# Preface

Underlined words have definitions in Appendix A: Definitions

**Bolded** words r just cos i wanna emphasise

# 1 Computer Architecture

In a computer, data is stored as **bits**. 8 bits = 1 byte, 4 bits = 1 nibble[1]

Kilobyte (kB) = $10^3$ bytes, Kibibyte (KiB) = $2^{10}$ bytes

Megabyte (MB) = $10^6$ bytes, Mebibyte (MiB) = $2^{20}$ bytes

So on and so forth. You need to memorise until Exabyte / Exibyte

Kilo → Mega → Giga → Tera → Peta → Exa

Kibi → Mebi → Gibi → Tebi → Pebi → Exi

Parts of a computer:

| | |
|---|---|
| **Processor** | It executes instructions. The frequency (Hz) is how many instructions it can carry out in 1 second. **Multi-core** processors have > 1 processing unit and can perform more than one instruction at the same time |
| **Memory** | Whenever a computer needs to store data it goes here. Generally, this refers to Random Access Memory (RAM). It's easily changed and **volatile**[2]. It stores data by assigning it an address. |
| **Secondary Storage** | It's long term storage for large amounts of data and is **non-volatile** and cheaper than RAM but it's much slower.<br>**Magnetic**: More vulnerable to damage from magnetic fields, heat, impact, and natural deterioration over time. It's heavier and bulkier than optical and solid-state media, and stores up to TBs of data. It's cheap per GB, and is slower than solid state media. E.g. Hard disk<br>**Optical**: More vulnerable to damage from scratches and natural deterioration over time, but more resistant to heat and impact. It's portable as it's small and light, but only stores up to GBs of data. It's cheaper than SSDs per GB, but more ex than magnetic media, and slower than solid state media e.g. digital versatile disk (DVD)<br>**Solid-state**: Most vulnerable, and most resistant to impact and temperature changes. It's portable and lightweight due to it being small and light, and can store up to TBs of data. It's the most expensive per GB, but faster than all other media. E.g. Solid-state drive (SSD) |
| **Data bus** | Bi-directional, transports data between CPU and memory |
| **Address bus** | Uni-directional, specifies memory address info. So when the processor wants to access data on say address 2, the processor |

[1] Don't write this tho, its not in tb and is very informal

[2] Data is lost if power supply is interrupted

| | |
|---|---|
| | will send 2 on the address bus, the data bus will return the data on address 2 in the memory to the processor. If the processor wants to override / update data at address 2, it will send data down the data bus and 2 on the address bus. The memory then updates address 2 with the data from the data bus |
| **Input devices** | Basically like things that provide the processor with data e.g. a microphone, a mouse, .etc |
| **Output devices** | Basically things that the processor outputs data to e.g. a screen, speakers, .etc |

Connection interfaces:

| Interface | Use | Types | Speed |
|---|---|---|---|
| Universal Serial Bus (USB) | Powering and/or communicating with external devices | Type A, B, C<br>Micro and Mini | 2.0: 480 Mbit/s<br>3.2: 20 Gbit/s<br>4: 80 Gbit/s |
| High-Definition Multimedia Interface (HDMI) | For delivering audio/video data to compatible devices (monitors, TVs) | HDMI standard<br>HDMI Mini<br>HDMI Micro | 1.3-1.4b: 10.2 Gbit/s<br>2.0-2.0b: 18 Gbit/s<br>2.1: 48 Gbit/s |
| Peripheral Component Interconnect Express (PCIe) | For communication between internal expansion cards | PCI-e x1<br>PCI-e x4<br>PCI-e x8<br>PCI-e x16 | Increases with the number of lanes, up to x16. Different lanes have diff speeds based on the version of PCIe<br>5.0: 4 GB/s<br>6.0: 8 GB/s<br>7.0: 17 GB/s |

PCIe connectors are **within** the motherboard.

# 2 Data representation

We like to use denary in our day to day lives. Computers like **binary**, though.

## 2.1 Bin n Dec

Denary to 1-byte binary: Example 169

Sum of place values:

Draw a table with 8 (1-byte) columns and 2 rows, one column for each digit. In the top row write $2^n$, with n = 0 at the right-hand side and increasing by one for every place it moves to the left

Then fill it in from the left

$169 - 2^7 = 41$, so there's $2^7$. So put 1 there.

$41 - 2^6 = -23$, so there isn't $2^6$. Put 0 there.

$41 - 2^5 = 9$, so there is a $2^5$. Put 1 there. Rinse, repeat, and you get:

| $2^7$ | $2^6$ | $2^5$ | $2^4$ | $2^3$ | $2^2$ | $2^1$ | $2^0$ |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |

So, $169_{10} = 1010\ 1001_2$

So you just do the reverse for binary to denary.

To find the denary value of $1010\ 1001_2$,

$2^7 + 2^5 + 2^3 + 2^0 = 169$, so $1010\ 1001_2 = 169_{10}$

## 2.2 Hex n Bin

Another one we like to use in computers is **hexadecimal**, since two letters can represent 1 byte.

To convert Bin to Hex, you map each hexadecimal digit to its 4 bit binary equivalent:

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

So $1010\ 1001_2 = A9_{16}$, $A9_{16} = 1010\ 1001_2$

So in binary u just add leading zeros (10 → 0010), split into 4 bits and use the table

## 2.3 Hex and Dec

From Hex to Dec: It's pretty easy. Same thing, using sum-of-place values: $A9_{16}$, you split it to get $A_{16}$ and $9_{16}$. Convert it into denary to get $10_{10}$ and $9_{10}$, then multiply from the right by $16^n$ starting with n = 0, so $10 \times 16^1 + 9 \times 16^0 = 169$. So $A9_{16} = 169_{10}$.

