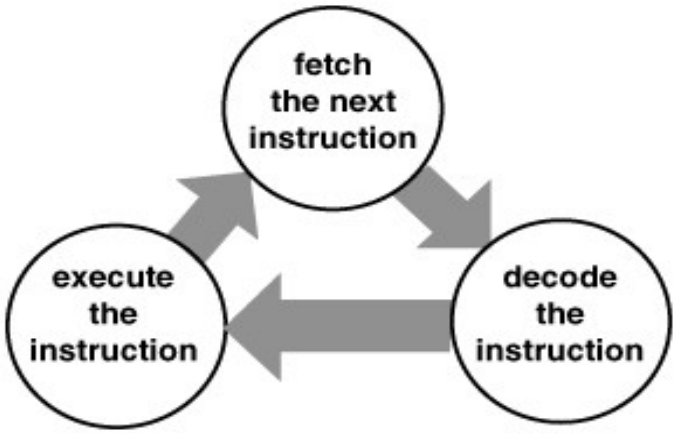
- **Two parts**
  - **Control Unit (CU):**
    - Directs the computer system to execute stored program instructions. Must communicate with memory and ALU. Sends data and instructions from secondary storage to memory as needed

  - **Arithmetic Logic Unit (ALU)**
    - Executes all arithmetic and logical operations.
    - Arithmetic operations (+, -, *, /)
    - Logical operations: Compare numbers, letters, or special characters using operators like >, <, =

19

CENTRAL PROCESSING UNIT

- Manages the instruction-execution cycle
  - FETCH – DECODE – EXECUTE
- Coordinates the activities of other devices



20

DATA STORAGE AND CPU

- **Two types of storage**:
  - **Primary storage (memory)**: Stores data temporarily. CPU refers to it for both program instructions and data
  - **Secondary storage**: Long-term storage. Stored on external medium, such as a disk
- CPU cannot process data from disk or input device. **It must first reside in memory**. Control unit retrieves data from disk and moves it into memory. Items are sent to ALU for processing. Control unit sends items to ALU, then sends back to memory after processing
- Data and instructions are held in memory until sent to an output or storage device or program is shut down

21

MAIN MEMORY

- (fast access) storage device for preserving binary **data and instructions**
- memory is divided into units such as **bytes** or **words**
- each is usually a standard size or **fixed-length**
- each memory word has a unique address for **random access**
- **Types:**

Random Access Memory (RAM)

- readable
- writeable
- usually volatile (e.g., Dynamic RAM or DRAM)
- general storage

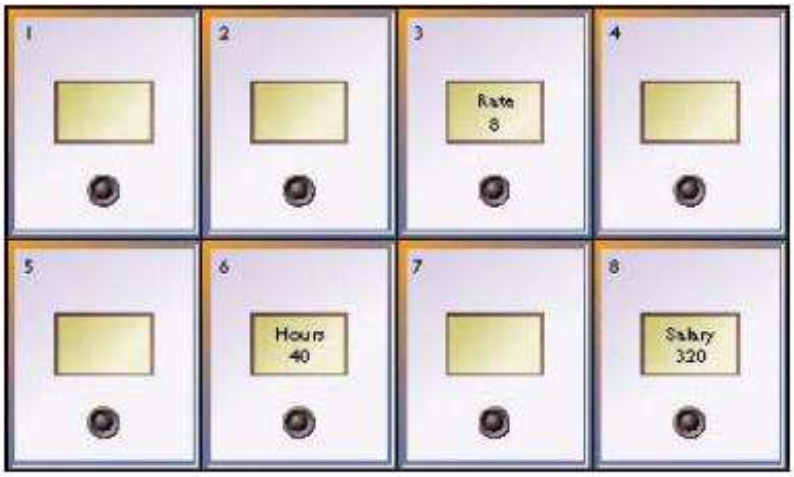
Read Only Memory (ROM)

