# Cambridge GCSE Notes
# 5090 Biology

Abrar Faiyaz Rahim, degrees pending

# Contents

# 1 Data representation

## 1.1 Number systems

Computers use logic gates and registers to store data, consisting of two states, on or off which boil down to whether electrons are flowing through a gate or a register. Thus, all data must be converted to binary to be processed by a computer.

The binary number system consists of the digits 0 and 1 only. Because of this, it is called a base 2 number system[1]. Denary is the numbers we use everyday and is a base 10 system with digits 0 through 9. Hexadecimal is a base 16 system with the digits 0 through 9 and letters A through F[2].

Converting decimal to binary consists of dividing the number to be converted by two, noting down the remainders, until the quotient is zero and writing the remainders from bottom to top.

Converting binary to decimal consists of multiplying each binary digit with two raised to the power of its place value[3]. So, to convert the number $(1101)_2$:

| 3 | 2 | 1 | 0 | place values |
|---|---|---|---|---|
| 1 | 1 | 0 | 1 | number |

hence,

$$(1)(2^3) + (1)(2^2) + (0)(2^1) + (1)(2^0) = 13$$

To convert denary to hexadecimal, we must first convert to binary. We must split the resulting binary number into four-bit chunks. Each four bit chunk we will again convert to denary and we will write the corresponding hexadecimal digits. 0 through 9 in denary is the same as that in hexadecimal, only 10 through 15 in denary is A through F in hexadecimal. For example, the above example with $(1101)_2$ is the same as $\mathrm{D}_{16}$.

To convert hexadecimal back to binary, we must split it it into each digit, convert each part into binary, from there we may convert to denary.

Hexadecimal is easier to understand than binary because it is more concise, has more unique characters and hence mistakes in code can be easily found when sifting through hexadecimal code rather than binary. Hexadecimal codes are also used as colour codes in the format `#RRGGBBAA` where `RR` are two digits representing the red component of the colour, `GG` being the green, `BB` and `AA` being blue and alpha respectively.

To add binary numbers, note that:

$$\begin{gathered} 1 + 1 = 10 \\ 10 + 1 = 11 \\ 1 + 0 = 0 + 1 = 1 \\ 0 + 0 = 0 \end{gathered}$$

[1] A number system with base $n$ has $n$ digits.
[2] A number $x$ will be represented as $(x)_n$ where $n$ is the base of the system it is written in.
[3] $2^n$, where $n$ is the place value.

remember to carry any overflowing digits.

Every register has a maximum value, which corresponds to how many bits it can hold, an $n$-bit register has a maximum value of $2^n$. In addition, due to carrying over of bits, some bits may be outside that limit and hence will *overflow* and be ignored.

Bit shifts consist of simply shifting the bits in a register to the left or right:

| 0 | 0 | 1 | 0 |
|---|---|---|---|

Above is a number $(0010)_2 = (2)_{10}$ before a shift. After right shift:

| 0 | 0 | 0 | 1 |
|---|---|---|---|

it becomes $(0001)_2 = (1)_{10}$, meaning it has been divided by two. After left shift:

| 0 | 1 | 0 | 0 |
|---|---|---|---|

it becomes $(0100)_2 = (4)_{10}$, meaning it has been multiplied by two.

To represent negative numbers in binary, we use the two's complement method. Given $x_{10}$ that we must convert, we will first convert the value of $|x_{10}|$ to binary. We will then flip each binary bit that results, and add one to that result. This representation works as adding $|x_{10}|$ and the result will result in zero. So, as seen in a previous example $13_{10}$ is $1101_2$. To find the binary value of $-13$ we will flip the bits, getting $0010_2$. Adding one results in $0011_2$. This is a proper binary representation as $1101_2 + 0011_2 = 0000_2 = 13_{10} - 13_{10} = 0_{10}$. Note that this only works with a fixed register length, as a remaining bit overflowing is what allows the sum to be zero.

## 1.2 Text, sound and images

Text must be converted to binary to be processed by a computer. Each letter is assigned a denary value, which is converted to binary in the computer. The American Standard Code for Information Interchange (ASCII) specifies 256 denary numbers corresponding to 256 symbols. ASCII is an example of a character set, another of which is Unicode, which consists of a greater range of characters, including emojis and other languages. Each letter in a text file using the ASCII character set can hence be within 8-bit ($2^8 = 256$) but Unicode requires far more bits per character.

