

Markscheme

Mock examination — Set 1

Computer science

Standard level

Paper 1

This markscheme is **confidential** and for the exclusive use of the teachers marking this mock examination.

It must **not** be shared with candidates before the mock examination has been sat.

General marking instructions

1. Follow the markscheme provided, award only whole marks and mark only in **RED**.
2. Where a mark is awarded, a tick/check (✓) **must** be placed in the text at the **precise point** where it becomes clear that the candidate deserves the mark. **One tick to be shown for each mark awarded.**
3. Sometimes, careful consideration is required to decide whether or not to award a mark. In these cases write a brief comment in the margin to support your decision, especially for moderation purposes. It should be remembered that the script may be returned to the candidate.
4. Personal codes/notations are unacceptable.
5. Where an answer to a part question is worth no marks but the candidate has attempted the part question, write a zero next to the answer. Where the candidate has not attempted the part question, write “NR”.
6. Ensure that you have viewed **every** page including any additional sheets.
7. A mark should not be awarded where there is contradiction within an answer. Write “CON” next to the contradiction.

Subject details: Computer Science SL Paper 1 markscheme (Mock examination — Set 1)

Mark allocation

Section A: Candidates are required to answer **all** questions. Total 38 marks.

Section B: Candidates are required to answer **all** questions. Total 12 marks.

Maximum total = 50 marks.

General

A markscheme often has more specific points worthy of a mark than the total allows. This is intentional. Do not award more than the maximum marks allowed for that part of a question.

When deciding upon alternative answers by candidates to those given in the markscheme, consider the following points:

- Each statement worth one point has a separate line and the end is signified by means of a semi-colon (;).
- An alternative answer or wording is indicated in the markscheme by a “/”; either wording can be accepted.
- Words in (...) in the markscheme are not necessary to gain the mark.
- If the candidate’s answer has the same meaning or can be clearly interpreted as being the same as that in the markscheme then award the mark.
- Mark positively. Give candidates credit for what they have achieved and for what they have got correct, rather than penalizing them for what they have not achieved or what they have got wrong.
- Remember that many candidates are writing in a second language; be forgiving of minor linguistic slips. In this subject effective communication is more important than grammatical accuracy.
- Occasionally, a part of a question may require a calculation whose answer is required for subsequent parts. If an error is made in the first part then it should be penalized. However, if the incorrect answer is used correctly in subsequent parts then **follow through** marks should be awarded. Indicate this with “FT”.

General guidance

Issue	Guidance
Answering more than the quantity of responses prescribed in the questions	• In the case of an “identify” question, read all answers and mark positively up to the maximum marks. Disregard incorrect answers. • In the case of a “describe” question, which asks for a certain number of facts <i>eg</i> “describe two kinds”, mark the first two correct answers. This could include two descriptions, one description and one identification, or two identifications. • In the case of an “explain” question, which asks for a specified number of explanations <i>eg</i> “explain two reasons ...”, mark the first two correct answers. This could include two full explanations, one explanation, one partial explanation <i>etc.</i>

Section A

1. (a) Award **[1]** max

Random-access memory (RAM);
Read-only memory (ROM);
Registers;

Note: Do not accept secondary storage (e.g. SSD, HDD, flash drive).

(b) Award **[2]** max

*Award **[1]** per point up to **[2]** max*

The data/instruction requested by the CPU is not found in the cache;
...so it must be fetched from a slower level of memory (L2/L3 cache/RAM);
...which takes longer than a cache hit/the CPU may have to wait (stall);

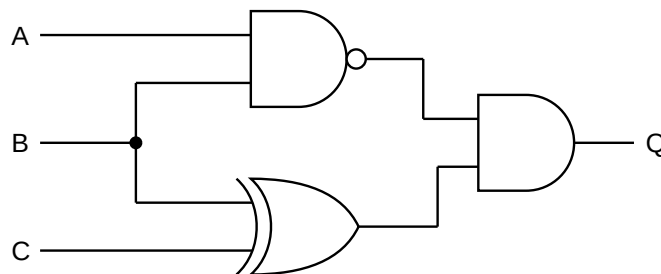
A copy of the data (and nearby data) is then placed in the cache;
...so that later requests for the same data result in a cache hit;