- readable
- permanent
- Nonvolatile
- special-purpose storage for data and instructions

22

MAIN MEMORY

- Each memory location has an address: A unique number, much like a mailbox. May contain only one instruction or piece of data . When data is written back to memory, previous contents of that address are destroyed
- Location is referred to by number but programming languages use a symbolic (named) address called variables, such as **Hours** or **Salary**



23

REGISTERS

Special-purpose
High-speed
Temporary storage
Located inside CPU

Instruction register
Holds instruction currently being executed

Status Register
Holds status of ALU operations

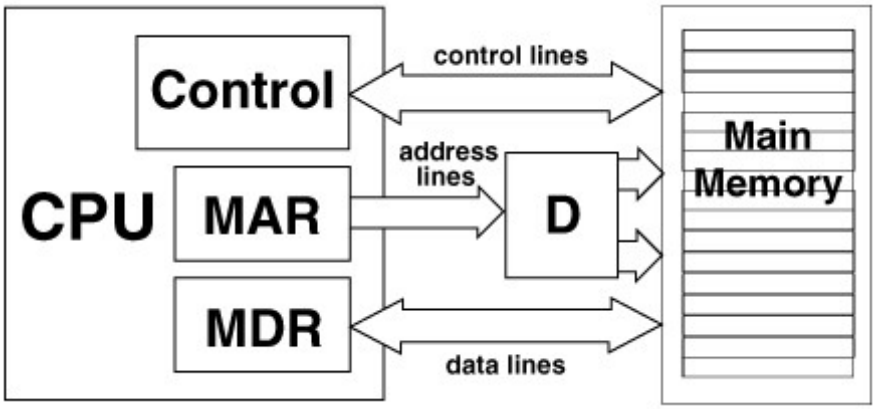
Data register

Holds data waiting to be processed

Holds results from processing

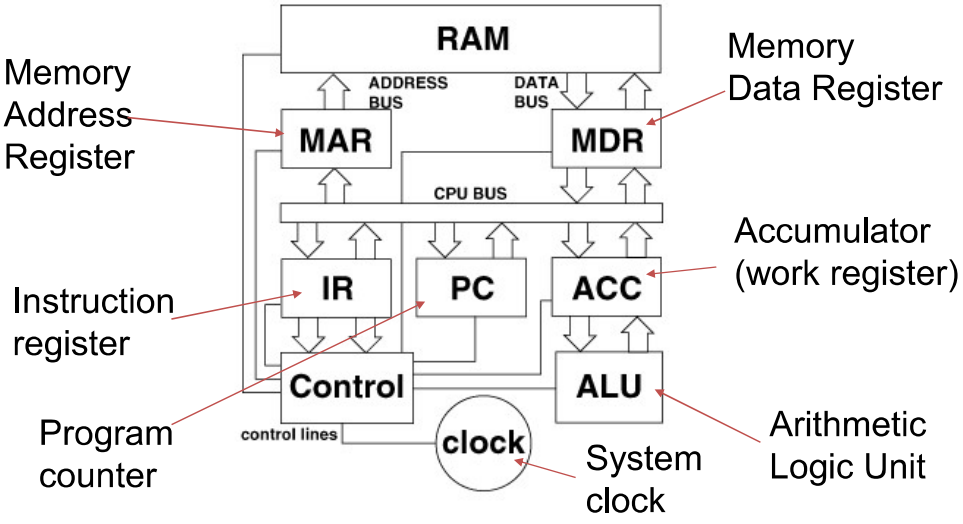
24

THE CPU AND MAIN MEMORY



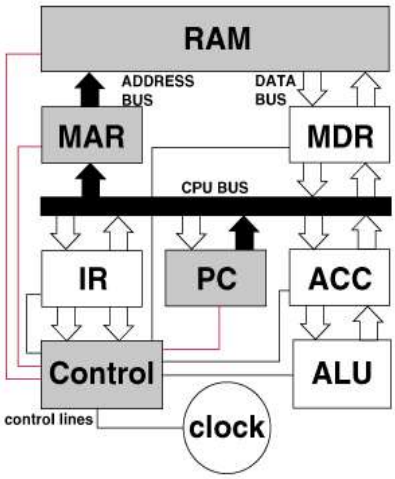
25

INSIDE THE CPU



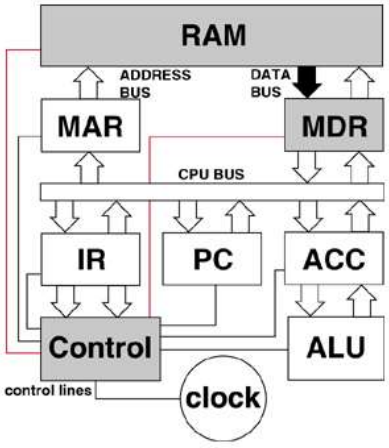
26

FETCH THE INSTRUCTION



1. Address of the next instruction is transferred from PC (Program Counter) to MAR (Memory Address Register)

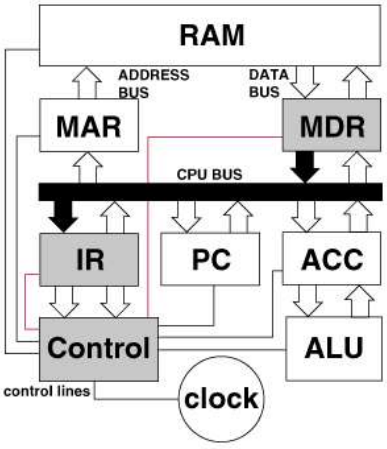
2. The instruction is located in memory



3. Instruction is copied from memory to MDR (Memory Data Register)