To represent sound, the computer uses "sound values" which are recorded at certain intervals, known as the sample rate. The number of bits used per sound value is called the sample resolution. The quality of the sound file depends on these two values. To find the size of a sound file:

$$\text{size} = \text{sample rate} \times \text{sample resolution} \times \text{number of samples}$$

Intuitively, we can say the higher the sample rate, resolution and number, the larger the sound file.

An image is a series of pixels, arranged in a matrix-like pattern. Each pixel is a value consisting of red, blue and green which determines it's colour. The

more the bits per pixel, the higher quality the picture. Bits-per-pixel is referred to as colour depth. The number of pixels in the image is called its resolution. The product of these give the file size, which increases with increase in any of these values.

$$\text{size} = \text{colour depth} \times \text{resolution}$$

## 1.3 Data storage and compression

Every 1 or 0 in a binary number is a bit. 4 bits is a nibble. 8 bits is a byte. 1024 bytes is a kibibyte (KiB). 1024 KiB is one mibibyte (MiB). 1024 MiB is a gibibyte (GiB), so on: tebibyte (TiB), pebibyte (PiB) and exibyte (EiB).

Data compression exists to reduce the size of the file, which results in less bandwidth required during file transmission, less storage space required and shorter time to transmit said file. Compression may be lossy or lossless.

Lossy compression consists of permanently removing data by reducing colour depth or sample rate or resolution.

Lossless compression reduces the file size without permanent loss of data, where it groups together repeating pixels with where they occur, called run length encoding (RLE).

# 2 Data transmission

## 2.1 Types and methods of data transmission

Large amounts of data, before transmission, is broken into smaller pieces called packets.

A packet consists of a packet header, a payload, and a packet trailer. The packet header holds the packet number, required for rearrangement of the packets after complete transmission, the sender's address and the destination address. The payload is the data being transmitted and the trailer holds any error-checking systems and also notifies the reciever that the packet has ended.

Data is transmitted across networks, which consist of multiple devices including routers. Routers decide the route of a packet, which it decides based on which route is busy. As a result, some packets may have taken a smaller route and may have arrived before others, causing them to be out of order. The recieving device uses the packet numbers to rearrange the transmitted data once the last packet has arrived.

Data can be transmitted as serial, parallel, simplex, half-duplex or full-duplex.

Serial transmission is the transmission of data in one bit at at a time. Data as a result, arrives in order, interference is less likely due to there being only one wire and hence is cheaper as only one wire need be bought and used. Data is also less likely to be skewed for the same reasons. However, transmission in this method is slow, and additional data called start bits and stop bits may need to be sent to inform the recieving device when the transmission has started and stopped.

Parallel data transmission consists of bits being transmitted simultaneously, across multiple wires. Data transmission in this method is quicker and since computers transmit data in parallel, there is no need for conversion. However, since multiple wires are being used the data may arrive out of order and may be skewed. Interference is more likely as well. Multiple wires are also pricey.

Simplex data transmission is where data can only be transmitted in one direction.

Half-duplex transmission is where data can be transmitted in both directions, but not simultaneously.

Duplex transmission is where data can be transmitted in both directions, simultaneously.

The Universal Serial Bus (USB) interface is an industry standard used to transmit data in serial. Such interfaces have only one correct connection, meaning no errors in connection can be made. The speed of transfer is also high in such interfaces. Howewver, a USB cable can only be 5 metres long, beyond which extenders must be used. Though transmission is fast, it is not as fast as ethernet.

## 2.2 Methods of error detection

Errors can arise during data transmission due to interference, which result in data loss, gain or change. So, to ensure correct data has been transmitted, we must check for errors.

Parity bits are used at the beginning or end of every byte, leaving 7 bits of actual data. Odd or even parity may be used. The method consists of counting the number of 1-bits in the 7 bits of data, if the parity is odd and the number of 1s is even, the parity bit is set to 1 to make the number of 1s in the number odd. Same applies for even parity. Below are examples where the leftmost bit is the parity bit.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | even parity |
| 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | odd parity |

The checksum method consists of using a certain algorithm to find a value using the data transmitted. The value is calculated on the sender's end and transmitted with the data, the reciever also calculates it and compares with the value the sender transmitted, if they match, data is errorless, otherwise an error is there in the data.