(c) Award **[3]** max

*Award **[1]** for a NAND gate with inputs A and B;*

*Award **[1]** for an XOR gate with inputs B and C;*

*Award **[1]** for an AND gate whose inputs are the outputs of the NAND gate and the XOR gate, giving output Q;*



*Note: Accept an AND gate followed by a NOT gate in place of the NAND gate. Accept an equivalent circuit of AND, OR and NOT gates in place of the XOR gate. Input B may be drawn as one line that branches to both gates. Award **[2]** max if correct gates are drawn with non-standard symbols or without input/output labels (A, B, C, Q).*

(d) Award **[1]** max

16 777 216/2²⁴;

Note: Accept “about 16.7 million”.

(e) Award **[2]** max

*Award **[1]** for identifying an advantage and **[1]** for expansion*

Files can be accessed from any workstation on the network;
...so several artists can work on the same project/shared assets without copying files between machines;

There is a single, up-to-date copy of each file;
...reducing the risk of artists working on out-of-date versions;

Data is better protected;
...a NAS often uses redundant drives (RAID)/central backup, so files are not lost if one drive fails;

Storage capacity can be increased easily;
...by adding drives to the NAS rather than upgrading every workstation;

Access can be controlled centrally;
...permissions can be set so that only authorized staff can open each project;

(f) *Award [2] max*

Award [1] per point up to [2] max

Each process is given a fixed time slice/quantum of CPU time in turn;
...when its time slice ends, the export is pre-empted and moved to the back of the queue;
...so the long export cannot monopolize the CPU;
...the drawing software receives the CPU frequently/after a short wait, so the animator notices no delay;

2. (a) Award [2] max

Award [1] per point up to [2] max

Reliability: if a link between two floor switches fails, data can travel along another path;
A fault in one cable/workstation only affects that workstation;
Scalability: new workstations can be added to a floor switch without disrupting the rest of the network;
Fewer data collisions: each workstation has its own link to a switch;
Transmission speed: the extra links between floor switches can share the traffic load (e.g. with layer 3 switching/link aggregation);
Cost: cheaper than connecting every device in a full mesh (only the four switches are meshed);
Easier fault finding/management: a problem can be isolated to one floor;

Note: Accept other reasonable advantages linked to reliability, transmission speed, scalability, data collisions or cost.

(b) Award [2] max

Award [1] per point up to [2] max

TCP is connection-oriented/reliable;
...it checks that every segment arrives (acknowledgements) and resends any that are lost or corrupted;
...it uses sequence numbers so that the data is reassembled in the correct order;
A single missing or incorrect part would corrupt the frame file/make the image unusable;
UDP does not guarantee delivery or order, so parts of a file could be lost;
For a file transfer, accuracy matters more than speed/low latency;

(c) Award [3] max

Award [1] per point up to [3] max

The website sends its digital certificate to the browser;
The certificate contains the website's public key;
The browser checks that the certificate was signed by a trusted certificate authority (CA)/is valid and has not expired;
...which confirms that the public key belongs to the studio's website;
The browser generates a (random) symmetric session key;
...and encrypts it with the website's public key before sending it;
Only the website's private key can decrypt the session key;
...so an attacker who intercepts it cannot read it;
Both sides then use the session key to encrypt the data (because symmetric encryption is faster);

Alternative answer (key exchange, e.g. TLS 1.3):

The browser and the website exchange (Diffie–Hellman) key shares;
The website signs its key share with the private key that matches its certificate;
...the browser checks the signature using the public key in the CA-signed certificate, so the share must have come from the studio's website;
Both sides combine their own secret value with the other's key share to compute the same session key;
...the session key itself is never transmitted, so it cannot be intercepted;

Note: Award [2] max if the response does not describe how the session key itself is protected (certificate-only answers). Points may be taken from either answer.