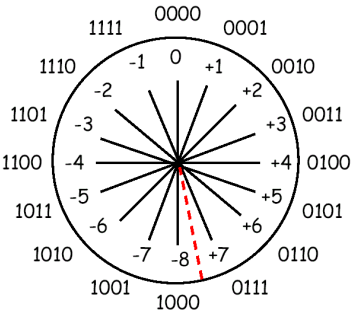
From Den to Hex, **this is where shit hits the fan.**

The easiest version is **division by 16**. So taking $169_{10}$, you just repeatedly divide by 16, keeping the remainder until your quotient is 0. Turn all your remainders into hex, and read from top to bottom. E.g.:

| Denary | Quotient | Remainder |
|---|---|---|
| 169 | 9 | $10_{10} = A_{16}$ |
| 9 | 0 | $9_{10} = 9_{16}$ |

Therefor $169_{10} = A9_{16}$

## 2.4 Two's complement




This is basically how you represent **negative numbers**.

The idea is like a cycle: If the first bit is 0, you're dealing with a positive number. Otherwise, you're dealing with a negative number.

So in 4-bit two's complement, 0001 is +1, 1111 is -1. 0010 is +2, 1110 is -2.

To calculate it, let the number be x. Then what you do is you take $2^n - x$, where n is the number of bits.

So if $x = 3_{10} = 011_2$ (3 bits), then $-3_{10} = 2^3{}_{10} - 011_2 = 8_{10} - 3_{10} = 5_{10} = 101_2$. Hence in **3-bit two's complement, $101_2 = -3$**.

The easier way is to just flip all the bits and add 1 (only do this for calculations, **never** for explanation on like how do you do this)

So like for $011_2$, flip the bits to get $100_2$, then add 1 to get $101_2$.

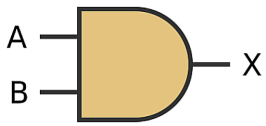
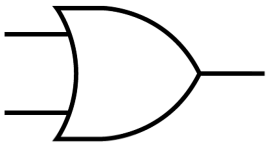
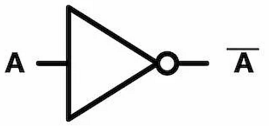
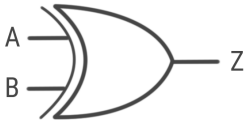
Or for like 0000 $0101_2$ (5), to get -5, flip the bits: 1111 $1010_2$, add 1: 1111 $1011_2$. So -5 in 8-bit two's complement is 1111 $1011_2$

## 2.5 Text

In text data, every symbol, letter, digit or space is called one character. Character encoding assigns each character a number, which allows you to transmit text in bits and bytes. The syllabus one is the (extended) **American Standard Code for Information Interchange (ASCII)**, which uses **1** byte for every character. So it can store characters from 0000 $0000_2$ to 1111 $1111_2$, which is $00_{16}$ to $FF_{16}$, or $0_{10}$ to $255_{10}$. Extended ascii usually uses hexadecimal.

# 3 Logic gates

Basically **boolean** math

| Operator | Drawing | Symbol | Output |
|---|---|---|---|
| AND |  | $A \cdot B$ | 1 if A and B is 1, 0 otherwise |
| OR |  | $A + B$ | 1 if A and/or B is 1, 0 otherwise |
| NOT |  | $\bar{A}$ | 1 is A is 0, 0 otherwise |
| XOR |  | $A \bigoplus B$ | 1 if A or B but not both is 1, 0 otherwise |

Memorise this ^

For stuff like NAND, NOR, XNOR, just put a tiny circle at the end of the operator.
Order of operations (1st to last): Parentheses, NOT, AND, OR

Truth tables: They show every possible combination of outputs

So for the equation: $X = (\overline{A + B} \cdot B) \bigoplus \overline{A \cdot B}$

| | | (Working space) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| A | B | P = A + B | Q = $\bar{P}$ | R = Q x B | S = A x B | T = $\bar{S}$ | Y = $R \bigoplus T$ | X |
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |

From this table, the equation is: NOT A AND NOT B OR NOT A AND B OR A AND NOT B

These are the outputs you get 1. (A/B = 1, NOT A / NOT B = 0)

(Technically simplified it's NOT (A AND B) but you don't need to know that)

## 3.1 Simplifying and properties, and presentation

| Property | AND | OR | XOR |
|---|---|---|---|
| Annulment | $A \cdot 0 = 0$ | $A + 1 = 1$ | - |
| Identity | $A \cdot 1 = A$ | $A + 0 = A$ | $A \bigoplus 0 = A$ |
| Idempotent | $A \cdot A = A$ | $A + A = A$ | - |
| Self-inverting | - | - | $A \bigoplus A = 0$ |
| Complement | $A \cdot \bar{A} = 0$ | $A + \bar{A} = A$ | $A \bigoplus \bar{A} = 1$ |
| Commutative | $A \cdot B = B \cdot A$ | $A + B = B + A$ | $A \bigoplus B = B \bigoplus A$ |
| Associative | $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ | $A + (B + C) = (A + B) + C$ | $A \bigoplus (B \bigoplus C) = (A \bigoplus B) \bigoplus C$ |
| Distributive | $A \cdot (B + C) = (A \cdot B) + (A \cdot C)$<br>$A \cdot (B \bigoplus C) = (A \cdot B) \bigoplus (A \cdot C)$ | $A + (B \cdot C) = (A + B) \cdot (A + C)$ | - |
| Absorption | $A \cdot (A + B) = A$ | $A + (A \cdot B) = A$ | - |
| De Morgan's | $\overline{A \cdot B} = \bar{A} + \bar{B}$ | $\overline{A + B} = \bar{A} \cdot \bar{B}$ | - |

They'll give you this list in the paper, **except for Associative, Distributive, De Morgan's and double negation**. (Double negation: $\bar{\bar{A}} = A$)

How do you do proving? It's like A math proving identities AKA LHS = ⋯ = RHS

So like prove A x (A+B) = A (Without using absorption)

LHS

= A x (A+B)

= (A+0) x (A+B) (Identity property of OR)

= A + (0 x B) (Distributive property of OR)

= A + 0 (Annulment property of AND)

= A (Identity property of OR)

