Computer Science (9618) - AS Level

https://keddy-edtech.org/

1. Information Representation

1.1. Data Representation

- The two fundamental characteristics of any number system are:
  - A base: The number of different digits that a system can use to represent numbers
  - Place value: The specific-value of a digit based on its position within a number
- Denary - Base 10
- Binary Systems - Base 2
  - Possible bits (binary digits): 0 and 1
  - All data and characters are represented in binary

128	64	32	16	8	4	2	1
0	0	0	0	0	0	0	0

- For example, 65 in binary is *01000001*

- Denary vs. Binary prefixes:

Denary Prefix	Factor Value	Binary Prefix	Factor Value
kilo- (k)	×103 ×106	kibi- (Ki)	×210 ×220
mega- (M)	×109 ×1012	mebi- (Mi)	×230 ×240
giga- (G)		gibi- (Gi)	
tera- (T)		tebi- (Ti)	

- Binary Coded Decimal (BCD)
  - Binary representation where each positive denary digit is represented by a sequence of 4 bits (nibble).
  - Only certain digits are converted to BCD because particular digits represent a digit greater than 9.
  - For example, to represent 429 in BCD:
  - Convert each individual digit of the number to its binary equivalents -
    1. *4 = 0100*
    2. *2 = 0010*
    3. *9 = 1001*
  - Concatenate the 3 nibbles (4-bit group) in order to produce BCD: *0100 0010 1001*
- Practical applications
  - A string of digits on any electronic device displaying numbers (like in calculators)
  - Accurately measuring decimal fractions
  - Electronically coding denary numbers
- Two's Complement
  - We can represent a negative number in binary by making the most significant bit (MSB) a sign bit, which indicates whether the number is positive or negative.
  - Converting from negative denary to binary two's complement (example -42):
    1. Find the binary equivalent of the denary number (ignoring the negative sign) | 42 = 101010
    2. Add extra 0 bits before the MSB, to format the binary number to 8 bits | 00101010
    3. Convert binary number to one's complement (flip the bits) | 11010101
    4. Convert binary number to two's complement (add 1) |1010101 + 1 = 11010110
  - Converting binary two's complement into denary (example 11010110):
    1. Flip all the bits | 00101001
    2. Add 1 | 00101010
    3. Convert binary to denary and put a negative sign) | -42
- Maximum positive number in 8 bits: 127
- Maximum negative number in 8 bits: -128
- Hexadecimal Systems - Base 16:
  - Possible digits: 0 to 9 and A to F, where A to F represent denary digits 10 to 15 Practical applications:
    - Defining colours in HTML
    - Defining Media Access Control (MAC) addresses
    - Assembly languages and machine code
    - Debugging via memory dumps
  - Example - A5 in Denary = (16×10) + (1×5) = *165*
  - Example - 65 in Hexadecimal = 65÷16 = 4 | Remainder 1 | ∴ = *41*
- Character Sets:
  - A character set generally includes upper & lower case letters, number digits, punctuation marks, and other characters.
  - Character sets use different binary representations for each character via character encoding

---

- Character Encoding Standards:

ASCII	Extended ASCII	Unicode
Only the English alphabet can be represented	ASCII's extension - Also includes most European languages' alphabets	Superset for ASCII & extended ASCII - recognized by various global languages
Each character encoding takes up 7 bits, hence 128 possible characters	ASCII extended to 8 bits, hence 256 possible characters	It has an excellent range of characters, using 2 or 4 bytes per character
Smaller storage space	2 or 4 times more storage space per character	

1.2. Multimedia - Graphics, Sound

- Bitmap Images
  - *Definition:* Bitmap images are created by assigning a solid color to each pixel using bit patterns.
  - *Encoding:*
    - Bit patterns are generated by considering each grid row as a series of binary color codes corresponding to each pixel's color.
    - These bit patterns are 'mapped' onto the main memory.
  - *Pixels:* The smallest picture element whose color can be accurately represented by binary code.
  - *File Header:* Bitmap images contain a file header with metadata, including image size, number of colors, etc.

- Image Resolution
  - *Definition:* the number of pixels that make up an image. Example:- 4096x3192 pixels.
  - *Effect:* Higher resolution results in sharper, more detailed images.

- Screen Resolution
  - *Definition:* The number of pixels that can be viewed horizontally and vertically on a device's screen.
  - *Calculation: Number of pixels=width×height* (e.g., 1680 × 1080 pixels).

- Colour depth
  - *Definition:* The number of bits used to represent the colour of a single pixel.
  - *Calculation:* An image with n bits has 2^n colours per pixel (e.g., a 16-color bitmap has 4 bits per pixel because 24 = 16)
  - *Effect:* Increasing color depth improves color quality but also increases file size.
  - *File Size Calculation:*
    - File Size = Number of Pixels × Color Depth
    - Convert bits to bytes by dividing by eight if necessary.
  - *Applications:* Scanned images and general computer usage, where small file sizes and easy manipulation are beneficial.

- Vector Graphics
  - *Definition:* Made up of drawing objects.
  - *Drawing Objects:* Mathematically defined constructs (e.g., rectangles, lines, circles).
  - *Drawing List:* A set of commands defining the vector (e.g., points, lines, shapes, and their attributes that define the image.)
  - *Properties:* Basic geometric data that determine the shape and appearance of each object.
  - *Encoding:* Data is encoded using mathematical formulas to generate properties for drawing lines and curves to create the image.
  - *Scalability:* Objects can be resized without losing quality, unlike bitmaps.
  - *Applications:* Company logos and other graphics that require scaling.

- Sound
  - *Analogue vs Digital:*
    - Analogue data consists of continuous electrical signals.
    - Digital data consists of discrete electrical signals.
  - *Sound Signals:* Vibrations through a medium, inherently analogue due to infinite detail.
  - *Conversion:*
    - Analogue signals are converted to digital signals by sampling.
    - The sound wave's amplitude is sampled at set time intervals.
    - These samples are encoded as a binary number sequence, providing a digital representation of the sound wave

---

- Sampling Rate
  - *Definition:* Number of samples taken per unit of time.
  - *Effect:* Increasing the sampling rate improves the accuracy of the digitized sound wave representation but increases file size.

- Sampling Resolution
  - *Definition:* Number of bits used to encode each sample.
  - *Effect:* Increasing sampling resolution improves the accuracy of digitized sound waves but increases file size.

- Bit Rate
  - *Definition:* Number of bits used to store 1 second of sound.
  - *Calculation: Bit Rate* = $Sampling\ Rate * Sampling\ Resolution$
  - *Length of Sound:* Measured in seconds.

1.3. Compression

- Definition: Compression is the process of reducing file size without significant loss in quality, resulting in:
  - Reduced time needed to search for data.
  - Faster transfer of compressed files, using less bandwidth than uncompressed files.

- Lossless Compression
  - *Definition:* A type of compression that allows original data to be perfectly reconstructed from a compressed file by utilizing some form of replacement.
  - *Examples:* Bitmap (.bmp), vector graphic (.svg), .png images, text file compression, database records.
  - *Run-Length Encoding (RLE):*
    - Definition: A form of lossless compression used for compressing text files and bitmap images.
    - Mechanism: Reduces file size by encoding sequences of adjacent, identical elements (characters in text files and pixels in bitmap images) into two values: run count and run value.
    - Example: RLE of a bitmap image.

- Lossy Compression
  - *Definition:* A type of compression that irreversibly eliminates unnecessary data.
  - *Effect:* File accuracy/quality is lower than with lossless compression, but file size is significantly reduced (often to about 10% of the lossless size).
  - *Examples:* Sound files (.mp3), .jpeg images.
  - *Mechanism in Sound Files:*
    - Perceptual Coding: Removes parts of the sound that are less audible or discernible to human hearing, as used in .mp3 compression.

