1. Computer Architecture

Name of unit	Symbol	Size in bytes
kilobyte	kB	10^3 1,000
kibibyte	KiB	2^{10} 1,024
megabyte	MB	10^6 1,000² = 1,000,000
mebibyte	MiB	2^{20} 1,024² = 1,048,576
gigabyte	GB	10^9 1,000³ = 1,000,000,000
gibibyte	GiB	2^{30} 1,024³ = 1,073,741,824
terabyte	TB	10^{12} 1,000⁴ = 1,000,000,000,000
tebibyte	TiB	2^{40} 1,024⁴ = 1,099,511,627,776
petabyte	PB	10^{15} 1,000⁵ = 1,000,000,000,000,000
pebibyte	PiB	2^{50} 1,024⁵ = 1,125,899,906,842,624

processor: It processes data and follows instructions given

data bus: transports data between memory & processor; bi-directional

address bus: transports required memory location from processor to memory; uni-directional

memory: stores data and instructions temporarily for immediate use by the processor

memory (RAM): - volatile
- loses its data when computer is switched off
- faster data

secondary storage: stores large amounts of data that will not be lost when the power supply is interrupted

secondary storage:
- non-volatile
- long term storage [up to terabyte]
- slower speed
- more durable

→ types:
- hard disk drives
- solid state drives
- flash drive / USB drives [portable]

Type	Magnetic	Optical	Solid-State
Description	Data is stored on a magnetic material that can be read or written by a magnetic "head" Example: hard disk Example: magnetic tape	Data is stored as very small pits or indentations that can be read or written by a laser Example: digital versatile disc (DVD)	Data is stored in electronic circuits that have no moving parts Example: memory cards Example: solid-state drive
Durability	More vulnerable to damage from • magnetic fields, • heat, • impact, and • natural deterioration over time	Vulnerable to damage from • scratches, and • natural deterioration over time More resistant to • heat • impact	Most durable; Most resistant to • impact, and • temperature changes
Portability	Heavier and bulkier than optical and solid-state media	Portable due to small size and light weight	Portable due to small size and light weight
Typical Capacities	Up to TBs of data	Up to GBs of data	Up to TBs of data
Cost per GB	Lowest	Lower cost than solid-state media but higher cost than magnetic media	Highest
Speed	Slower than solid-state media	Slower than solid-state media	Faster than magnetic and optical media

Table 2.2 Three main types of storage media

USB: used for powering and/or communicating with external devices
e.g. printers, mouse, keyboards
communicating powering

HDMI: used for delivering audio and video data to compatible devices
e.g. monitors, TVs

PCI Express: for communicating with internal expansion cards
e.g. graphics card

main memory (RAM): The purpose is being able to directly access the processor, and it has volatile storage

secondary storage: The purpose is to provide long-term storage for files, and it has non-volatile storage

2. Data representation

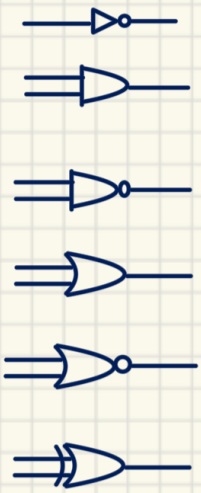
smallest non-negative 2-digit hexadecimal no. : 00
largest 2-digit hexadecimal no.: FF

common uses for hexadecimal number system:
- RGB colour code
- ASCII code
- IP address
- MAC address

two's complement:
1. convert to binary form
2. flip all bits
3. add 1 *

Describe how characters are represented in binary.
Each character has a unique Extended ASCII number value, which can be represented up to 256 distinct characters. The extended ASCII value of the character is converted to a 8 bit binary number.