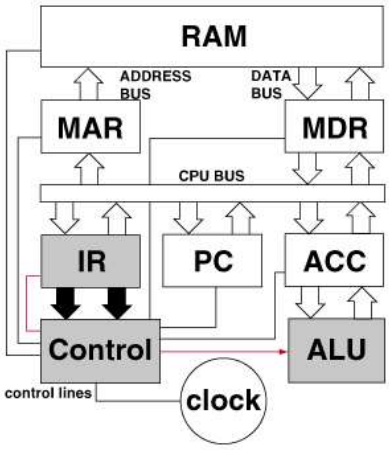
27

DECODE THE INSTRUCTION



4. Instruction is transferred to and decoded in the IR (Instruction Register)

EXECUTE THE INSTRUCTION



5. Control unit sends signals to appropriate devices to cause execution of the instruction

28

SYSTEM CLOCK

- **Function of the system clock**: synchronizes all the components on the mother board. If the clock is high, all the components do their work. Because the CPU needs to perform more operations per time than the mother board, the CPU clock is much higher.

- If the clock speed (or clock rate) is say 2GHz, it means the clock pulses (clock cycles) 2 billion times per second. That is, under ideal circumstances, the processor/CPU should execute 2 billions instructions per second. However, in real world scenarios, the CPU is often stuck waiting for other components to answer back, especially RAM and disk. Delays can occur every time a program has to make a decision about which code path to go (conditionals)

Note: Modern processors try to beat this by predicting which path will be taken before that point is actually reached, but if the prediction turns out to be wrong you incur additional penalties.