2. Communication

2.1. Networks, including the Internet

LAN	WAN
A network that connects devices within a small geographical area, often within the same building.	A network that connects devices within a larger geographical area, such as a city, country, or globally.
Only private ownership.	Private or public ownership.
Transmission medium: Twisted Pair Cables , Coaxial Cables or Wi-Fi.	Transmission medium: PSTN or Satlink.
Higher data transfer rate.	Lower data transfer rate.
Less congestion.	Higher congestion.

- Networking devices: Interconnected devices that enable fast data transmission within a network.
- Networking benefits:
  - File sharing: Easily share data between different interconnected devices.
  - Resource sharing: Use network-connected output devices like printers or share software within the network.
  - Higher storage: Files can be stored in network-connected storage mediums.

- Client-Server Model
  - Server-based network: A dedicated server provides applications (administration of users, security, and resources) for the client computer to utilize.
  - Client-server Applications:
    - Printer: Manages print jobs from client computers.
    - File Sharing: Clients access software and user data files stored on the server.
    - Proxy server
    - Email server: For sending, receiving, and storing emails.
    - Database server: Manages DBMS.
    - Domain controller server:
      - Manages user accounts (IDs & passwords).
      - The client sends a login request to the server, which processes and grants the request if the user ID & password are recognized.

- Thin Clients vs. Thick Clients

Thin Clients	Thick Clients
A client that solely runs on the resources provided by the server and has no local storage.	An independent client that does not require the server to run.
Only provides input and receives output; processing is done by the server.	Thick client processes most of the application locally.
Smaller purchase cost: inexpensive, demanding hardware is not required.	Can function even if no server is connected (works offline).
Improved security: Cannot run unauthorized, harmful software.	No lag related to network problems.

- Peer-to-Peer network model (P2P)
  - Definition: A decentralized network where each connected computer stores data and operates independently as a 'peer', acting as both a client and a server.
  - Applications: Internet and Ad hoc networks.

- Client-Server vs. Peer-to-Peer models

Client-Server	Peer-to-Peer
Centralized backup.	Lesser initial setup cost.
Files & resources centralized in server: Prevents illegal resource usage.	Lesser network traffic: Each peer can simultaneously receive data from different sources.
Improved security: Files are stored on a central server, which would be regularly scanned for malware.	It can work even if a device goes down, but the client-server model can't work if the server goes down.

- Network Topologies
  - Bus
    - A single line (bus) connects all devices with terminators at each end.
    - Other computers can read data being sent between any two computers.
    - Unsuitable for heavy traffic due to frequent collisions.
  - Star
    - Consists of a central server (switch) with all other computers connected via dedicated connections.
    - The server can send packets to different devices simultaneously and bidirectionally.
    - No collisions are possible.
  - Mesh
    - Every device (node) is directly interconnected with each of the other devices (nodes).
    - Commonly used for wireless networks, such as the Internet, through the mesh connection of routers.
  - Hybrid
    - A combination of two or more topologies.
    - Example: A connection between two or more LANs of different topologies.

	Benefits	Drawbacks
Copper Cable	Less expensive and easier to install. Flexible. Easier to make terminations.	Doesn't perform well with small charges. Affected by electromagnetism.
Fiber-Optic Cables	Greater bandwidth, improved security, lightweight, easy to install, and less signal boosting are required; used in long-distance communications.	Needs expensive optical transmitters and receivers.

- Wired Networks
  - Use copper (twisted-pair cable or coaxial cable) or fiber-optic cables.
  - Cables are connected to an Ethernet port on the network router.

	Benefits	Drawbacks Low
Radio waves	Can travel over large distances with a wide range of wavelengths. Relatively inexpensive. Used for TV signals and mobile phone communications.	frequency means less data can be transmitted at one time. Affected by interference from radio stations with similar frequencies.
Microwaves	Larger bandwidth allows more data transfer.	Expensive to build emitting towers. Physical obstacles can interfere with signals.
Satellites	Cost-effective for long-distance communication, used in satellite phones and radio broadcasts.	Susceptible to interference. Expensive setup.

- Wireless Networks
  - Use radio waves (including WiFi), microwaves, and satellites to connect devices to networks without cables.

Real-time	On-demand
The event is captured live via a video camera that is connected to a computer.	Existing digital files are converted to encoded bit-streaming format for broadcasting on the internet by uploading to a dedicated server.
Video signal converted to an encoded streaming video signal.	A link for encoded video is placed on the website, and the user clicks on the link to view encoded streaming video.
Encoded video signal uploaded from computer to a dedicated streaming server via cables or high-speed wireless internet connection.	The data is streamed to a buffer in the user's computer, and the buffer stops the video from being paused as the bits are streamed.
The server then sends live images to all users requesting them as a real-time video. It cannot be paused, fast-forwarded, etc.	As the buffer is emptied, it's filled again, thus providing continuous viewing.
	Can be paused, fast-forwarded, etc.

- Ethernet
  - The most common wired medium for data transmission in LANs or WANs.
  - Typically used in bus topology; data collisions are managed by the CSMA/CD (Carrier Sense Multiple Access with Collision Detection) method.
  - CSMA/CD Process:
    - Device checks if the channel is busy before transmitting.
    - If busy, the device waits a random time before retrying.
    - During transmission, the device listens for other transmissions.
    - If a collision occurs, transmission is aborted, and both devices wait random times before retrying.

- Bit Streaming
  - Sequence of digital signals (bits) transferred over a communication path at high speeds, requiring a fast broadband connection and buffers.
  - Bit Streaming Types:
    - Real-time: Live events captured and transmitted directly.
    - On-demand: Pre-existing files are converted and streamed as requested.
  - Importance of High Broadband Speed/Bit-Rate
    - The user has to download and display bits at the same time.
    - Higher quality media requires faster speeds due to larger data frames as well.
    - Real-time streaming needs higher speeds due to simultaneous data requests coming from multiple different users.

- Cloud Computing
  - On-demand provision of computing services over the internet, including infrastructure, and platforms.
    - Infrastructure: Storage capacity and higher processing power.
    - Platform: Software, testing & debugging resources.

Public cloud vs. Private cloud

Public cloud	Private Cloud Owned and
Access provided by third-party service providers, shared among multiple users.	maintained by a single organization, providing exclusive access. Can be managed internally
Managed by cloud service providers using large server farms.	by the organization itself, or outsourced.

Benefits	Drawback
Less technical knowledge required, easy to implement.	Cannot access the resources/data stored on the cloud if there are bandwidth issues.
Flexibility to scale with organization's growth mindset.	Poor data privacy, since there may be data leakage in the multi-tenant architecture (public clouds).

- World Wide Web (WWW):
  - Description: Collection of web pages stored on websites.
  - Function: Protocols are used to transmit data across the ~~WWW~~.
- Internet (Interconnected Network):
  - Description: Massive, open network of networks.
  - Protocol: Uses TCP/IP protocol, which uses IP addresses to identify devices connected to the internet.
  - Access: Provided by Internet Service Provider.
  - Communication Methods: Wired, radio, and satellite.

- Router in a Network:
  - Function: Connects two networks together which operate under the same protocols (for example, IP).
  - Connections: Allows internal connections between LANs or external connection from the main LAN to a WAN.
  - Additional Roles: Acts as a gateway and firewall.
  - Setup: Usually attached to a server or switch in a LAN.
  - IP Address Translation: Translates private IP addresses to public IP addresses and vice versa.