Echo checks consist of data being sent to reciever, and the reciever sending it back the to sender. The sender compares it's recieved data with that which it sent, if an error is found, the data is re-sent.

Check digits are used to check for errors during data entry. The check digit may already have been calculated and may exist in a database. During entry, the check digit is recalculated and compared with that stored in the database, if they are not the same, wrong data has been entered. This is mostly used during entry of data using barcodes and ISBN, the standard for book codes.

The automatic repeat query (ARQ) can be have negative or positive feedbacks. ARQ using negative feedback follows:

1. Sending device transmits packet
2. The reciever checks for errors
3. If no errors are found, no further action is taken
4. Otherwise negative acknowledgement is sent to sender.
5. Sender recieves negative acknowledgement and re-sends the data.
6. A timeout is set by sending device during transmission, if it recieves no acknowledgement after timeout, it will stop listening for acknowledgements.

Using positive feedback:

1. Sending device transmits packet
2. Recieving device checks for errors
3. If none found, positive acknowledgement is sent

4. If no acknowledgement sent within certain time, data is resent, this may happen for a set number of times before data transmission is stopped.

## 2.3 Encryption

Data must be encrypted during transmission so as to protect important data from being tapped. Encryption is done upon plain text to turn it to cipher text using encryption key.

Symmetric encryption uses the same key to encrypt and decrypt, the key is transmitted with the data to enable the recieving device to decrypt the message.

Asymmetric encryption uses public and private keys, which are generated together and a certain user's public key can encrypt the message in such a way that the decryption of that message can only be done by that user's private key. So, before transmission, the reciever sends the sender his public key, the sender encrypts data using it, and transmits the data. The reciever decrypts it using his private key.

# 3 Hardware

## 3.1 Computer architecture

The central processing unit (CPU) is where all instructions and data are processed, from inputs to produce outputs.

Microprocessors are processing units on single chips which are integrated circuits made to perform many of the functions of a CPU, but to a limited degree.

In a computer with the Von Neumann architecture, the CPU consits of two units, the arithmetic logic unit (ALU) and the control unit (CU). It has five registers: the program counter (PC), memory address register (MAR), memory data register (MDR), current instruction register (CIR) and accumulator (ACC). It also has the address, data and control buses.

In the Von Neumann architecture, instructions are processed in a fetch-decode-execute (FDE) cycle.

Data and instructions stored in the the random access memory (RAM), such instructions or data may be stored in the hard drive, before processing which they will be brought into RAM. The PC is a register which stores the address of the next instruction to be processed. This address, during the fetch stage, must be brought into the MAR, which is done by the address bus. The address of the instruction, now in the MAR, is used to find the data/instruction in RAM. The address bus goes to that location and sends the data stored there (be it an instruction or some values) to the MDR using the data bus. Using the data bus, the MDR sends the instruction to the CIR, which is part of the control unit, which is responsible for decoding the instruction.

Once the CU has recieved the instruction, it decodes it using an instruction set which is the set of all commands that the CPU is capable of executing. These commands are usually in machine code.

The execute stage may involve some mathematical calculations, done by the ALU. Any values needed to be stored temporarily are stored in the ACC.

The CU synchronises all the parts of the CPU to do what they need to, using the control bus.

Each CPU has a cache, multiple cores and an internal clock.

Each core in a CPU consists of the parts described above, hence the more the cores, the more instructions that can be processed simultaneously, increasing CPU performance.

The internal clock controls the speed at which instructions are processed. A clock speed of 1 Hz means 1 instruction is processed per second. The more the clock speed, the faster the CPU. However, increasing the clock speed beyond a certain value can cause overheating of the CPU and result in damage.

The cache is where instructions most commonly performed by the CPU is stored. The larger the capacity of the cache, the more instructions that can be stored, more instructions can be accessed faster, hence increasing CPU speed.

Embedded systems are used to perform dedicated functions, such as domestic appliances, cars, security systems, lighting systems or vending machines. This

is different to a personal computer in that these can perform many different tasks and are not usually hyper-specific.

## 3.2 Input and output devices

Input devices are those used to enter data into a computer system, including text, images and sound.