= RHS

LHS = A x (A+B)

= A (Absorption property of OR)

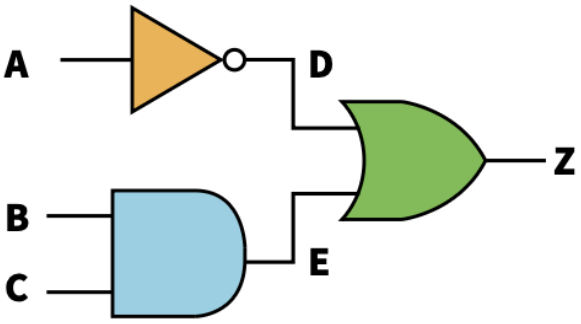
= RHS

## 3.2 Logic circuit diagrams

How to draw 101

Just right, think a bit, find all your intermediate steps n shit

So, for $Q = \bar{A} + B \cdot C$




You need NOT A, and B AND C, so you need to put A into a NOT gate followed by an OR gate, and that OR gate should be connected to B AND C which is B and C connected into an AND gate. Or:

To draw lines that overlap but don't intersect, just draw the line. If they do, put a dot there.

To derive a logic statement from a logic circuit diagram, just work backwards.

Z = D OR E

D = NOT A

E = B AND C

Therefore Z = NOT A OR B AND C

Yeah that's basically it

## 4-8 aka practical

Algorithm, Program, Source code, Input (algorithms), Output (algorithms), solution, Command line interface (CLI), Graphical user interface, Syntax highlighting, Kernel (Jupyter), Notebook (Jupyter), Comment, Variable, Expression, Nested function call, Return value, Method, Object, Binary operator, Operand, Operator, Unary operator, Floating-point number (float), integer (int), Type casting, module, string (str), escape code, replacement field, immutable, substring, dictionary (dict), key, control flow, flow line, flowchart, symbol (flowchart), construct, sequence construct, structured programming language, branch, iterate, iteration construct, selection construct, loop, parameters, signature (programming), local va

riable, global variable, computational thinking, decomposition, generalisation, delimiter, bug, debugging, test case, testing, dry run, trace table, backtracking,

In theory paper when they ask right "Specify an input and an output and requirements to be valid" JUST SAY THE INPUT LAH

Like theres a program that takes in two numbers and adds them together, an input: The first number, x. Requirement: X should be numerical and not alphabetical

**Validation**: Types of checks:

**Presence**: Checks if input exists (whether it's whitespace)

**Format**: Checks if input is in the correct format (5-character string, .etc

**Range**: Checks if input is within the specified range (1-10, 1000-9999, .etc)

**Existence**: Checks if input lies in a database

**Length**: Checks if input is the correct length (5 characters long)

**Algo design**:
Before you start coding, you should **understand the problem**.

**Decomposition**: breaking down the problem into smaller parts

**Modular decomposition**: breaking it down into like modules, e.g. in the idea of getting all grades from a C5 to an A1, that wld be like term 1 get Math and AMath to A1, term 2 get bio and physics to A1, term 3 get humans to A1, .etc

**Incremental decomposition**: doing it step-by-step. In the same example, that wld be like term 1 all B4, term 2 all A2, term 3 all A1, .etc

Then you **generalise** it so that it' s applicable to more cases

**Testing and debugging**:

We discriminate against **bugs** and love **debugging**.

Types of errors:

**Syntax**, **Logic** and **Run-time**.

Syntax errors are caused by incorrect source code that violate the rules of the language, e.g. forgetting a colon, indentation

Logic errors cause the program to not behave as expected, e.g. multiplying instead of adding

Runtime errors are errors that occur while a program is running, e.g., dividing by zero, dividing a string

# 9 Excel

# 10 Networking

## 10.1 Types of network

### 10.1.1 Wired vs wireless

There are two classifications for networks, based on transmission medium:

**Wired**: A network of devices connected by physical cables, typically using **ethernet**

**Wireless**: A network of devices in which signals are transmitted without the use of physical cables

| Factor | Wired | Wireless |
|---|---|---|
| Cost | Initially cheaper but becomes more expensive as the network grows due to the cost of cables | Initially expensive due to the cost of wireless networking hardware but becomes more cost-effective as the network grows |
| Speed of transmission and bandwidth | Faster and higher bandwidth as cables provide dedicated connection | Generally slower and lower bandwidth due to possible interference from radio waves or microwaves; varies according to user location in relation to network |
| Reliability | More reliable as data transmission is less affected by interference | Less reliable due to potential interference from radio waves and microwaves or blockage from physical obstructions |
| Security | More secure as the network is less susceptible to interception and hacking from a distance | Less secure due to possible intrusions by hackers who do not need physical access to a cable or network hardware |
| Mobility of users | Lower as network connections are fixed at specific spots and users cannot move to other locations | Higher as users can move about freely within the range of the wireless network |
| Scalability | More cumbersome to add new devices to the network due to the need for additional cables | Easier to add new devices to the network as no additional hardware is usually needed; however there is a limit as interference increases with the number of wireless devices |

### 10.1.2 LAN vs WAN

**LAN** is much much smaller, usually only like a home, school or an office.

**WAN** is much much bigger, like over multiple geographical locations, generally consist of multiple smaller networks like **LAN**s. There can be public or private **WAN**s, private ones can be built with encryption, allowing for confidentiality

## 10.1.3 Client-Server vs Peer-to-Peer

Client-server: One or more **higher-capacity** devices will act as servers, everyone else are clients. Every server contains data and other resources to be shared with clients, and also fulfills requests by clients. Usually corporations will buy devices **dedicated** to being servers

Peer-to-Peer: All devices are considered as equals and load is distributed evenly among all devices, all devices act as both a client and a server