(d) Award [4] max

Award [1] per point up to [4] max

The firewall filters packets using rules about source/destination IP addresses, ports and protocols;

...it does not normally examine the content of the data, such as an attachment;

Email traffic comes from a trusted/whitelisted mail server on an allowed port/protocol;

...so the packets carrying the attachment matched an “allow” rule and passed through;

The email may have been encrypted, so the firewall could not inspect its contents;

The malware was activated by an authorized user’s action inside the network;

...a firewall cannot stop users being tricked (e.g. by a convincing email);

A blacklist only blocks known malicious addresses/domains, so an email from a new or unlisted sender is not blocked;

3. (a) Award **[2]** max

Award **[1]** per field

Frames: integer/INT;

DayRate: decimal/DECIMAL(6,2)/NUMERIC/real/float/currency/money;

Note: Do not accept “number” on its own. Do not accept integer for DayRate.

(b) Award **[3]** max

Award **[1]** for selecting ShotID, DueDate and LastName from SHOT and ANIMATOR with a correct join condition (SHOT.AnimatorID = ANIMATOR.AnimatorID)

Award **[1]** for BETWEEN '2027-03-01' AND '2027-03-31' or an equivalent condition

Award **[1]** for ORDER BY DueDate (ASC)

```
SELECT SHOT.ShotID, SHOT.DueDate, ANIMATOR.LastName
FROM SHOT
INNER JOIN ANIMATOR ON SHOT.AnimatorID = ANIMATOR.AnimatorID
WHERE SHOT.DueDate BETWEEN '2027-03-01' AND '2027-03-31'
ORDER BY SHOT.DueDate ASC;
```

Alternative solution using an implicit join:

```
SELECT ShotID, DueDate, LastName
FROM SHOT, ANIMATOR
WHERE SHOT.AnimatorID = ANIMATOR.AnimatorID
AND DueDate >= '2027-03-01' AND DueDate <= '2027-03-31'
ORDER BY DueDate;
```

Note: Accept any date format, either quote style and LIKE '2027-03%' for the date condition. Ignore minor syntax errors. Table prefixes are only required where a field name is ambiguous. Do not award the first mark if both tables are listed without a join condition (this returns every combination of shot and animator).

(c) Award **[2]** max

Award **[1]** for UPDATE SHOT SET AnimatorID = 'A05'

Award **[1]** for a WHERE clause that selects only the shots of A03 that are not approved

```
UPDATE SHOT
SET AnimatorID = 'A05'
WHERE AnimatorID = 'A03' AND Status <> 'Approved';
```

*Note: Accept NOT Status = 'Approved', Status != 'Approved' or Status IN ('In progress', 'Not started'). Do not award the WHERE mark for mixed AND/OR without brackets (e.g. AnimatorID = 'A03' AND Status = 'In progress' OR Status = 'Not started', which also reassigns S105). Award **[1]** max for statements that name the shots directly, e.g. WHERE ShotID IN ('S102', 'S106') or separate statements for S102 and S106.*

(d) Award **[2]** max

Award **[1]** per point up to **[2]** max

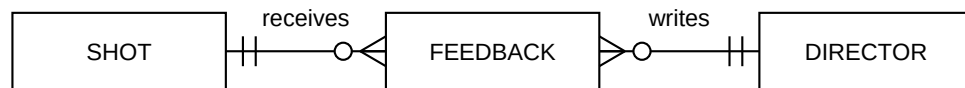
ReleaseDate depends on Film, which is not the primary key/is a non-key attribute;
...this is a transitive dependency ($\text{ShotID} \rightarrow \text{Film} \rightarrow \text{ReleaseDate}$);
3NF does not allow a non-key attribute to depend on another non-key attribute;
The release date would be repeated for every shot of the same film;
...which could lead to inconsistent data/update anomalies if a release date changes;