Barcode scanners are input devices which throw light onto barcodes, and depending on which parts are reflected and which are not, scan the information stored in the barcode into the computer.

Digital cameras use light and colour sensitive cells to form pixels, which, when arranged as a matrix can form an image.

Keyboards have buttons, called keys, which, when pressed, signal the computer that a certain letter has been input.

Microphones have sensors that produce different electric signals depending on the sound level around it, which can be input into the computer.

Optical mice work using a red-light source and a sensor to pick up said source. Using the changing patterns of light sensed, the location of the mouse is calculated and inputted into the computer.

A quick response (QR) code scanner does the same as a barcode, but instead of a linear series of dark and light lines, a QR code consists of dark and light squares on a matrix.

Touchscreens are of three types: resistive, capacitive and infrared.

Resistive touchscreens have two layers, when pressure is applied, the two layers come in contact and an electric circuit is completed, the position of contact can also be calculated. Such screens do not support multi-touch, give poor visibility in sunlight, and are easily scratched. One must also press down quite hard to use such screens. However, they are dust and water resistant, and can be used with anything (gloves, stylus, whatever).

Capacitive touchscreens have an electric field around them. When a finger is place in said field, the field is disturbed and the position of the finger can be calculated. Such screens have good image clarity, even in sunlight. They are resistance to scratched and support multi-touch. However, they only work with bare fingers or special styluses and are sensitive to electromagnetic radiation as those too disturb the electric field.

Infrared touchscreens consist of infrared rays passed across the screen in a matrix pattern. When a finger touches the screen, these rays are blocked and the sensors which do not recieve their respective rays can be used to calculate the position of the fingers. Such screens allow multi-touch, are durable and work with scratched or cracked screens. Contrastingly, they are water/moisture sensitive, accidental activation is possible if anything blocks the screen and are sometimes affected by light.

Two dimensional (2D) scanners are used to scan flat documents, done by lighting the document producing an image which is made electric by photosensitive cells.

Three dimensional (3D) scanners scan solid objects and produce digital images of actual objects, they can be used in computer aided design (CAD).

The result of computer processing is shown by output devices.

Actuators are output devices that consist of solenoids whose electromagnetic fields and electric fields result in a force acting on something.

Digital light processing (DLP) projectors pass light through micromirrors whose arrangement and number define the resolution of the digital image. The light after passing through these mirrors passes through a colour filter to produce colour on the projected image.

Inkjet printers have print heads which spray droplets of ink onto the paper to form characters. Ink cartridges from which they derive the ink, colours and all (using magneta, cyan and yellow combinations). A mechanism to move the print head side to side and a paper feed which simply gives the printer paper to print on. Such printers use either thermal bubble technology, which uses heat to partially vapourise ink to form a bubble and pass some of it onto the paper. When the bubble collapses, negative pressure draws more ink into the print head. Piezoelectric technology uses crystals, which, when they vibrate, force ink onto the paper.

Laser printers use ink powder instead of liquid ink. The paper to be printed upon is given an electric charge opposite to that in the ink, in places where the printing must occur. As a result, ink sticks to the paper. Once printing is completed, the charge on the paper is removed to prevent paper from sticking to the charged ink drum inside the laser printer.

Light emitting diodes, LEDs, can be arranged with red, blue and green components to create colours in light. Different amounts of different colours can result in images being produced.

Liquid crystal display (LCD) projectors pass light through red, green and blue mirrors. LCD screns, depending on the amount of colour on each pixel, block and allow light to pass through and finally it all converges to form an image through a prism.

LCD screens use changed electric fields to produce different colours. They are backlit by either LEDs or CCFL technology.

Loudspeakers convert stored sound values in the computer to physical sound waves by using those sound values to vibrate a physical cone which can then produce corresponding sound waves.

3D printers can print directly and additively, moving the print head wherever it requires and adding some material in those places. Sometimes, the material is first printed as a powder and then it is all made sturdy using glue. They can be used to make custom parts, prosthetic limbs, art, etc.

Sensors are input devices which read conditions of their surroundings and convert them to digital values using a digital to analogue converter (DAC).

Acoustic sensors are like microphones which take readings of sound levels in surroundings.

Accelerometers measure tehe change in velocity using a piezoelectric cell whose output varies with change in velocity.

Flow sensors produce output based on the amount of fluid passing around it per unit time.