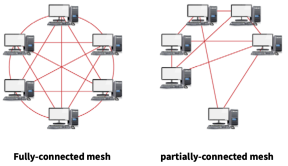
Comparison: Just memorise these and you're GOOD TO GO

| | Client-Server | Peer-to-Peer |
|---|---|---|
| Cost | High as the use of specialised high-performance servers would be needed | Low as basic devices can act as servers to share resources |
| Set-up | Harder to set up as specialised software and operating system features are needed for servers | Easy to set up as no specialised software or operating system features are needed |
| Security | High as rights can be controlled centrally at a server | Low as security is handled by each device and not by a central server |
| Control of resources | Centralised control of all data and resources | All computers store data locally |
| Backups | Easy to schedule backups of all shared files at regular intervals | More effort is required to access and back up resources as they are stored locally within each computer instead of centrally in a server |
| Bandwidth | Typically high as servers have **centralised control** to use bandwidth more efficiently and offer more reliable connections, but this relies on the capabilities of the server | Varies depending on how data needs to be transmitted; bandwidth may be reduced if a single device must handle many large request, but it can be increased if large requests can be divided into smaller requests that are handled by multiple devices simultaneously |

## 10.1.4 Star vs Mesh topology

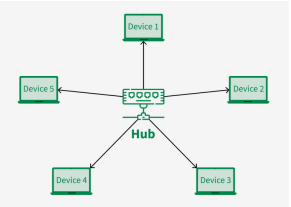
There's more than these but only these two are in our syllabus

**Mesh** topology (**Fully-connected**): Devices are all connected to each other and cooperate to relay messages to their intended destinations.



**Mesh** topology (**Partially-connected**): Some devices are connected to several devices, but not to all

**Star** topology: a network hardware like a switch or a wireless access point are at the centre of the network with connects to all other computers.




Common comparison factors:

| Factor | Star | Mesh |
|---|---|---|
| Application in home networks | Can be used for both wired and wireless home networks | Only suitable for wireless home networks, especially home automation |
| Complexity | Less complex | More complex |
| Cost | Typically lower cost depending on medium; only one piece of central network hardware is needed but cables for a wired network can be costly as the network grows | Typically higher cost; more network hardware needed |
| Robustness | Less robust as the central hardware is a single point of failure that will stop the entire network if it fails | More robust as there would typically be multiple different routes between devices, and the network would continue to function if any one device or hardware fails |
| Range | Typically smaller; limited by cable length and distance to the central hardware | Typically larger; each addition of network hardware extends the network's range without sacrificing speed |
| Standards | Highly standardised so hardware from different companies can usually work together | Currently not highly standardised so hardware from different companies may not work together |

## 10.2 Protocols n Packets

Devices in a network communicate using network protocols, and split up data into packets which are sent independently of each other

### 10.2.1 Parity bit

In an **even parity** bit system, the total number of 1-bits (including parity bit) should be even. In an **odd** parity bit system, the total number of 1-bits shld be **odd**.

E.g. for a byte 00001111

Even parity system: 000001111 or 000011110

Odd parity system: 100001111 or 000011111

## 10.2.2 Checksums

Basically, you assign a value to each number or smt, then do a **specific algorithm** which gives you the checksum. Then you transmit the checksum as well

Example in textbook: You sum up the denary values of all the data and take the result modulus 256. Then you send that.

But rlly u can do anyth uw

## 10.2.3 Echo checks

The receiver sends back a **complete copy** of all the data received to the sender, the sender compares it to what was sent, if anyth is different, sender re-sends

Advantage is that it's a lot more thorough and can detect more kinds of errors compared to like parity bits and checksums

However they roughly **double** the amount of traffic and time needed to complete each transmission, and without further checks, the sender cannot tell if the error occurred during transmission or when the echo was transmitted. And it doesn't deal with a situation when the data is completely lost and the receiver receives zilch

## 10.2.4 Automatic Repeat reQuests

This is basically everything rolled into one

The sender sends data with some corruption shit like parity check / checksum.

If the receiver processes it and determines no corruption, it sends a positive acknowledgement (**ACK**) back to the sender.

If it's corrupted, it sends a negative acknowledgement to the sender (**NACK**). Sender re-sends.

If the sender receives no acknowledgement after a certain amount of time (**timeout period**), sender assumes data was lost and resends data

Advantage: addresses and recovers situations where data is lost and receiver receives zilch

Disadvantages: By itself, it only deals with packet loss, it relies on other corruption checks as well like parity check and checksum. Also resending data when a timeout occurs could waste bandwidth if the receiver got the data and its find but the ACK was delayed or got lost