- LAN-Supporting Hardware:
  - Switch:
    - Connected to all devices in a LAN.
    - Can simultaneously broadcast information to all devices.
  - Server:
    - Device/software that provides specific functions for computers in the network.
  - Network Interface Card (NIC):
    - Provides each device (end-system) in the wired LAN with a unique MAC address to uniquely identify it on the network.
    - Allows each device to connect to the network.
  - Wireless Network Interface Card (WNIC):
    - Provides each end-system of a wireless (WiFi) LAN a unique network address to identify it.
  - Wireless Access Points (WAP):
    - Allows devices to connect to the LAN via WiFi instead of using a cable.
    - Usually built into the router.
  - Cables:
    - A wired transmission medium that allows communication in wired networks.
  - Bridge:
    - Connects two LANs which work using the same protocol, which can be two segments of the same network.
    - Stores network addresses for all devices (end-systems) between the two networks.
    - Looks for the receiving device before it sends the message.
  - Repeater:
    - Connects two cables.
    - Regenerates the sent data signal over the same network before the signal weakens (attenuation) to prevent it from being corrupted.

- Internet-Supporting Hardware:
  - Modems:
    - Allows a device to connect to the Internet via a telephone line.
    - Function: A transmitter uses a modem to convert digital signals (from the transmitting device) to analogue signals sent down the telephone line. A receiver uses a modem on the other end to convert the analogue signals to digital signals so the receiving device can understand the data.
  - PSTN (Public Switched Telephone Network):
    - Refers to all telephone networks.
    - Channel: Used between two endpoints for the call duration via circuit switching.
    - Resilience: Lines are active even during a power outage.
    - Communication: Bi-directional.
  - Dedicated Lines:
    - Telecommunication path between endpoints.
    - Not shared with multiple users; it's bought/leased.
    - Function: Able to host websites as well as carry phone calls. Allows continuous, uninterrupted access to the Web.
  - Cell Phone Network:
    - Wireless networks spread over land areas divided into (hexagonal) cells.
    - Base Stations: Each cell is served by at least one base station (transceiver), which uses a different frequency range compared to adjacent cells to transmit data.
    - Capacity: Larger capacity is possible since the same frequencies can be used in non-adjacent cells.
    - Transmission: Radio waves are usually used for transmission. Can be broadcast in all directions over a wide area.
    - Portable Transceivers: Devices like mobile phones can communicate and access the internet via base stations.

- IPv4 vs. IPv6

IPv4	IPv6
32-bit address, split into 4 blocks by "."	The 128-bit address is divided into eight 16-bit blocks by ":"
Each block could have a value between 0 and 255 (00 to FF in hex).	Each block can have 4 hex values ranging from 0000 to FFFF.
For example, 255.0.1.255.	IPv6 can be shortened by removing at least (≥) 2 blocks containing only zeroes. For example: "2001:0db8:85a3:0000:0000:8a2e:0070:7334" can be shortened to "2001:0db8:85a3::8a2e:0070:7334"

- IPv4 Functionality:
  - IP Address Structure:
    - Network Identifier (netID): Identifies the network to which the host (device) is connected.
    - Host Identifier (hostID): Identifies the host within the network.
    - Classful Addressing: Used for IPv4, where different bit lengths for identification impose restrictions on available addresses.
  - Subnetting:
    - Definition: The practice of dividing a network into two or more networks.
    - Structure: IP addresses are broken down into three parts by not changing the netID but partitioning the host ID into a subnet ID and host ID.
      - Subnet ID: These bits are used to identify each subnet within the network.
      - Subnet Masks: Numbers that hide (mask) the netID of a system's IP address and leave only the host part as the machine identifier, allowing data to be routed within the subnet to the appropriate host.

- Public and Private IP Addresses:
  - Public IP:
    - Provided by the ISP.
    - Unique and can be accessed across the internet.
  - Private IP:
    - Issued by the LAN's router.
    - Unique within the LAN and can only be accessed within the LAN.
    - NAT (Network Address Translation): Required for private IP addresses to access the internet directly.
    - Security: Private IPs are more secure than public IPs since they are not directly accessible on the Internet and are hidden by NAT.
    - Address Range: The range of IP addresses used for private IP addressing can never be assigned to public IP addresses.

- Static vs. Dynamic IP addresses

Static	Dynamic
IP address never changes.	The IP address will change at regular periods.
Static IP addresses are valid when websites need to remember a device for a long time, e.g VPNs whitelisting.	Dynamic IP address is relatively more secure, hence used where data privacy is quite important.
Faster upload/download speeds.	Maintaining the cost of the dynamic IP address is a lesser.

- URL (Uniform Resource Locator)
  - Unique reference address for the exact location of an internet resource on the WWW
  - Protocol: Enables the browser to know what protocol is used to access information in the domain.
  - Host-name: The domain name.
  - Location of Server: The path indicating the server location.
  - Domain Name Service (DNS)
    - Definition: A naming system used for computers or resources having an internet connection.
    - Structure: Consists of a hierarchy of DNS servers which have a URL database and their corresponding IP addresses.

3. Hardware

3.1. Computers and Their Components

- A general-purpose computer system comprises a processor, memory, and I/O functionality.
- The following essential features are needed in a computer -
  - Input: Takes in data from the outside world.
  - Output: Displays data for human understanding.
  - Primary Storage: Main memory storing critical program instructions and data.
  - Secondary Storage: Non-volatile storage for noncritical data.
  - Removable secondary storage:
    - File backup and archives
    - Portable transfer of files to a second device
- Embedded systems:
  - Miniature computer systems such as microprocessors that are often a part of a more extensive system.
  - Each embedded system performs a few specific functions, unlike general-purpose computers.

Benefits	Drawbacks
Reliable since there are no moving parts	Difficult to program functions since there is no interface
Require less power	Expensive expert help is needed for the repair
Cheap to mass-produce	

Principle Operations of Hardware Devices

Laser printer:

- A laser beam and rotating mirrors are used to draw an image of the page on a photosensitive drum
  - The image is converted into an electric charge, which attracts charged toner such that it sticks to the image
- Electrostatic-charged paper rolled against the drum
- Charge pulls toner away from drum and onto paper

- Heat applied in the fuser to fuse toner to the paper
- The electrical charge was removed from the drum, and excess toner was collected

  please note: inkjet printers have been removed from the syllabus

3D Printer:

- The process starts with a saved digital file that holds the blueprint of the object to be printed
- The object is then built by sequentially adding layers of a material (e.g. polymer resin) until the object created
- The object is then cured (e.g. resin-made objects are hardened by UV light)

Microphone:

- Incoming sound waves enter the screen and cause vibrations in the diaphragm
- Vibrations cause the coil to move past a magnetic core
- Electrical current is generated, which is then digitized

Speaker:

- Takes electrical signals and translates them into physical vibrations to create sound waves
- The electric current in the voice coil generates an electromagnetic field
- Change in digital audio signal causes current direction to change, which changes field polarity
- Electromagnet is either attracted or repelled to a permanent magnet, causing a diaphragm that is attached to the coil to vibrate
- Vibration transmitted to air in front of the speaker
- The degree of vibration determines the amplitude and frequency of the sound wave produced

Magnetic Hard Disk:

- Hard disks have platters whose surfaces are covered with a magnetisable material.
- Platters are mounted on a central spindle and rotated at high speed
- The surface of platters is divided into concentric tracks & sectors, where data is encoded as magnetic patterns
- Each surface is accessed by read/write heads
- When writing, current variation in the head causes magnetic field variation on the disk
- When reading, magnetic field variation from the disk produces current variation in the read head

Solid State (Flash) Memory:

- Most use NAND-based flash memory
- Consist of a grid of columns & rows that has 2 transistors at each intersection
- Two transistors:
- Floating Gate: stores electrons, and the presence or absence of charge (electrons) represents either 1 or 0
- Control Gate: controls charge (electrons) flow for read/write

Optical Disc Reader/Writer:

- The disc surface has a reflective metal layer and is spun
- The tracking mechanism moves the laser assembly
- The lens focuses laser onto the disc A laser beam shone
- onto a disc to read/write Tracks have sequences of
- amorphous and crystalline states
  on the metallic layer
- When reading, the reflected light from the different states on the track is encoded as bit patterns
- When writing, the laser changes surface to crystalline and amorphous states along the track, corresponding to 1s or 0s.