Thus increasing clock frequency does increase the operations per second, but number of operations performed by the CPU is largely defined based on the underlying CPU architecture itself and the actual code running on the CPU

29

TYPES OF PROCESSOR OPERATIONS

- **Data Movement Operations**
  - moving data from memory to the CPU
  - moving data from memory to memory
  - input and output
- **Arithmetic and Logical Operations**
  - integer arithmetic
  - comparing two quantities
  - shifting, rotating bits in a quantity
  - testing, comparing, and converting bits
- **Program Control**
  - starting a program
  - halting a program
  - skipping to other instructions
  - testing data to decide whether to skip over some instructions

30

OPERATIONS WITHIN DIGITAL SYSTEMS(1/2)

Four operations to which information may be subjected to in digital systems are:

1. **Storage**: Computer stores data and instructions either temporally or permanently using storage devices such as registers and main memory
2. **Input / Output**:
   - The computer performs input or output or both operations also internally
   - **Input operation** – Data or instructions are given to the computer through an input device such as the keyboard for processing by the computer
   - **Output operation** – Processed data is given to external systems through an output device such as monitor

31

OPERATIONS WITHIN DIGITAL SYSTEMS(2/2)

3. **Transmission**:
   - Data is transmitted or communicated from one device to another along communication channels
   - In some instances, data and/or information is brought from far away through space through communication devices such as satellite systems

4 **Processing**: The act of processing data describes the manipulation of the representation of the data given to the computer in order to produce expected output.

32

Data Representation and Processing

34

DEFINITIONS (1/2)

- **Data**:
  - It refers to any or all facts, numbers, letters and symbols that describe an object, idea conditions, situations, event or some other factors
  - Data is intended to depict a real life phenomenon but in themselves do not make any sense
  - In computing, data refers to symbols (materials) to be operated upon by the computer

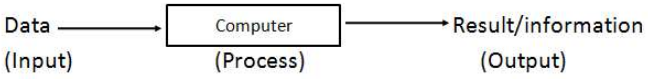
- **Information**:
  - It refers processed data that makes sense and meaningful to the recipient of the information

35

DEFINITIONS (2/2)

- **Mathematically: Information = Data + Meaning**

- Data and information are sometimes used interchangeably and that the context in which either is used is what matters.



- **Knowledge**:
  - It is derived from operation of data and/or information that goes a long way to strengthen decisions that users make
  - Knowledge is a generalization of a body of ideas derived from operations performed on data and/or information

36

DATA REPRESENTATION AND PROCESSING (1/2)

- Basically, data is an abstract concept and it is natural to associate symbols with many concepts.

  Examples: What is '2' and 'B'?

  2 is not the abstract number two. But it is a symbol that represents it. Similarly, B is a symbol representing the second letter of the alphabets

- Classification of data processing system:
  - **Discrete (or Digital) Systems**: They use symbols to represent numbers or quantities
  - **Analogue Systems**: They use continuous physical measures such as length or voltage to represent numbers or quantities.

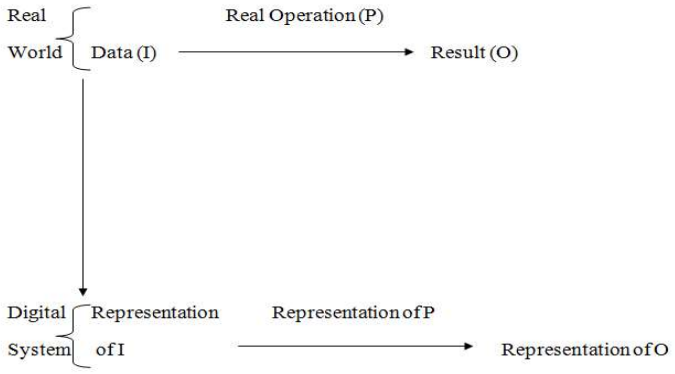
NB: There are also systems called **hybrids** which combine the methods of both discrete and analogue systems

37

DATA REPRESENTATION AND PROCESSING (2/2)

- Data processing in digital systems is done in such a way that if a real operation (P) is performed on the data (I) to give result (O), then a representation of P is performed on a representation of I to give a representation of O.