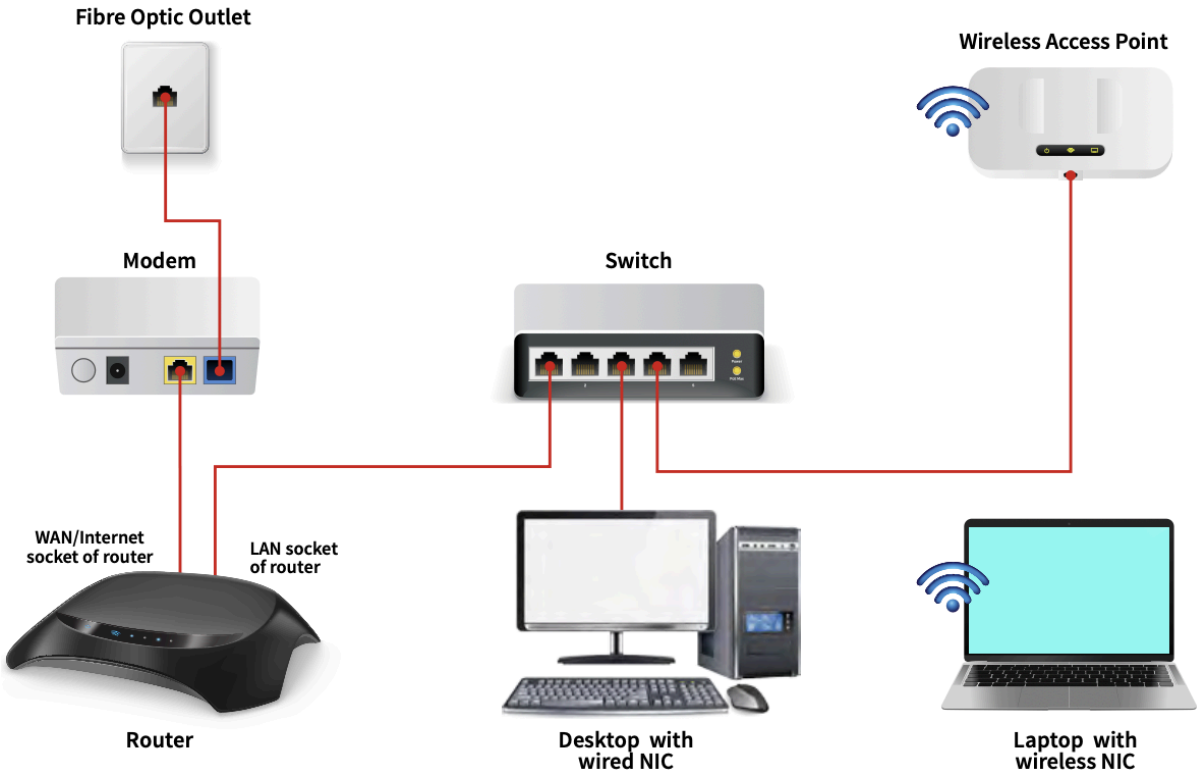
# 10.3 Network hardware

| | |
|---|---|
| Modem (Modulator-Demodulator) | Converts digital data into a form suitable for transmission and back |

| | |
|---|---|
| Network Interface Controller (NIC) | Provides hardware interface to enable the transfer of data between a device and network. Also gives u a Media Access Control (MAC) address |
| Network Switch | Constructs a network by connecting multiple devices together. It looks at MAC addresses, not IP addresses |
| Wireless Access Point (WAP) | Allows wireless devices to connect to a wired network, usually has an ethernet connection to a switch or smt |
| Router | Forwards packets between separate networks. Handles data using IP addresses instead of MAC addresses |

## 10.4 Addresses

| | |
|---|---|
| Media Access Control (MAC) | Unique, permanent addresses of the NIC. Made up of 6 bytes, represented in hexadecimal |
| Internet Protocol v4 (IPv4) | Unique, dynamic address assigned to a connected device. 4 bytes, displayed in denary e.g. 192.168.68.62 |
| Internet Protocol v6 (IPv6) | Unique, dynamic address assigned to a connected device. 16 bytes, displayed in groups of 4 hexadecimal digits separated by colons. E.g. 2001:0db8:0000:0000:0020:0024:0bad:c0de (By convention, use lowercase. Most ppl shorten the 0s, so u can write it as: 2001:db8::20:24:bad:c0de) |




# 11 Security & Privacy

3 principles: **Confidentiality**, **Integrity**, **Availability**. See definitions in Appendix A

Privacy, personal data n security see definitions also

## 11.1 Threat actors

| | |
|---|---|
| Human actions | Physical means: Like a storage medium, you drop it causing physical impact or like subject it to extreme temps causing data corruption<br>Or maybe you forgot to eject storage media before physical removal |
| | Non-physical: Overwriting data due to multiple users working on the same file that is shared on a network<br>Or data corruption due to incorrect data entry |

Malware

| | |
|---|---|
| Adware | Malware that installs itself without the user knowing and displays unwanted advertisements |
| Spyware | Hidden malware that secretly collects users' personal data and transmits this information to the attackers without the user's knowledge |
| Trojans | |
| Viruses | |

| | |
|---|---|
| Phishing | The use of emails, messages and fake websites that appear to be from reputable companies to steal personal data from users |
| Pharming | The interception of requests from a legitimate website and redirection of those requests to a fake website to steal personal data |
| Cookies | Small pieces of data stored by the web browser when a user visits a website |

Defenses

| | |
|---|---|
| Anti-malware | A program that detects, remove and stop malware from running |
| Firewall | Something that blocks traffic to and from unauthorised senders or receivers |
| Encryption | The process of encoding a message such that a secret key is needed to decode the message |
| Personal Data Protection Act (PDPA) | A law that governs the collection, protection, and use of personal data in Singapore |
| Computing practices (Strong Passwords) | Shld include Numbers, Lowercase letters, Uppercase letters, n Symbols. Makes it harder to hack |
| Computing practices (MFA) | **Multi factor authentication** such as using biometrics and security questions |

Scam email elements:

| | |
|---|---|
| Requests for personal data | "Email, password" |
| Suspicious links | "Go to … to sign up" |
| Errors | Wrong grammar |
| Generic greetings | "Dear customer" |

# 12 Intellectual property

For definition: Look at Appendix A definitions

Copyright and its infringement: Look at Appendix A definitions

## 12.1 Software Licenses

| | |
|---|---|
| **Proprietary Software** | Commercial software, most of the legal protections under copyright are retained. E.g. Microsoft Excel, macOS, WindowsOS |
| **Shareware** | Demonstration proprietary software that is distributed for free but for a specific |

| | |
|---|---|
| | evaluation period only |
| **Freeware** | Proprietary software that is available for use at no cost, **or** limited version of a paid proprietary software |
| **Free and Open Source Software (FOSS)** | Software where users have the freedom to change, copy, study and share software. Free refers to freedom to **copy, modify and distribute**, not price. |
| **Public Domain software** | Not a license, but basically this is when there is NO copyright at all. |

Software piracy: Check definitions

What u do when u hv a crack: ask help form a cher or parent to check its legality cos a lot of it can be replaced by an alt open source program[3]

Avoid copyright infringement: Use public domain material instead, and for educational purposes, limit use of it to 10% and properly cite it

# 13 Impact of Computing

Rlly, just bs this. Like I'm not even kidding. Just BS it.