Touchscreen:

- Considered as both an input & output device
- There are two main-types:

Resistive	Capacitive
Consists of two charged plates	Made from materials that store electric charge
Pressure causes plates to touch, completing the circuit	When touched, the charge is transferred to the finger
Point of contact registered with coordinates used to calculate the position	

- Virtual (Reality) Headset:
  - Virtual headsets consist of 2 lenses, (an LCD) display, a circuit board with sensors, a cover and foam padding
  - The display provides a simulation of a 3D environment generated by a 3D graphics package
  - The user can 'move' in the virtual environment by moving their head or using controllers
- Buffers:
  - A queue that temporarily stores data to balance input/output speed of data, while the cache is the short-term memory storage that stores frequently used data.
- Random Access Memory vs. Read-Only Memory

RAM	ROM
Volatile memory: loses content when power is turned off	Non-volatile memory: does not lose content when power is turned off
It can be read and altered	It can only be read
Used to store currently executing program	Used for storing OS kernel and boot-up instructions

- Types of RAM - Static RAM vs. Dynamic RAM

SRAM	DRAM
Doesn't need to refresh; hence, it uses less power and faster access time	Has to be refreshed; it has slower access times and needs higher power
More complex circuitry, hence more expensive	Only a single transistor & capacitor, hence less expensive to purchase
Each bit is stored in a flip-flop	Each bit is stored as a charge
Has lower data density	Has higher data density
Used in cache memory	Used in main memory

- Types of ROM – PROM vs. EPROM vs. EEPROM

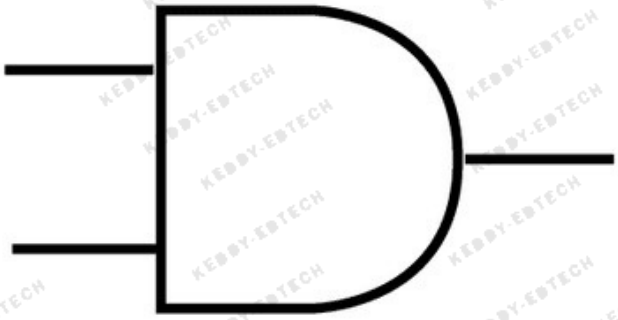
PROM	EPROM	EEPROM
Programmable ROM	Erasable Programmable ROM	Electrically Erasable Programmable ROM
It can be programmed only once after it is created	It can be erased by UV light exposure and can then be reprogrammed	It can be erased by an electrical signal and can then be reprogrammed
Data cannot be erased or deleted	Chip has to be removed for reprogramming	Can update data without removing the chip.

- Monitoring and Control Systems
  - Monitoring System:
    - Monitors some state external to the computer system
    - No changes were made to the environment by the system, and hence, no feedback
  - Control System:
    - Regulates the behaviour of other devices or systems
    - Event-driven system: the controller alters the system's state in response to some event
    - Time-driven system, where the controller takes action at a specific point in time
  - Hardware typically used in a system:
    - Sensor that measures an (analogue) property and transmits it to a processing unit, generally as an electrical or optical signal
    - Actuators that switch on/off heavy appliances (e.g. heater to heat/fan to cool)
    - ADC that converts analogue signals to digital signals
    - Transmission cable to transfer signals
  - Feedback Systems:
    - Output from the system affects the input of sensors
    - Ensures the system operates within the given criteria
    - Enabling the system output to affect subsequent system inputs may cause a change in the actions taken by the system
    - This enables the system to adjust conditions in a continuous process automatically

3.2. Logic Gates and Logic Circuits

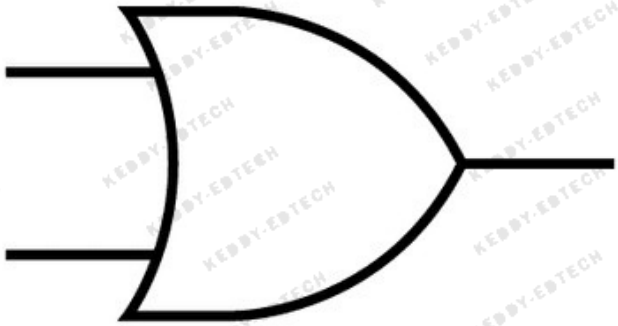
- Logic Gates: use one or more inputs and produce a single logical output
- AND gate: If both inputs are high, the output is high ($A \cdot B$)

A	B	Output
0	0	0
0	1	0
1	0	0
1	1	1



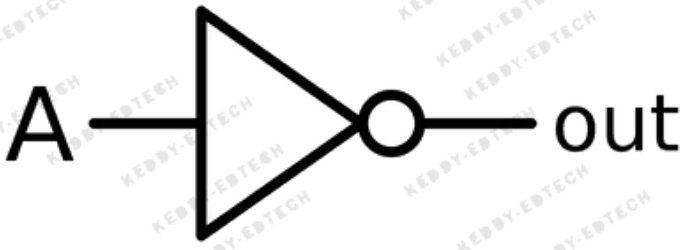
- OR gate: If either input is high, the output is high (*A+B*)

A	B	Output
0	0	0 1 1 1
0	1	
1	0	
1	1	



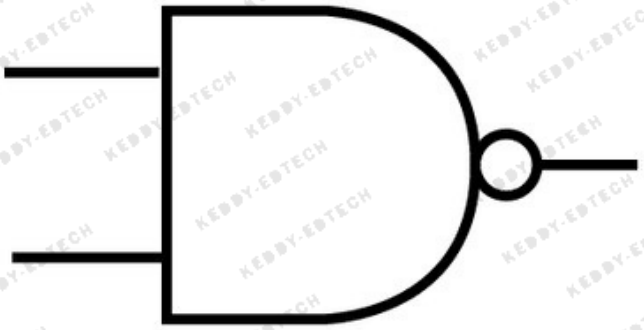
- NOT gate: an inverter (A)

A	Output
1	0 1
0	



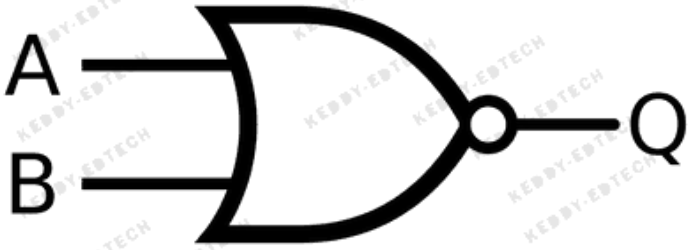
- NAND gate: (A•B)

A	B	Output
0	0	1 1 1 0
0	1	
1	0	
1	1	



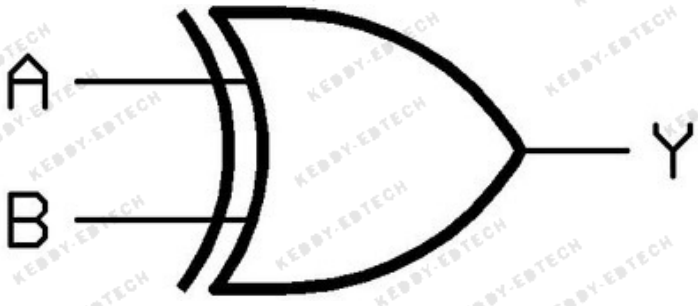
- NOR gate: (A+B)

A	B	Output
0	0	1 0 0 0
0	1	
1	0	
1	1	



- XOR gate: ($A \oplus B$)

A	B	Output
0	0	0 1 1 0
0	1	
1	0	
1	1	



4. Processor Fundamentals

4.1. Central Processing Unit (CPU) Architecture

Von Neumann model

- Von Neumann realized data & programs are indistinguishable and can, therefore, use the same memory.