Gas sensors use various methods to output the amount of the gas being monitored.

Humidity sensors measure the water vapour in a sample of air based on the conductivity of that sample.

Using ultrasonic or capacitive methods, level sensors sense the depths of liquids or caverns or whatever.

Light sensors use photoelectric cells to produce an output depending on the level of light being sensed.

Magnetic field sensors output the change in magnetic field.

Moisture sensors measure water levels in soil based on resistance of sample.

pH sensors use changes in voltages in sample to output a pH value.

Pressure sensors are transducers which generate different electric currents based on pressure applied.

Proximity sensors sense the presence of a nearby object.

Temperature sensors produce signals with change in temperature.

## 3.3 Data storage

Primary storage is that which can be directly accessed by the CPU. It consists of the RAM and the read only memory (ROM). The ROM contains instructions and programs to boot up to computer and the BIOS. RAM is where instructions to be processed are stored. Hence, RAM is an ever-changing part of storage, and is volatile in that whatever is stored in RAM will be lost once the computer is powered down. The ROM is non-volatile and the data stored in it never changes.

Secondary storage is that which is never directly accessed by the CPU, it too, is non-volatile. Permanent data is stored as secondary storage.

Secondary storage is done in magnetic, optical and solid state storage methods.

Magnetic storage consists of disks called platters on which, magnetism is used to store data. The disk is separated into sectors and tracks and the disk spins to allow data to be read and written. An electromagnetic head is used to read or write the data by magnetising dots on these tracks and sectors, an example of such storage are hard disk drives.

Optical storage utilises lasers to read and write data from a circular disk. Writing data consists of burning physical pits into the disk itself using the laser. Reading data is also done by a laser using the pits and lands and the data that they store. Examples of such storage include digital versatile disks (DVDs), compact disks (CDs) and Blu-rays.

CDs, read/write DVDs (DVD-RW) and Blu-ray disks have a single track, whereas DVD-RAM have multiple concentric tracks.

Such storage media require moving parts, which can go wrong. They are not very portable either as the read-write equipment for the technologies can be quite heavy.

Solid state storage lacks moving parts and uses semiconductor chips to store bits. Such storage is often called flash memory. Cells and transistors laid out in a grid, parallely in NOR structures and serially in NAND structures store the data. Using currents sent through the control gates of transistors, each transistor can hold the value of 0 or 1.

Hard disk drives are cheaper, and longer lasting than SSDs. They are also trusted technology. However, SSDs need not get upto speed, have faster read/write cycles, run quieter and cooler, are more portable and lighter, and are more compact.

When there is not enough space in RAM to hold all the data that it must for a certain task, virtual memory is created in the secondary storages. The excess data to be stored in this storage is structured into data structures called "pages". These pages are transferred back to the RAM when need be.

Cloud storage consists of servers in a remote location away from the user, on which the user may store what they want. In comparison to local storage, that which is stored in the cloud can be accessed from anywhere if one simply has an Internet connection. Data stored locally can only be accessed from the computer it is stored on and the network(s) it is connected to. For businessses, companies providing cloud storage can be more economical than storing their things themselves, sustaining their servers themselves and hiring personnel to do so, themselves. Cloud storage may be risky as such servers can be tapped by eavesdroppers.

## 3.4 Network hardware

To access networks[4], such as the Internet, a device requries a network interface card (NIC). NICs may be wired or wireless.

During manufacture, NICs are assigned media access control addresses (MACs), which are unique for each device and include the manufacturer's code and the device specific serial code. It falls into the following format:

```
XX:XX:XX:XX:XX:XX
```

or,

```
XX-XX-XX-XX-XX-XX
```

where each `XX` is an 8-bit hexadecimal number.

An internet protocol (IP) address is used to identify each unique device on the same network. Such addresses may be static, in that they will never change, or dynamic, in that they change each time the device connects to the network. IP addresses have two versions, IPv4, which is now quite old and deprecated, consisting of four 8-bit denary numbers seperated by a period (`.`). IPv4 falls into the following format:

```
XX.XX.XX.XX
```

[4] A network consists of routers, devices and transmission media

IPv6 falls into the following format:

`XXXX:XXXX:XXXX:XXXX:XXXX:XXXX:XXXX:XXXX`

where `XXXX` is a sixteen bit hexadecimal number.