3. Logic Gates



NOT gate
AND gate
NAND gate
OR gate
NOR gate
XOR gate

De Morgan's Theorem:
1. $\overline{A \cdot B} = \overline{A} + \overline{B}$
2. $\overline{A + B} = \overline{A} \cdot \overline{B}$

Distributive:
$A \cdot (B + C) = (A \cdot B) + (A \cdot C)$
$A + (B \cdot C) = (A + B) \cdot (A + C)$
$A \cdot (B \oplus C) = (A \cdot B) \oplus (A \cdot C)$

4. Programming

parameters
function name

```
def area (num1, num2):
XX return (num1 * num2)
print area (12,2)

print (ord ("65"))
```
↳ number to letter

```
print (chr ("A"))
```
↳ letter to number

add variable to dictionary:
```
scores [`John'] = 174
```

remove variable from dictionary:
```
del scores ['Bala']
```

to see all variables on seperate lines:
```
scores = {'Alex': 176, 'Bala': 180, 'Siti': 177}
for name in scores:
XX print (name)
```

read

```
with open ('file.txt', 'r') as f:
    f.read().split ('\n')

with open ('file.txt', 'w) as f:
    f.write (' ')
```

5. Input validation

- ask for input again
- exit the program

common validation checks:

1. Length check
   - length of inputs need to meet certain requirements

2. Range check
   - limit the input to a particular range of values
   - uses these operators:
     - less than (<)
     - less than or equal to (<=)
     - greater than (>)
     - greater than (>=)

3. Presence check
   - ensures all the required inputs are provided
   - uses special functions or compare inputs to values such as None or an empty string(" ")

4. Format check
   - requires the input to satisfy additional complex requirements

5. Existence check
   - checks if the input already exists

Common validation checks:

- .isdigit( ) = checks for digit
- .startswith( )
- .islower( )
- .isspace( ) = checks if every character is a space
- .lower( )
- .upper( )
- .endswith( )
- .find(' ')
- .isalnum( ) = checks every character is a letter or number
- .isalpha( ) = checks if all is letters

6. Testing and Debugging

Types of program errors:

1. Syntax error
   - errors caused by incorrect source code that violate the rules of the language
   - e.g. – spelling mistakes [case sensitive]
     – incorrect indentation
     – invalid variable names
     – incorrect sequence of symbols

2. Logic error
   - errors that cause the program to not behave as expected
   - e.g. – use of a flawed or incomplete algorithm

3. Run-time errors
   - errors that occur while a program is running
   - e.g. – incorrect use of commands
     – input data that has not been properly validated (str to int)
     – conditions outside of the program's control

Designing test cases:

1. normal conditions
   → expected output is printed out
2. boundary conditions
   → between a range
3. error conditions
   → test something that is out of range for data validation

7. Algorithm design

Finding the minimum / maximum in a list
w/o using the min( ) or max( ) functions
→ minimum_value / maximum_value

Splitting without .split( )

```
S = input(" Enter string: ")
delimeter = input(" Enter delimeter: ")
result = [ ] # empty list
current_el = " "
for char in s:
    if char != delimeter:
        current_elt= char
    else:
        result t= [current_el]
        current_ el = " " # initialise again!!

result t= [current_el]
print (result)
```

8. Software Engineering

Traditional "waterfall model"
1. Gather requirements
2. Design solutions
3. Write code
4. Test and refine code
5. Deploy code

9. Excel Spreadsheets

= ceiling.math → rounds a number to its closest whole number
= floor.math → rounds down
= mod(,) → returns remainder
= power(,) → put power first
= quotient(,)
= rand( ) → returns number from 0 to 1.
= randbetween(,) → returns number from (,)
= round(,)
↳ decimal place
= sqrt( ) → returns square root
= sum( ) → returns total
= sumif(range, criteria, " ")
↳ optional
= average( )
= averageif (range, criteria)
= count( ) → counts number of cells
= counta( ) → counts everything
= countblank( ) → returns number of empty cells
= countif (range, criteria)
= large (range, position)
↳ 1 will be the largest, 2 will be second largest
= max (range) → same as large (range, 1)
= median (range) → no error even if there is text
= min (range) → smallest number in range
= mode.sngl( ) → highest frequency
= rank.eq (number, array, 0/1) → if number is repeated, it is considered a number on its own
↓ descending ↳ ascending
= small (range, position)
↳ from smallest to largest
= concat( ) → joins 2 texts together
= left (text, number of characters)
= right (text, number of characters)
= mid (text, starting number, number of characters)
= find ("NA", banana) → will return error
→ case-sensitive
= search ("NA", banana) → not case-sensitive