- Von Neumann's architecture uses a single processor. It
- follows a linear sequence of fetch–decode–execute operations for the set of instructions, i.e. the program.
- To do this, the processor uses registers.

Registers

- Registers: smallest unit of storage of microprocessor; allows fast data transfer between other registers
- General Purpose registers
  - Used to temporarily store data values which have been read from memory or some processed result
  - Assembly language instructions can use it
- Special Purpose Registers
  - Some are accessible by assembly language instructions
  - Only holds either data or memory location, not both
  - Particular purpose registers include:
    - Program Counter (PC): holds the address of the next instruction to be fetched
    - Memory Data Register (MDR): holds data value fetched from memory
    - Memory Address Register (MAR): Holds the address of the memory cell of the program which is to be accessed
    - Accumulator (ACC): holds all values that are processed by arithmetic & logical operations.
    - Index Register (IX): Stores a number used to change an address value
    - Current Instruction Register (CIR): Once program instruction is fetched, it is stored in CIR and allows the processor to decode & execute it
    - Status Register: holds results of comparisons to decide later for action, intermediate and erroneous results of arithmetic performed

The Processor (CPU)

- Arithmetic and Logic Unit (ALU): part of the processor that processes instructions which require some form of arithmetic or logical operation
- Control Unit (CU): part of the CPU that fetches instructions from memory, decodes them & synchronizes operations before sending signals to the computer's memory, ALU and I/O devices to direct how to respond to instructions sent to the processor
- Immediate Access Store (IAS): memory unit that the processor can directly access
- System Clock: a timing device connected to a processor that synchronises all components.

Buses

- Set of parallel wires that allow the transfer of data between components in a computer system
  - Data bus: bidirectional bus that carries data instructions between processor, memory, and I/O devices.
  - Address bus: unidirectional bus that carries the address of the main memory location or input/output device about to be used, from processor to memory address register (MAR)
- Control bus
  - Bidirectional
  - used to transmit control signals from the control unit to ensure access/use of data & address buses by components of the system does not lead to conflict

Performance of Computer System Factors

- Clock Speed
  - Number of pulses the clock sends out in a given time interval, which determines the number of cycles (processes) the CPU executes in a given time interval
  - Usually measured in Gigahertz (GHz)
  - If the clock speed is increased, then the execution time for instructions decreases. Hence, more cycles per unit time, which increases performance.
  - However, there is a limit on clock speed since the heat generated by higher clock speeds cannot be removed fast enough, which leads to overheating.
- Bus Width
  - Determines the number of bits that can be simultaneously transferred
  - Refers to the number of lines in a bus
  - Increasing bus width increases the number of bits transferred simultaneously, increasing processing speed and performance.
- Cache Memory
  - Commonly used instructions are stored in the cache memory area of the CPU. If the cache memory size is
  - increased, more commonly executed instructions can be stored, and the need for the CPU to wait for instructions to be loaded reduces. Hence, the CPU executes more cycles per unit of time, thus improving performance.
- Number of Cores
  - Most CPU chips are multi-core — have more than one core (essentially a processor)
  - Each core simultaneously processes different instructions through multithreading, improving computer performance.

Ports

- Hardware which provides a physical interface between a device with CPU and a peripheral device
- Peripheral (I/O) devices cannot be directly connected to the CPU, hence connected through ports.

- Universal Serial Bus (USB): Can connect both input and output devices to the processor through a USB port.
- High Definition Multimedia Interface (HDMI)
  - Can only connect output devices (e.g. LCD) to the processor through a HDMI port
  - HDMI cables transmit high-bandwidth and high-resolution video & audio streams through HDMI ports
- Video Graphics Array (VGA)
  - Can only connect output devices (e.g. second monitor/display) to the processor through a VGA port
  - VGA ports allow only the transmission of video streams but not audio components

Fetch-Execute (F-E) cycle

- Fetch stage
  - PC holds the address of the next instruction to be fetched
  - The address on the PC is copied to MAR
  - PC is incremented
  - Instruction loaded to MDR from the address held in MAR
  - Instruction from MDR loaded to CIR
- Decode stage: The opcode and operand parts of instruction are identified
- Execute stage: Instructions executed by the control unit sending control signals
- Register Transfer Notation (RTN)
  - MAR ← [PC]
  - PC ← [PC] + 1
  - MDR ← [[MAR]]
  - CIR ← [MDR]
  - Decode
  - Execute
  - Return to start
  - *Square brackets: value currently in that register*
  - *Double square brackets:* CPU is getting value stored at the address in the register

Interrupts

- A signal from a program seeking the processor's attention
  - Handles the interrupt by controlling the processor
  - Different ISRs used for different sources of interrupt

A typical sequence of actions when an interrupt occurs:

The processor checks the interrupt register for interrupt at the end of the F-E cycle for the current instruction

If the interrupt flag is set in the interrupt register, the interrupt source is detected

If the interrupt is low priority, then an interrupt is disabled

If interrupting is a high priority:

All contents of registers of the running process are saved on the stack

PC is loaded with the ISR and is executed

Once ISR is completed, the processor pops the registers' contents from the stack, and the interrupted program continues its execution.

Interrupts re-enabled and

Return to the start of the cycle

4.2. Assembly Language

Assembly language: low-level programming language with instructions made up of an op code and an operand

Machine code: code written in binary that uses the processor's basic machine operations

Relationship between machine and assembly language: every assembly language instruction (source code) translates into exactly one machine code instruction (object code)

Symbolic addressing

Symbols used to represent operation codes

Labels can be used for addresses

Absolute addressing: a fixed address in memory

Assembler

Software that changes assembly language into machine code for the processor to understand

The assembler replaces all mnemonics and labels with their respective binary values (that are predefined before by the assembler software)

One pass assembler

Assembler converts mnemonic source code into machine code in one sweep of program

Cannot handle code that involves forward referencing

- Two pass assembler: software makes 2 passes through the code
  - On the first pass:
    - Symbol table created to enter symbolic addresses and labels into specific addresses
    - All errors are suppressed
  - On the second pass:
    - Jump instructions access memory addresses via table
    - Whole source code translates into machine code
  - Error reported if they exist
- Grouping the Processor's Instruction Set

Op Code	Operand	Explanation
Addressing		
LDM	#n	Immediate: Load n into ACC
LDD		Direct: load contents at address into the ACC
LDI		Indirect: load contents of address at given address into ACC
LDX		Indexed: load contents of given address + IR into ACC
Data Movement		
STO		Store contents of ACC into address
Arithmetic Operations		
ADD		Add contents of register to ACC
INC		Add 1 to contents of the register
Comparing		
CMP		Compare contents of ACC with that of given address
CMP	#n	Compare contents of ACC with n
Conditional Jumps		
JPE		Jump to address if compare TRUE
JPN		Jump to address if compare FALSE
Unconditional Jumps		
JMP		Jump to given address
I/O Data		
IN	Input any character and store ASCII value in ACC	
OUT	Output character whose ASCII value is stored in ACC	
Ending		
END	Return Control to operating system	