Just note:

POFMA (Protection from Online Falsehoods and Manipulation Act):

if you spread falsehoods you are breaking it, and then u can get imprisoned, fined, or both. You may either get a "correction direction" where you just post a correction post that clarifies shit or in worse cases a "stop communication direction" which forces u to take down the original post

Yeah just see definitions. But gang you're like Gen Z wait actl Gen $\alpha$ BUT my point still stands there's no way you don't know what stuff like "feed", "engagement rate" means

Ethical ways of promoting content

- Creating content that is informative, entertaining and/or inspiring to generate interest among users
- Being inclusive, sensitive and/or respectful to avoid driving away users

Unethical ways of promoting content

- Spreading falsehoods or half-truths that purposely leave out important context to attract attention
- Posting controversial opinions that are intentionally divisive or insensitive, so users are enticed to comment
- Impersonating notable people or companies

[3] This is what you SHOULD do, but if you're taking Cp+ you probably have a lotta them so idk idc

# 14 Emerging Technologies

AI: See definition

Uses of AI:

| | |
|---|---|
| **Facial recognition** | AI can detect if a photo or video contains human faces and locate those faces in a frame, and also differentiate between different faces. Used in security systems such as for authenticating users to unlock devices |
| **Voice recognition** | AI can convert spoken language into text, as well as differentiate between different voices, which is used to operate computers in a hands-free fashion, such as giving voice commands to a smartphone. Also used to **automate** transcription jobs, like generating meeting notes |
| **Image classification** | AI can analyse images n organise them into categories. E.g. Social media sites can detect and block undesirable content, doctors can use it to detect abnormal conditions from medical images (X-rays) |
| **Spam filtering** | AI can identify and separate unsolicited and unwanted emails (spam / junk mail) from legitimate emails, used by email services to protect users from phishing, malware, other cyberattacks that requires users to open an email |
| **Game playing** | AI can analyse the state of a game and serve as an opponent or a substitute for another human player. Often used in video games to provide players with a challenge / training through "bots" |
| **Content generation** | AI can create new content (including text, images, sounds and videos), based on a provided input prompt (which can also include text, images, sounds and videos). Can be used to brainstorm ideas, produce reports, write code, create artwork, imitate voices, produce animations, rlly a lot. This AI is called generative AI |

ML: Basically a **technique** in **AI**. You **train** a computer on **training data** to get a **model**. When you apply it on new data, that's called an inference. ML vs traditional shit:

| Aspect | ML | Traditional Programming |
|---|---|---|
| Approach | Provide a large amount of data for the system to identify patterns and make its own rules | Write explicit rules and instructions for given scenarios |
| Flexibility | Adapts to new data, improves over time with more data | Requires manual modification to change behaviour |

| | | |
|---|---|---|
| Predictability | Depends on type of ML system, output is probabilistic and may vary each time for the same input | Output is typically deterministic and repeatable for the same input |
| Speed | Training speed is slow, inference speed depends on the size of the model and available resources | Development and execution are usually faster for well-defined problems with clear rules and requirements |
| Cost | Typically higher due to the need for specialised hardware (GPUs) and high power usage | Typically lower cost due to lower hardware and power requirements |
| Ideal use case | Solving problems with complex patterns that are difficult or impossible for humans to formulate explicit rules | Solving well-defined problems with clear rules and requirements |

## 14.1 Ka Ni Na

kNN: **k-nearest neighbour** (The plus of computing+)

The definition of Nearest Neighbour (NN) method is in definitions. No k there cos not in textbook

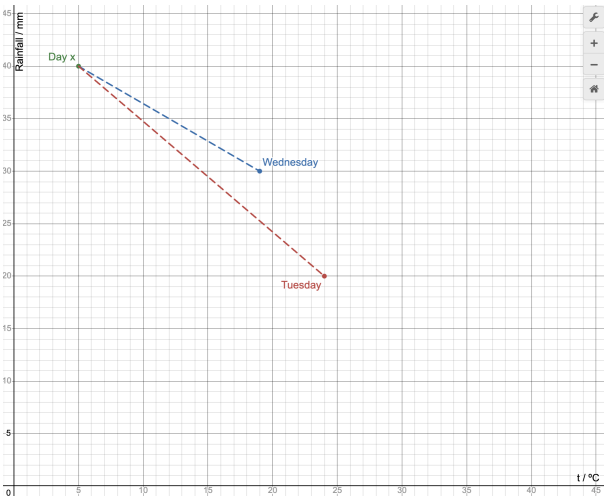
A **data point** is an instance of data. Each one has features and a label.

E.g. (bs only ah)

There was 20mm of rain at 24°C on a **Tuesday**

There was 30mm of rain at 19°C on a **Wednesday**

Is the day it rained 45mm at 5°C a tuesday or wednesday?




Using 1-NN (1 nearest neighbour), it's a **Wednesday**. Why? Look:

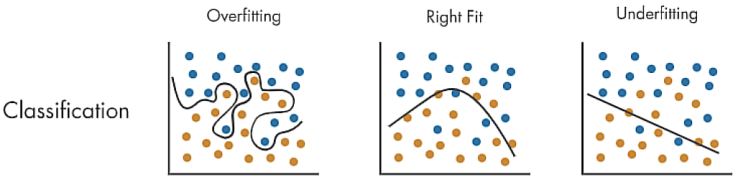
There's a lotta calculation, but from here it's obvious the blue line (distance from a Wednesday) is shorter than the red line. So it's chosen to be a Wednesday. Equal distance → coin flip

In kNN, basically, you take k number of neighbours, and choose the label that's the most common. So like if you did 3 nearest neighbours, 2 were **Wednesday** and 1 was **Tuesday**, you'd choose **Wednesday**. For tiebreakers, there's many ways, but you can choose like which label has the **shortest** distance which is p common.