= vlookup (value, range, column number, TRUE/FALSE)
↓ → exact match
→ approx. match
= hlookup (value, range, row number, TRUE / FALSE)
= index (range, match (value, range))
= days (end_time, start_time)
= now ( ) → date + time
= today ( ) → date

10. Networking

wired network vs wireless network

- a network of devices connected by physical cables
- 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 the 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
Physical organisation	Tends to look more disorganised due to cables running across floors	More organised without cables

Local area network (LAN):
→ small geographical area
e.g. home/office / school

Wide area network (WAN):
→ large-scale geographical area
e.g. telephone lines

Client-Server network:

advantages
- centralised control of data and resources
- easy to schedule backups of all shared files at regular intervals
- security may be enhanced with the use of specialised software or operating system features that are designed for servers

disadvantages
- higher initial cost due to the need for a server
- administrative costs needed for the maintenance of server and clients

advantages
- cheaper to set up as there is no cost related to dedicated servers
- easy to set up as no specialised software or operating system features are needed.

disadvantages
- more effort is required to access and back up resources as they are stored locally within each computer instead of centrally in a server
- security is an issue as access rights are not administered by a central server

Factor	Star Topology	Mesh Topology
Applications in Home Networks	Suitable for both wired and wireless home networks	Typically 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 be typically be multiple different routes between devices, and the network would continue to function if any one device or hardware fails

Factor	Star Topology	Mesh Topology
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

Protocols: Standards & rules that govern communication over a network

Echo check:
When data is received, the receiver sends a copy of what was received back to the sender so the sender can compare it with what was orignally sent. If the 2 sets of data are different, the sender sends the data again
* manual checking

Automatic Repeat Requests:
When data is received using ARQ, the receiver sends a positive acknowledgement (ACK) back to the sender if no errors are detected. On the other end, the sender waits for an acknowlegement (either tve or -ve) from the receiver to arrive within a predetermined amount of time

Modem:
- performs modulation & demodulation
- modulation:
  converts digital data into a form suitable for transmission
- demodulation:
  → reverse process of modulation
  → converts transmitted signals back digital data
- provides Internet Service Provider (ISP) to the protocols used by local network (LANs) in a home network

Network Interface Controller (NIC):
- provides the hardware interface to enable the transfer of data between a device and a network
- any device that connects to a network needs to have an NIC
- for home networks, an NIC typically connects to a network via an Ethernet cable or wirelessly using Wi-Fi

Network switch:
- constructs a single network by connecting multiple devices together
- for home network, switches are typically used to construct a LAN using Ethernet cables
- uses MAC addresses

Wireless Access points:
- lets wireless devices connect to a wired network
- typically has an Ethernet connection to either a switch or a router and allows wireless devices to join the network using Wi-Fi

Router:
- forwards packets between separate networks
- keeps the connected networks separate and forwards packets between them using Internet Protocols

Figure 10.31 Setting up a modem, a router, a switch and a wireless access point to form a home network

Figure 10.32 Simplified home network using a wireless router

11. Security and Privacy

Security:
- practice of protecting the confidentiality, integrity and availability to data
  → data in general!

principle	definition
confidentiality	protection from unauthorised access
integrity	protection from unauthorised modification
availability	ability to access the data in a timely and uninterrupted fashion

Privacy:
- practice of protecting the confidentiality and control of personal data

Human actions:
- Physical:
  → exposure of storage media to physical impact / extreme temperatures
  → data corruption
- Non-physical:
  → overwriting of data due to multiple users working on the same file that is shared on a network
  → data corruption due to incorrect data entry

Adware:
- malicious software
- installs itself without user's knowledge and displays unwanted advertisements

Spyware:
- hidden malware
- secretly collects personal information about its users and transmits this information to attackers without the users' knowledge

Cookies:
- small pieces of data stored by the web browser when a user visits a website

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 sent from a computer to a legitimate website and the redirection of those requests to a fake website to steal personal data

Anti-malware programs:
- detect when malware is about to be run and stop it from running;
- detect malware that is already running and tries to stop it;
- scans the user's computer to detect and remove malware

Encryption:
- process of encoding a message so that a secret key is needed to decode the data

Firewall:
- a device or computer program that prevents unauthorised access to or from a network

Personal Data Protection Act (PDPA):
- a law that governs the collection, protection and use of personal data in Singapore