#denotes immediate addressing

B denotes a binary number, e.g. B01001010 & denotes a hexadecimal number, e.g. &4A

- Modes of Addressing
  - Direct Addressing: loads contents at address into ACC
  - Indirect Addressing: The address to be used is at given address. Load contents of this second address to ACC
  - Indexed addressing: form the address to be used as + *the contents of the IR (Index Register)*
  - Relative addressing: next instruction to be carried out is an offset number of locations away, relative to address of current instruction held in PC; allows for relocatable code
  - Conditional jump: has a condition that will be checked (like using an IF statements)
  - Unconditional jump: no condition to be followed, simply jump to the next instruction as specified

4.3. Bit Manipulation

- Binary numbers can be multiplied or divided by shifting
- Left shift (LSL #n)
  - Bits are shifted to the left to multiply
  - E.g. to multiply by four, all digits shift two places to left
- Right shift (LSR #n)
  - Bits are shifted to the right to divide
  - E.g. to divide by four, all digits shift two places to right
- Logical shift: zeros replace the vacated bit position
- Arithmetic shift: Used to carry out multiplication and division of signed integers represented by bits in the accumulator by ensuring that the sign-bit (usually the MSB) is the same after the shift.
- Cyclic shift: the bit that is removed from one end by the shift is added to the other end.

Bit Masking

- Each bit can represent an individual flag.
- ∴ by altering the bits, flags could be operated upon.
- Bit manipulation operations:
  - Masking: an operation that defines which bits you want to keep and which bits you want to clear.
  - Masking to 1: The OR operation is used with a 1.
  - Masking to 0: The AND operation is used with a 0.
  - Matching: an operation that allows the accumulator to compare the value it contains to the given value in order to change the state of the status register.
- Practical applications of Bit Masking:
  - Setting an individual bit position:
    - Mask the content of the register with a mask pattern which has 0 in the 'mask out' positions and 1 in the 'retain' positions.
    - Set the result with the match pattern by using the AND command with a direct address.
  - Testing one or more bits:
    - Mask the content of the register with a mask pattern which has 0 in the 'mask out' positions and 1 in the 'retain' positions.
    - Compare the result with the match pattern by using the CMP command or by "Checking the pattern".
  - Checking the pattern
    - Use AND operation to mask bits and obtain resultant.
    - Now subtract matching bit pattern from resultant.

    The final 'non-zero' result confirms the patterns are not the same else vice versa.

5. System Software

5.1. Operating System

- Need for OS
  - A set of programs designed to run in the background on a computer system which
    - Controls operation of computer system
    - Provides a user interface
    - Controls how computer responds to user's requests
    - Controls how hardware communicate
    - Provides an environment in which application software can be executed
  - OS hardware is unusable without an OS, as the OS acts as an interface since it controls communication between user and hardware

Key Management Tasks

- (Main) Memory Management
  - Memory protection to ensure 2 programs do not try to use same memory space
  - Paging
  - Use of virtual memory
- File Management
  - Provides file naming conventions
  - Maintains a directory structure
  - Allocates space to particular files

- Security Management
  - Proves usernames & passwords
  - Ensures data privacy
  - Prevents unauthorized access
  - Carries out automatic backup
- Hardware (input/output/peripherals) Management
  - Installation of appropriate driver software
  - Controls access to data sent to and from peripherals
  - Receives & handles interrupts from hardware devices
- Process Management
  - Enables multiprogramming and multitasking
  - Resolution of conflicts when 2 or more processes requires the same resource
  - E.g. via Round-robin method

Utility Software

- Disk Formatter
  - Prepares a hard disk to allow data to be stored on it
  - Deletes any existing data on disk
  - Performs formatting, process where computer 'draws lines' on disk surface to split it into small areas
- Virus checker
  - Checks for and then removes any viruses found
  - Constantly checks all incoming and outgoing files
- Defragmentation Software
  - Files can be big so have to be stored in multiple sectors, which can result in fragmentation (contents of file scattered across >2 non-contiguous sectors)
  - Fragmentation slows down disk access and thus the performance of the entire computer.
  - Defragmenting software works by physically reorganizing disk contents (files) such that they are stored in contiguous sectors.
  - This defragmentation reduces number of movements of the read/write heads require to access the disk contents, hence increasing computer performance
  - The defragmentation also creates larger contiguous free space regions
- Disk contents analysis/disk repair software
  - Software utility for visualization of disk space usage
  - Gets size for each folder and files, and generates a graphical chart showing disk usage distribution according to folders or other user defined criteria.
  - Allows disk to report errors (e.g. "bad sector")
  - Software will attempt to offer a solution
- File Compression
  - Reduces file size by removing redundant data in files
  - Causes improvements in the computer's performance by reducing the data that needs to be stored

Back-up Software

Makes copy of files on another storage medium in the event of a hard drive failure, user error, disaster or accident.

Should be a regular process

Can provide synchronization between devices

Program Libraries

Pre-written code that can be linked to a software under development without any amendments

Can perform common or complex tasks

Takes the form of classes

Benefits:

Saves time: less code needs to be written

Smaller testing time: pre-tested and used by others

Library file is a complex algorithm which the user does not need to understand to use it

Dynamic Link Library (DLL) files

Shared library file that contains code and data

Code saved separately from the main .EXE file, reducing the .EXE file's size

Code only loaded to main memory when required

DDL file can be made available to several applications simultaneously, thus reducing strain on memory

DLL files act as modules in more complex programs, making it easier to install and run updates

5.2. Language Translators

Assembler

Software that translates assembly language statements into machine code (binary) for execution

The mnemonics used translates into machine opcodes

Process simple because assembly language has a one-to-one relationship with machine code.

Compiler and Interpreter

Compiler	Interpreter
Translates a high-level language program to machine code.	Translates and executes a high-level language program, line-by-line.
Creates a .exe file which can be easily distributed.	No .exe file created.
Once compiled, .exe file does not need to be compiled again, resulting in faster execution.	Execution very slow – translated each time program run.
Reports all errors at the end of compilation: difficult to locate errors ∴ development process long.	Debugging easier/faster, since it stops translating when it reaches an error. This allows real time error correction.
Only be produced when all errors are fixed.	Can run program any time, even before code finished.
Used when development is completed.	Used during development.

- One Pass Assembler
  - A one-pass assembler is a type of assembler that processes the source code in a single scan. During this scan, it defines symbols and literals, storing them in the symbol and literal tables, respectively. It keeps track of the location counter to assign memory addresses to variables and instructions, processes pseudo-operations as they appear, and simultaneously translates the source code into machine code, effectively generating the object code while resolving addresses and symbolic references on the go.
- Two Pass Assembler
  - A two-pass assembler is a program that reads the source code twice to generate object code. In the first pass, it scans the code to define symbols, assign memory addresses, and gather information about literals and variables. In the second pass, it converts symbolic op-codes into their corresponding numeric op-codes, resolves symbol values, generates data for literals, and produces the final object code by fully translating the source code.
- Two-step translation
  - Java and some other high level language programs may require two-step translation, i.e., they will be partially compiled and partially interpreted
  - Java code first translated to bytecode by Java compiler
  - Bytecode finally interpreted by the Java Virtual Machine to produce machine code

- Integrated Development Environment (IDE) features
  - Coding
    - Context-sensitive prompts: Displays choice of keywords and available identifiers appropriate at current insertion point and provides choices in alphabetical order
    - Highlights undeclared/unassigned variable identifiers
  - Initial Error Detection
    - Dynamic syntax checks: Automatic checking and highlighting of syntax errors, as soon as line typed
    - Type checking & parameter checking
  - Presentation
    - Prettyprint: Automatic indentation and color-coding of keywords
    - Expand and Collapse code blocks: Saves excessive scrolling if collapsed, and easy to see global variable declarations and main program body when collapsed
  - Debugging
    - Single stepping: Executes program line-by-line to see the effect of each statement on variables
    - Breakpoints: Pauses program at a specific line to ensure program operates correctly up to that line
    - Variables/expressions Report Window: Monitors variables for comparing values.