# 4 The internet and its uses

## 4.1 The internet and world wide web

The internet is upon which the world wide web is based. In other words, the internet is the infrastructure, on which stands the collection of websites and webpages which make up the world wide web.

In short, the internet is a very large global network which allows users access to the world wide web. The world wide web is the collections of all websites and webpages available to the public.

Each website and webpage in the world wide web has a unique text based address called its uniform resource allocator (URL). A URL consists of three main components: the protocol, domain name and webpage or file name. It is arranged into a format resembling what follows:

```
protocol://www.domain-name/webpage
```

The way users obtain webpages from web servers is based on the hypertext transfer protocol (HTTP). It consists of users requesting the domain name server (DNS) for a certain website, and the web server responding with the IP of that website. Another protocol, built on top of HTTP is HTTPS, standing for hypertext transfer protocol secured. This provides another layer of security using digital certificates which are given to website owners by certificate authorities. The certificate is given if and when the website passes certain security measures tested by said authorities. This layer increases security in that the user will now check the authenticity of the digital certificate of the website, which the DNS will provide to the user, before exchange of website files.

The user does all this exchanging of website files by means of a web browser. A web browser is a piece of software into whose address bar the URL of the desired website is typed in. The web browser then requests the DNS for the website with that URL, if it exists, the DNS gives the browser the site's digital certificate of the site. If the certificate is authentic, the browser requests the webpage's IP address. Otherwise, the user is warned and asked if they want to proceed. In HTTP, this exchanging of certificates is absent and the webpages are exchanged with no intermediary security phases.

Websites are written in a programming language called HTML, standing for hypertext markup language. HTML uses "tags" to define colour, layout etc. The primary function of web browsers is to render and display to the user, the website written in this language. Modern web browsers provide some additional functions, such as they allow the user to bookmark certain websites and declare certain sites as favourites. The user's browsing history will also be recorded by the browser to make backtracking easier. Multiple tabs are offered for simultaneous browsing. Cookies are temporary files that websites store on the user's machine, managing such files is done by the browser. Navigation tools such as going to the previous webpage or to the next and such is provided by the web browser, alongside the most basic function of allowing the user to type in their desired address.

Cookies are files on the stored by a website on the user's device. They can be of two types, session cookies and persistent cookies.

Session cookies are created when a user enters a webpage, and are deleted as soon as the user exits. An example of usage of such cookies is in online shopping websites, where the "cart" of the user is stored as long as the user is in the site itself, once the user exits, all that they had stored in their shopping cart will be lost, as the session cookie storing that data is lost.

Persistent cookies are not deleted once the user exits the webpage, they are permanent. However, such cookies have expiration times/dates, beyond which they will be deleted. These are used to store users' personal data and allow for automatic logins for the user on certain websites. They may also be used to store the user's preferences and such.

## 4.2 Digital currency

A digital currency is that which exists electronically and not physically. An example of such is Bitcoin, which is a form of cryptocurrency. Banks are centralised authorities, in that there is an authority at the centre of it all that oversees all transactions. Cryptocurrencies are decentralised, in that there is no central authority. Transactions of cryptocurrencies are kept track of using a system called blockchain. A blockchain is essentially a list of all records of transactions. These are called digital ledgers. With each transaction, a record is appended to the blockchain, including a digital signature with time and date of transaction, once this data has been entered into the blockchain, it cannot be tampered with. If one wishes to view the data, they may, but the data that has been entered into the blockchain has been encrypted, making it very difficult to change the data.

## 4.3 Cyber security

When browsing the internet, users are subject to potential perpetrators, in that they may intercept and eavesdrop on users' data exchanges. Perpetrators use many strategies to do so. One such strategy is a brute force attack.

A brute force attack, in general, means trying every possible combination of something until it works. This is mostly done for getting the user's password, where every possible combination of every letter, number and symbol is tried until the correct permutation is stumbled upon.

Perpetrators may intercept users' data by use of packet sniffers, installed onto routers through which data packets pass. Packet sniffers are software which examine data packets which pass by to check for any possibly useful data. If contained data seems useful, this packet will be sent back to the perpetrator.