12. Intellectual Property

Intellectual property:
- creations of the mind that have value but can exist purely as data with no physical form
- must have:
  1. Originality and Creativity
  2. Fixation
  3. Ownership

Copyright:
- legal rights of owners to control the use and distribution of their intellectual property under the Copyright Act

License:
- An official description of activities that are authorised / forbidden by the owner of intellectual property

Freeware:
- Proprietary software that is available for use at no cost

Free and open-source software (FOSS):
- Software where users are given freedom to change, copy, study and share the software and its source code

Proprietary Software:
- Commercial software for which most of the legal protections under copyright are retained

Shareware:
- Demonstration software that is distributed for free but for a specific evaluation period only

Software piracy:
- illegal use and distribution of copyrighted computer programs in a manner that is forbidden by their license

Crack:
- A program that modifies software so that the software cannot detect that it is being used illegally

Software privacy:
- The crime of copying, distributing and / or using copyrighted software in a manner that is not permitted by its license

13. Impact of computing

Algorithms and social media:
1. To increase the site's relevance to users
2. To incentivise the creation of relevant posts
3. To increase advertising revenue

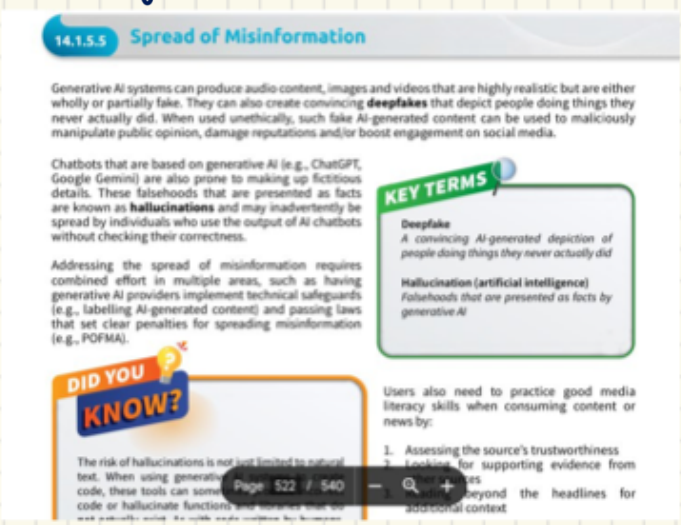
Protection from Online Falsehoods and Manipulation Act (POFMA):
- if found to have engaged in prohibited activities under the Act
  → subjected to be fined or imprisoned or both

14. Emerging Technologies

Artificial Intelligence (AI):
- ability of a computer to perform complex tasks without constant human guidance and improve its performance as more data is collected

Machine Learning (ML):
- technique used in AI and explain the difference between machine learning and traditional programming

Common tasks performed by AI:
- Facial recognition
- Voice recognition
- Image classification
- Spam filtering
- Game playing
- Content generation



14.1.5.5 Spread of Misinformation

Generative AI systems can produce audio content, images and videos that are highly realistic but are either wholly or partially fake. They can also create convincing **deepfakes** that depict people doing things they never actually did. When used unethically, such fake AI-generated content can be used to maliciously manipulate public opinion, damage reputations and/or boost engagement on social media.

Chatbots that are based on generative AI (e.g., ChatGPT, Google Gemini) are also prone to making up fictitious details. These falsehoods that are presented as facts are known as **hallucinations** and may inadvertently be spread by individuals who use the output of AI chatbots without checking their correctness.

Addressing the spread of misinformation requires combined effort in multiple areas, such as having generative AI providers implement technical safeguards (e.g., labelling AI-generated content) and passing laws that set clear penalties for spreading misinformation (e.g., POFMA).

KEY TERMS

Deepfake
A convincing AI-generated depiction of people doing things they never actually did

Hallucination (artificial intelligence)
Falsehoods that are presented as facts by generative AI

DID YOU KNOW?

The risk of hallucinations is not just limited to natural text. When using generati... code, these tools can some... code or hallucinate functions ...

Users also need to practice good media literacy skills when consuming content or news by:

1. Assessing the source's trustworthiness
2. Looking for supporting evidence from ...ces
3. ...beyond the headlines for additional context

Page 522 / 540