6. Security, Privacy and Data Integrity

6.1. Data Security

- Data Security: ensuring data is protected against loss and unauthorized access.
- Data Integrity: making sure that data is valid and does not corrupt after transmission
- Data Privacy: ability to determine what data is shared with a third party
- Data Security and Computer System Security

Data Security	System Security
Protection of data on a computer system	Protection of the computer system
To prevent corruption of data and prevent hackers from using data	To prevent access of viruses to the system and prevent hackers from entering your computer system
E.g. encryption	E.g. ID & Password

*Threats to Computer & Data Security**

- Malware
  - software intentionally designed to damage a computer or computer network
  - Includes Virus & Spyware
  - Virus: Software that replicates itself by inserting a copy of itself into another piece of software, which may cause the computer to crash and can lead to deletion or corruption of data
  - Spyware: software that gathers information about users' online and offline activity, including accessed sites, applications, and downloaded files.
  - Risk restriction: Ensure anti-virus and anti-spyware software is installed, regularly updated and run.
- Hacking
  - illegal access to a computer system
  - Hackers can obtain user's confidential data which can cause identity theft
  - Can lead to the deletion or corruption of data
  - Risk restriction: Use strong passwords and ensure firewall
- Phishing
  - Attempt through emails to obtain user's confidential data which can cause identity theft
  - Risk restriction: Ignore suspicious mails and ensure firewall criteria include SPAM filters, blacklist, etc.
- Pharming
  - Redirects user to a fake website that appears legitimate to gain confidential data
  - Risk restriction: use a reliable ISP; check that links are genuine and ensure https is present in the URL

Computer System Security Measures

- User Accounts and Passwords
  - Usernames & passwords to deny access to unauthorized users
  - User-assigned privilege, which accesses to only the user's workplace, preventing the user to have admin rights.
  - Can assign privileges to files so users with low privileges do not have access.
- Firewalls
  - Hardware or software that filters information travelling between the computer system and the internet
  - (software) firewall can make decisions about what to allow and block by detecting illegal attempts by specific software to connect to the internet
- Authentication
  - Process of determining whether someone is who they claim to be.
  - Helps prevent unauthorized access
  - Log-on using digital signatures, passwords and biometric scans.
- Digital Signatures
  - The data is hashed using a agreed hashing algorithm, to produce a digest. The sender encrypts the digest using the sender's private key to create the digital signature. The message and the signature are sent to the receiver. The receiver uses the sender's public key to decrypt the signature to reproduce the orignal digest. The receiver then applies the same hashing algorithm on the received message to produce a second digest. The receiver compares the second digest with the one from the digital signature. If both are the same, the document is authentic and this proves the integrity of the message and the identity of the sender.
- Anti-virus software
  - Runs in the background to detect & remove viruses.
  - Checks files for known malicious patterns
- Anti-spyware software: detects & removes spyware.
- Encryption:
  - Conversion of data to code by encoding it
  - It doesn't stop illegal access but appears meaningless
  - Necessary to use decryption software to decode data
- Data Security Measures
  - Encryption
  - Access Rights to data (authorization): different users assigned different authorization levels which prevent them from accessing all data ∴ increases security
- Data Backup
  - An exact copy of an original piece of data in case the original is lost or corrupted
  - Within the same computer system or at different site
- Disk-mirroring strategy
  - Real-time strategy that writes data to two or more disks at the same time.
  - If one fails, the other is still there to be read off of

6.2. Data Integrity

- Data validation and data verification help protect the integrity of data by checking whether the data entered is sensible and accurate, respectively.
- Data Validation: checks if data entered is valid, but not its accuracy

- Data Validation Methods
  - Range check: data must be between a set of values
  - Format check: data must follow correct pattern/order
  - Length check: data must have exact no. of characters
  - Presence check: checks if some data has been entered
  - Existence check: data entered must exist
  - Limit check: checks whether a value entered is within acceptable minimum and maximum values.
  - Check digit: A digit is used as the answer to an arithmetic operation of other digits in data. If not matched, then data entered incorrectly
- Data Verification: checks data entered is accurate during data entry and data transfer
- Data Entry Verification Methods
  - Visual Check: Person manually compares original data with that entered to check if correct
  - Double Entry: Enter data into computer twice and compares.
  - If differences found, go back to raw data to fix error

Data Transfer Verification Methods

- Errors may occur when data moved in system.
- Parity Check
  - All data transmitted as bits
  - Number of 1s in a byte must always be either an odd number or an even number
  - Parity can be set either as even or odd
  - E.g. two communicating devices decide there will always be an odd number of 1s. A byte is received that has even number of 1s so error occurred and receiving device would ask for it to be sent again
  - Used also when data sent between parts of the CPU
  - Not foolproof: if 2 bits are transposed, data accepted
- Checksum Check
  - Data sent from one place to another as block of bytes rather than individual bytes
  - Computer adds together all bytes being sent
  - Any bits lost at most-significant end as carry ignored so answer is an 8-bit number
  - Checksum calculated before and after data sent
  - If two bytes different, error occurred therefore block of bytes must be sent again

7. Ethics and Ownership

- Ethics: a system of moral principles that guide behaviour based on philosophical views
- Computer Ethics
  - Regulate how computing professionals should make decisions regarding professional & social conduct.
  - A computing professional can be ethically guided by joining a professional, ethical body such as the BCS and IEEE, which have codes of conduct.
  - The BCS Code of Conduct provides the following guidance:
    - Public Interest: Both codes emphasize the importance of acting in the public interest and safeguarding the health, safety, and welfare of the public.
    - Judgement and Integrity: Professionals are expected to exercise sound judgement and maintain their integrity in their work.
    - Ethical Practice: Both codes stress the importance of ethical behaviour and decision-making.
    - Lifelong Learning: The IEEE code explicitly mentions the need for continuous learning, which aligns with the need to maintain professional competence.
  - The IEEE Software Engineering Code of Ethics defines eight principles:
    - PUBLIC: Software engineers shall act consistently with the public interest.
    - CLIENT AND EMPLOYER: Software engineers shall act in the best interests of their client and employer, while also considering the public interest.
    - PRODUCT: Software engineers shall ensure that their products meet the highest professional standards possible.
    - JUDGEMENT: Software engineers shall maintain integrity and independence in their professional judgement.
    - MANAGEMENT: Software engineering managers and leaders shall promote an ethical approach to the management of software development and maintenance.
    - PROFESSION: Software engineers shall advance the integrity and reputation of the profession, consistent with the public interest.
    - COLLEAGUES: Software engineers shall be fair to and supportive of their colleagues.
    - SELF: Software engineers shall participate in lifelong learning and promote an ethical approach to their practice.

- Ownership
  - Data ownership: having legal rights and complete control over a single piece or set of data elements.
  - Copyright gives the creators of some types of media rights to control how they're used and distributed.
  - Competitors can steal programming ideas and methods; software can easily be copied and bootlegged (sold illegally); hence, legislation is needed to protect the ownership, usage, and copyright of data.