- This is summarised in the figure below:



38

Internal Representation of Data

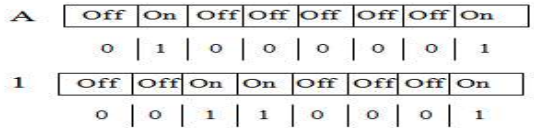
40

INTERNAL REPRESENTATION OF DATA

- Both instructions and the data which the computer deals with are made up of symbols called **characters** that are coded in the form suitable for electronic storage

- **Common Coding Schemes:**
  - **ASCII**: Stands for American Standard Code for Information Interchange. Most widely used standard. Used on virtually all personal computers
  - **EBCDIC**: Extended Binary Coded Decimal Interchange Code. Used primarily on IBM and IBM-compatible mainframes
  - **Unicode**: Designed to accommodate alphabets of more than 256 characters. Uses 16 bits to represent one character (65,536 possible values)

- Below is a typical storage code for the character A and the decimal digit 1



41

INTERNAL REPRESENTATION OF DATA

- The storage is accomplished through tiny magnetic or electrical components which takes one of the two states – "ON" or "OFF". Eight individual components (called a **byte**): each storing one bit of information is used to store a particular character

- **Transistors** are electronic switches that may or may not allow electric current to pass through. If current passes through, switch is on, representing a 1 bit Otherwise, switch is off, representing a 0 bit

- A Group of the individual components (usually transistors) is forced into the pre-established pattern of "ON" and "OFF" states corresponding to the character's code

- Thousands of groups of components are provided so that the computer will be able to store all the instructions in a program as well as the data it must process at any given time

42

BYTE AND WORD

- **Byte**: A group of 8 bits.
  - Each byte has 256 (2^8) possible values
  - For text, stores one character: Can be letter, digit, or special character
  - Memory and storage devices measured in number of bytes

- **Word**: The number of bits the CPU processes as a unit
  - Typically a whole number of bytes
  - The larger the word, the more powerful the computer
  - Personal computers typically 32 or 64 bits in length

- **Storage Size**
  - **Kilobyte**: 1024 (2^{10}) bytes
  - **Megabyte**: roughly one million (2^{20}) bytes
  - **Gigabyte**: roughly one billion (2^{30}) bytes
  - **Terabyte**: roughly one trillion (2^{40}) bytes

43

CALCULATIONS

Note: 1024 Kilo = 1 Mega and 8bits = 1B or a byte

1GB RAM = 1024MB RAM which contains $8(1024)^2$ transistors

(convert byte to bits by multiplying it by 8)

1. How many transistors (electrical components) are required by RAM of the following sizes:

   a. 128MB b. 512MB c. 5GB d. 128KB

2. How many characters will a RAM of the following sizes store when a byte is used to store a character:

   a. 156MB b. 225MB c. 2GB d. 98KB

44

47

Assignment

Describe how the computer handles images, audio and video files. Your report should be precise and concise without any ambiguity but should NOT be more than 200 words

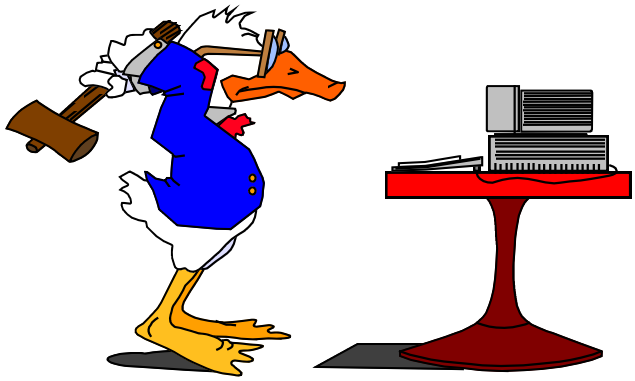
Deadline: Sunday 28th May, 2023, 11:59pm

Format: Use a text file.

Note:

- **Your work will be ignored if the file type is not text (.txt) and the report contains more than 200 words**
- **The work should be done individually.**

48



HIT ANY KEY TO END DAY 1