## 14.2 Risks

There are risks in being lazy and giving your job to an AI (duh)

Firstly watch out for overfitting and underfitting (see definitions in appendix)




Another problem is **bias**, which can be from **prejudice bias**, where training data reflects biases and prejudices that already exist in society, e.g. most nurses being women, in which an AI for recommending jobs may inadvertently suggest positions based on **stereotypes** instead of skills and experience, or **sample selection bias**, where the selection of training data is unbalanced and does not adequately represent the population, e.g. only training the model on what a certain group of people say which can result in **unfair outcomes for minority groups**

**Privacy, Surveillance, Copyright:**

ML requires a **lot** of data to work well, which may include personal data that are collected or used without clear permission of the individuals involved. In SG, the PDPA forces organisations to seek consent and disclose their usage of personal data, .etc.

For **AI-generated** content, who has the copyright? In Singapore, you need human authorship, so AI cannot copyright content.

Gen AI may also train on copyrighted materials without a license from copyright owners, but there's an argument that it's "fair use" of copyrighted materials because it's for research. As of 2024 there are many lawsuits against various providers of gen AI regarding the use of their copyrighted content for training, but the outcome of it is still uncertain

Gen AI may **also** generate copyrighted material. For example, generating content that closely resembles copyrighted materials, like an artist's style, and overfitted models may even just output the exact same copyrighted material, which could count as plagiarism and/or copyright infringement

## 14.3 Misinformation

Gen AI may **hallucinate**, which can cause misinformation to be spread if people do not check what AI tells them, as well as create **deepfakes** which can be used to maliciously manipulate public opinion as well as damage reputations.

# 15 Appendix A: Definitions

| | |
|---|---|
| **Address** | A number that is used to locate a byte in memory |
| **Address bus** | A bus that is used to specify memory address information |
| **Adware** | A type of malware that installs itself without the user's knowledge and displays unwanted advertisements |
| **Artificial Intelligence (AI)** | The ability of a computer to perform complex tasks without constant human guidance and improve its performance as more data is collected |
| **Authentication** | The process of verifying the identity of a user |
| **Authentication factor** | A category of evidence that is used for authentication: something the user knows or owns, or something that is measured from a physical part of the user |
| **Automatic Repeat reQuests (ARQ)** | An error detection and correction method where the sender automatically resends the data if no acknowledgement from the receiver is received within the timeout period |
| **Anti-malware** | Software to detect, remove and stop malware from running |
| **Availability** | The ability to access data in a timely and uninterrupted fashion |
| **Bandwidth** | The average number of bits of data that can be transmitted from a source to a destination over the network in one second; usually measured in megabits per second (Mbps) or gigabits per second (Gbps) |
| **Bias (AI)** | When an AI system produces results that are systematically prejudiced or discriminatory in an unfair manner |
| **Bi-directional** | Able to work in two directions, to and fro |
| **Biometrics** | A type of authentication based on the measurement of human physical characteristics |
| **Binary number system** | A number system that is made up of two unique digits |
| **Bit** | A binary digit that takes on a value of either 0 or 1 |
| **Boolean** | A data type representing either True or False |
| **Byte** | A binary number made up of eight bits |
| **Central processing** | The part of the computer that processes data and follows |

| | |
|---|---|
| **unit (CPU) or processor** | instructions |
| **Character (text)** | A unit of textual data, usually a symbol, letter, digit or space |
| **Character encoding** | The assignment of numbers to characters |
| **Checksum** | A calculated value that is used to determine the integrity of transmitted data |
| **Client** | A device that initiates a connection to a server to request for resources and services |
| **Computer** | A device that receives and processes data according to a set of instructions and produces the processed data |
| **Computer Architecture** | A description of how a computer is designed and built to function, including how its various parts are designed, organised and connected |
| **Confidentiality** | The protection of data from unauthorised access |
| **Cookies** | Small pieces of data stored by the web browser when a user visits a website |
| **Copyright** | The legal right of owners to control the use and distribution of their intellectual property |
| **Copyright infringement** | The use or distribution of copyrighted work without the permission of the copyright owner |
| **Crack** | A program that modifies software so that the software cannot detect that it is being used illegally |
| **Data (singular Datum)** | Information that is used in a computer program |
| **Data bus** | A bus that is used to transport data between memory and the processor |
| **Data corruption** | The introduction of errors and inconsistencies into data |
| **Data point (ML)** | A unit of information representing an instance or example relevant to a ML problem |
| **Deepfake** | A convincing AI-generated depiction of people doing things they never actually did |
| **Demodulation** | Conversion of transmitted signals into digital data |

| | |
|---|---|
| **Denary number system (Decimal number system)** | A number system that is made up of 10 unique digits |
| **Denial-of-service attack** | A type of attack where a server is overwhelmed with excessive requests such that it cannot respond or responds very slowly to legitimate requests |
| **Echo check** | An error detection method where the receiver sends a copy of what was received back to the sender |
| **Encryption** | The process of encoding a message so that a secret key is needed to decode the data |
| **Engagement rate** | A measure of how much other users actively interact with a social media post |
| **Ethernet** | The most commonly used wired network standard for local area networks |
| **Execute** | To follow or perform an instruction |
| **Feature (ML)** | A measurable property or characteristic of a data point |
| **Feed (Social Media)** | A feature of social media sites where posts from other users are displayed as a continuously updated stream of content |
| **Firewall** | A device or computer program that prevents unauthorised access to or from a network |
| **Free and open-source software (FOSS)** | Software where users are given freedom to change, copy, study and share the software and its source code |
| **Freeware** | Proprietary software that is available for use at no cost |
| **Gateway** | A device that serves as an entry and exit point for communication with devices that are not in the local network |
| **Generative AI** | AI that can create new content such as text, images, sounds and videos |
| **Hardware** | Physical components of a computer |
| **Hallucination (AI)** | Falsehoods that are presented as facts by gen AI |
| **Hard drive** | Secondary storage where data is stored on rigid rotating disks coated with a magnetic material |