Software Licencing

- Free Software Foundation:
  - A license gives users freedom to run, copy, distribute, study, change, and improve the software.
  - Condition: any redistributed version of software must be distributed with original terms of free use, modification, and distribution (aka copyleft)
- The Open Source Initiative:
  - The source code of open-source software is readily available to users under copyright; it does enable users to re-distribute the software.
  - The concept of an open-source program relies on the fact that users can review source code to eliminate bugs in it
- Shareware:
  - Demonstration software that is distributed for free but for a specific evaluation period only
  - Distributed on a trial basis and with an understanding that sometime later, a user may be interested in paying for it
  - Used for marketing purposes
- Commercial: Requires payment before it can be used, but includes all program's features, with no restrictions

Artifical intelligence (AI)

- Artificial Intelligence (AI): the ability of a computer to perform tasks in such a way that is conventionally associated with human intelligence:
  - AI can learn from past mistakes
    - they adapt to stop the same problem from occurring again
    - they learn to predict what might happen & raise alert
  - AI can learn to work more efficiently
    - when an action slows the system down, it can prevent this from happening again
    - when an action increases the speed of the system, it can repeat this when necessary to improve efficiency
- AI Applications
  - Developing autonomous mechanical products
  - Machine learning through data sets
- AI Impacts
  - Social
    - Replacement of manual labour with automation could lead to massive unemployment.
    - However, it could lead to increased leisure time.
  - Economic: Due to increased innovation and efficiency with automation provided by AI, there'd be lower manufacturing costs in general
  - Environmental: Detrimental impact on the environment due to robot manufacture with limited resources and its waste disposal

8. Database and Data Modelling

8.1. File Based System

- Data stored in discrete files, stored on computer, and can be accessed, altered or removed by the user

Disadvantages of File Based System:

- No enforcing control on organization/structure of files
- Data repeated in different files; manually change each
- Sorting must be done manually or must write a program
- Data may be in different format; difficult to find and use
- Impossible for it to be multi-user; chaotic
- Security not sophisticated; users can access everything

8.2. Database Management Systems (DBMS)

- Database: collection of non-redundant interrelated data
- DBMS: Software programs that allow databases to be defined, constructed and manipulated

Features of a DBMS:

- Data management: data stored in relational databases - tables stored in secondary storage
- Data dictionary contains:
  - List of all files in database
  - No. of records in each file
  - Names & types of each field
- Data modeling: analysis of data objects used in database, identifying relationships among them
- Logical schema: overall view of entire database, includes: entities, attributes and relationships
- Data integrity: entire block copied to user's area when being changed, saved back when done
- Data security: handles password allocation and verification, backups database automatically, controls what certain user's view by access rights of individuals or groups of users

Data change clash solutions:

- Open entire database in exclusive mode – impractical with several users
- Lock all records in the table being modified – one user changing a table, others can only read table
- Lock record currently being edited – as someone changes something, others can only read record
- User specifies no locks – software warns user of simultaneous change, resolve manually
- Deadlock: 2 locks at the same time, DBMS must recognize, 1 user must abort task

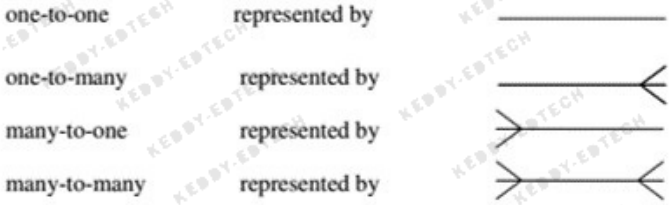
Tools in a DBMS:

- Developer interface: allows creating and manipulating database in SQL rather than graphically
- Query processor: handles high-level queries. It parses, validates, optimizes, and compiles or interprets a query which results in the query plan.

8.3. Relational Database Modelling

- Entity: object/event which can be distinctly identified
- Table: contains a group of related entities in rows and columns called an entity set
- Tuple: a row or a record in a relational database
- Attribute: a field or column in a relational database
- Primary key: attribute or combination of them that uniquely define each tuple in relation
- Candidate key: attribute that can potentially be a primary key
- Foreign key: attribute or combination of them that relates 2 different tables
- Referential integrity: prevents users or applications from entering inconsistent data
- Secondary key: candidate keys not chosen as the primary key
- Indexing: creating a secondary key on an attribute to provide fast access when searching on that attribute; indexing data must be updated when table data changes

8.4. Relational Design of a System



8.5. Normalization

1st Normal Form (1NF): contains no repeating attribute or groups of attributes. Intersection of each tuple and attribute contains only 1 value. Example:

DELNOTE

Num	CustName	City	Country	ProdID	Description
005	Bill Jones	London	England	1	Table
005	Bill Jones	London	England	2	Desk
005	Bill Jones	London	England	3	Chair
008	Mary Hill	Paris	France	2	Desk
008	Mary Hill	Paris	France	7	Cupboard
014	Anne Smith	New York	USA	5	Cabinet
002	Tom Allen	London	England	7	Cupboard
002	Tom Allen	London	England	1	Table
002	Tom Allen	London	England	2	Desk

2nd Normal Form (2NF): it is in 1NF and every non-primary key attribute is fully dependent on the primary; all the incomplete dependencies have been removed. Example:

DELNOTE

Num	CustName	City	Country
005	Bill Jones	London	England
008	Mary Hill	Paris	France
014	Anne Smith	New York	USA
002	Tom Allen	London	England

DEL_PROD

Num	ProdID
005	1
005	2
005	3
008	2
008	7
014	5
002	7
002	1
002	2

PRODUCT

ProdID	Description
1	Table
2	Desk
3	Chair
7	Cupboard
5	Cabinet

3rd Normal Form (3NF): it is in 1NF and 2NF and all non-key elements are fully dependent on the primary key. No inter-dependencies between attributes.

- MANY-TO-MANY functions cannot be directly normalized to 3NF, must use a 2 step process e.g.

becomes:

8.6. Data Definition Language (DDL)

- Creation/modification of the database structure using this language
  - written in SQL
- Creating a database:

```
CREATE DATABASE <database-name>
```

- Creating a table:

```
CREATE TABLE <table-name> (…)
```

- Changing a table:

```
ALTER TABLE <table-name>
```

- Adding a primary key:

```
PRIMARY KEY (field)
ADD <field-name>:<data-type>
```

- Adding a foreign key:

```
FOREIGN KEY (field) REFERENCES <table>(field)
```

- Example:

```
CREATE DATABASE 'Personnel.gdb'
CREATE TABLE Training
(EmpID INT NOT NULL,
CourseTitle VARCHAR(30) NOT NULL,
CourseDate Date NOT NULL,
PRIMARY KEY (EmpID, CourseDate),
FOREIGN KEY (EmpID) REFERENCES Employee(EmpID))
```

8.7. Data Manipulation Language (DML)

- Query and maintenance of data done using this language – written in SQL

Queries:

- Creating a query:

```
SELECT <field-name>
FROM <table-name>
WHERE <search-condition>
```

- SQL Operators:

=	Equals to
>	Greater than
<	Less than
>=	Greater than or equal to
<=	Less than or equal to
<>	Not equal to
IS NULL	Check for null values

- Sort into ascending order:

```
ORDER BY <field-name>
```

- Arrange identical data into groups:

```
GROUP BY <field-name>
```

- Joining together fields of different tables:

```
INNER JOIN
```

Data Maintenance:

- Adding data to table:

```
INSERT INTO <table-name>(field1, field2, field3)
VALUES (value1, value2, value3)
```

- Deleting a record:

```
DELETE FROM <table-name>
WHERE <condition>
```

- Updating a field in a table:

```
UPDATE <table-name>
SET <field-name> = <value>
WHERE <condition>
```

Bridging Dreams with Knowledge

Need more study material? Reach out anytime!