Using a botnet, perpetrators may attempt a distributed denial of service (DDoS) attack. This is achieved by sending malware (bad software, literally) to many devices to use those devices as "bots". Each bot will then target a certain web server and send to it multiple requests, simultaneously. A web server can

only handle so many requests, as a result the web server crashes. Such attacks can also be done on other pieces of network hardware other than web servers.

Hacking is the act of attempting unauthorised access to data. The perpetrator who does this is called a hacker. They do so by means of brute force attacks, exploiting vulnerabilities in systems or networks etc.

Malware is any form of malicious software designed to disrupt user's computer or data. Examples are: virus, worm, trojan horse, spyware, adware and ransomware.

- Virus: A computer program installed into unassuming user's hard drive. Such software replicates itself, causing the user's device to slow down. It may corrupt the user's files and use up all of the user's primary memory, resulting in system crash.
- Worm: Another self-replicating software, however such software is meant to find and exploit vulnerabilities in a network, replicating itself in those "holes". This uses up the network bandwidth, clogging and slowing down the whole network.
- Spyware: Software installed onto user's device for the purpose of spying on the user's actions is called spyware. An example is a keylogger, which records all key presses on user's keyboard. This may result in the perpetrator gaining access to victim's passwords.
- Trojan horse: A trojan horse is meant to be a disguise for other malware. The trojan horse looks to be like any other game or such, yet once it has been run, it releases malware into your system.
- Adware: Software that creates pop ups and banner advertisements on user's screen when they are online. These may be annoying, and these are done because the makers of such software are paid by whatever company they are advertising.
- Randsomware: This software encrypts the user's data and restricts the user's access to it, a sum of money is demanded from the user in exchange for their data. They sometimes threaten public display of user's data if money is not duly paid.

Pharming is done by perpetrators to gain access to user's personal data. Perpetrators install malware into user's machine, and whenever they visit certain websites, they are sent to another website, which, visually looks identical to that which they wanted to go to. Here, they put in their login details, which the perpetrator now knows and can access.

Phishing is similar to pharming, in that the user is sent to a fake website which mimics a real one. However, in phishing the user is sent an email, which contains a fake link, which mimics a real website. The rest of the process is identical to pharming.

Social engineering consists of people being decieved into providing perpetrators their personal data.

Such threats can be undone by some strategies.

Users may utilise access levels, which restrict user's access to certain files on a device depending on the user's "class". An administrator may have access to any and all files on a system, a user of that system may not have access to all of it.

Anti malware is a kind of software which looks for patterns of actions used by malware, if any such applications exist on the user's system, it is reported to the user and action is taken. Anti viruses are such software which locate, isolate and confirm whether the software is malicious. Anti spyware uses the same strategy, but specifically for spyware.

Authentication steps can be taken by the user. The user may set their password consisting of seemingly nonsense, using symbols letters, uppercase lowercase and numbers, making it harder for perpetrators to guess it. An example may be:

```
Kjalf89&(&!)YQOEnv,
```

The user may also set up two-step verification, where, during login, the user is asked to do some thing which only they can do. They may have to enter a code which was sent to a device which only they have access to.

The user may also set up biometric passwords, where biological data unique to the user only must be input. These require biometric devices which can scan, for example, the user's fingerprint, retina etc.

It has been seen that older software tend to have more vulnerabilities which perpetrators can exploit. Hence, user's should set up systems to automatically update their software to prevent being victims of such exploitation.

To avoid being phished or pharmed, users must check spelling and tone of emails, emails with typos and strange grammer and tone tend to be malicious.

Attached to links, the URL can be checked by hovering one's cursor over the link. The URL must be authentic, for the user to confidently click on it.

Firewalls scan incoming and outgoing data from a user's system, and criteria can be set for the firewall to examine said data. This prevents the user from downloading malicious software onto their machine.

Privacy settings can be set by the user on certain accounts to not have their information be sold out to companies.

Proxy servers can be used by web servers to act as a barrier to the web server. The proxy server examines each request before passing it through to the web server. A proxy server prevents multiple requests from the same IP, and when the server is bombarded with requests, it passes the requests through at a slower rate to prevent crashing.

The HTTPS protocol uses either the secure sockets layer (SSL) or transport layer security (TLS) to determine the security of websites. Checking a website's URL before accessing it can be useful in that, one can check whether the website is secured. This usually boils down to checking whether the website URL has a padlock beside the address bar and that the protocol is HTTP or HTTPS.