| | |
|---|---|
| **Hexadecimal number system** | A number system that is made up of 16 unique digits |
| **Inference (ML)** | The process of applying an ML model on new data to make predictions or decisions |
| **Input (Computer architecture)** | Data or instructions that the computer receives for processing |
| **Input device** | A hardware device that allows users to enter data and instructions into a computer |
| **Integer overflow** | Result when an operation produces an integer that exceeds the limit of the representable range |
| **Integrity** | The protection of data from unauthorised modification |
| **Intellectual Property** | Creations of the mind that have value but can exist purely as data with no physical form |
| **Internet Service Provider (ISP)** | An organisation that offers services for accessing the internet |
| **Label (ML)** | The target output of each data point |
| **Lane (PCIe)** | An interface for transferring data between a computer and an expansion card |
| **Leading zero** | The zero digit (0) that is to the left of the first non-zero digit in a number |
| **License** | An official description of activities that are authorised or forbidden by the owner of intellectual property |
| **Local Area Network (LAN)** | A network of computing devices connected within a small geographical area, typically within the same building, such as a home, school or office |
| **Logical operation** | An operation that acts on binary inputs to produce an output according to the laws of Boolean logic |
| **Logic circuit** | A circuit that consists of only one logic gate or multiple connected logic gates |
| **Logic circuit diagram** | A visual representation using symbols to show how the component logic gates are connected together in a logic circuit |
| **Logic gate** | An electrical circuit that performs logical operations based on one or more inputs and produces a single output |

| | |
|---|---|
| **Machine Learning (ML)** | A technique used in AI that enables a computer to perform complex tasks without explicit programming |
| **Malicious software (Malware)** | Software that is intentionally used to damage, disrupt, or gain unauthorised access to a computer system |
| **Memory (Main Memory)** | A device that is used to store data and instructions temporarily for immediate use by the processor |
| **Motherboard** | The main circuit board in a computer than connects all the components together |
| **Model (ML)** | A mathematical representation of the patterns and relationships that are found as a result of training an ML system |
| **Modem** | Device responsible for modulation and demodulation |
| **Modulation** | Conversion of digital data into a form suitable for transmission |
| **Nearest neighbour (NN) method** | A simple ML algorithm for classification that assigns to a data point the label of its closest neighbouring data point in the training data |
| **Network** | A system of two or more devices that are connected by a transmission medium for the exchange of data |
| **Network protocol** | A set of standards and rules that govern how two or more devices communicate over a network |
| **Overfitting (ML)** | When an ML model learns random noise present in the training data to the extent that it results in poor generalisation when applied to new data |
| **Output (Computer architecture)** | Intermediate or final results produced by the computer in the form of processed data |
| **Output device** | A hardware device used to display, project or print processed data from a computer |
| **Packet (Computing)** | A unit of broken-up data containing a header with information that are needed for transmission, such as the source and destination addresses |
| **Packet number (or sequence number)** | A number in a packet's header that indicates the position of that packet in the sequence of data being transmitted |
| **Parity check** | An error-checking technique which uses a parity bit to detect errors |

| | |
|---|---|
| **Parity bit** | An additional bit, which can be either 0 or 1 depending on whether an odd or even parity system is used, that is appended or prepended to a string of binary data before transmission |
| **Personal data** | Any data related to a person that may allow that person to be identified |
| **Personal Data Protection Act (PDPA)** | A law that governs the collection, protection and use of personal data in Singapore |
| **Post (Social Media)** | Any content (e.g., text, pictures, videos, links) shared by a user on social media |
| **Pharming** | The interception of requests sent from a computer to a legitimate website and the redirection of those requests to a fake website to steal personal data |
| **Phishing** | The use of emails, messages and fake websites that appear to be from reputable companies to steal personal data from users |
| **Privacy** | The practice of protecting the confidentiality and control of personal data |
| **Proprietary Software** | Commercial software for which most of the legal protections under copyright are retained |
| **Router** | Network hardware that forwards packets between separate networks |
| **Secondary Storage** | A way of storing large amounts of data that will not be lost when the power supply is interrupted |
| **Security** | The practice of protecting the confidentiality, integrity and availability of data |
| **Security token** | A device that is used specifically for authentication purposes |
| **Server** | A device that shares resources with and responds to requests from clients |
| **Service set identifier (SSID)** | A string of up to 32 bytes that identifies the Wi-Fi network created by a Wireless Access Point (WAP) |
| **Shareware** | Demonstration software that is distributed for free but for a specific evaluation period only |
| **Signature (Malware)** | Identifying characteristics that are used to detect a known version of some malicious software |

| | |
|---|---|
| **Software** | A set of instructions to perform specific tasks on a computer |
| **Software piracy** | The crime of copying, distributing and/or using copyrighted software in a manner that is not permitted by its license |
| **Spyware** | A type of hidden malware that secretly collects personal information about its users and transmits this information to attackers without the user's knowledge |
| **Sum-of-Product (SOP)** | A Boolean statement where the AND terms are ORed together |
| **Timeout (ARQ)** | A time limit used by the sender to wait for an acknowledgement before it is assumed that the data or acknowledgement itself was lost in transit |
| **Training data (ML)** | A collection of examples that is used to teach an ML model how to make predictions or decisions |
| **Transmission medium** | A means of connecting two or more computers together using copper cables, radio waves or light pulses, so that they may exchange data and interact with each other |
| **Truth table** | A table that shows the resulting output for every possible combination of inputs |
| **Two's complement** | System used by computers to represent both positive and negative numbers |
| **Two-factor authentication** | A type of authentication that uses evidence from both something the user knows and something the user owns |
| **Underfitting (ML)** | When an ML model is too simple to learn the underlying patterns of the training data |
| **Uni-directional** | Able to work in one direction only |
| **Volatile** | Lost when the power supply is interrupted |
| **Wide Area Network (WAN)** | A network of computing devices covering a large-scale geographical area, typically across multiple geographical locations |
| **Wired network** | A network of devices connected by physical cables |
| **Wireless access point (WAP)** | Network hardware that lets wireless devices connect to a wired network |
| **Wireless network** | A network of devices in which signals are transmitted without |

| | |
|---|---|
| | the use of physical cables |