(e) Award **[3]** max

Award **[1]** for a one-to-many relationship between SHOT and FEEDBACK;

Award **[1]** for a one-to-many relationship between DIRECTOR and FEEDBACK;

Award **[1]** for correct modality: optional (zero) at the FEEDBACK end of both the SHOT–FEEDBACK and DIRECTOR–FEEDBACK relationships, and mandatory one at the SHOT and DIRECTOR ends;



Note: Any recognized ERD notation is acceptable (e.g. crow's foot, Chen, or 1 and M labels with O for optional); relationship names are not required. Award **[1]** max if SHOT and DIRECTOR are joined by a many-to-many relationship without the FEEDBACK entity.

4. (a) *Award [2] max*

Award [1] per point up to [2] max

The 20 000 hand-labelled shots provide labelled training data (inputs with known correct outputs);
...the model learns the patterns that link the features of the images to the character labels;
...it can then predict/classify which characters appear in the unlabelled shots;
The categories (the characters) are known in advance, so this is a classification task;
Some labelled shots can be kept back to test the model's accuracy before it is used;

(b) *Award [2] max*

Award [1] for identifying an issue and [1] for expansion

Privacy;
...detailed data about each animator's working activity is collected and stored;
Consent;
...animators may not have agreed to be monitored/may feel unable to refuse their employer;
Fairness/algorithmic bias;
...edits per hour may not reflect the quality of creative work, so some animators may be unfairly predicted to miss deadlines/lose bonuses;
Transparency;
...animators may not know how the predictions are made, so they cannot challenge them;
Accountability;
...it is unclear who is responsible if a wrong prediction leads to an unfair decision;
Well-being/societal impact;
...constant monitoring may cause stress and reduce trust between staff and managers;
Security;
...the stored monitoring data could be accessed by unauthorized people;

Section B

5. (a) *Award [2] max*

Award [1] for identifying role and [1] for expansion

At each step, the CNN takes the current noisy image (and the step number/prompt) as input;
...and predicts the noise in the image/produces a slightly less noisy version of the image;

The predicted noise is removed/subtracted from the image;
...and the result is passed back to the CNN, repeating many times until a clear/coherent image is produced;

The CNN uses features/patterns learned from the training dataset;
...to gradually reconstruct edges/shapes/textures so that the final image resembles the training data/matches the prompt;

Generation begins with noise injection/an image of random noise;
...which the CNN refines over many steps into a coherent image;

The CNN is trained by gradually adding known amounts of noise to training images (as in the DDPM);
...and learning to predict/remove that noise, so that it can reverse the process when generating a new image;

Note: Answers that only state that the CNN “removes noise” should be awarded [1] max.

(b) *Award [4] max*

Award [1] for identifying a way and [1] for expansion

Use only licensed images;
...e.g. buy licences from stock image libraries/commission artists to create training images;

Use images that Visionary Studios already owns, or that are in the public domain/openly licensed;
...so that permission to use them for training is already in place;

Respect opt-outs from creators;
...exclude images from websites or artists that have indicated they must not be used for AI training;

Record the source and licence of every image (e.g. in a database of metadata);
...so that compliance can be checked/audited and images removed if a claim is made;

Filter the collected images automatically;
...e.g. remove images with watermarks, logos or copyrighted characters, or that match known copyrighted works;

Take legal advice on copyright law in the countries where the model is trained and used;
...because exceptions for training (e.g. text and data mining) differ between countries;

Mark as [2] + [2]

(c) Award **[6]** max

Responses may include some of the following points. The best ones will demonstrate research and give a balanced perspective before presenting a conclusion. Named examples, figures and cases below are illustrative only; candidates are not required to know them.

Benefits of training its own diffusion model

- Dataset curation: Visionary Studios controls which images are used, so it can use only licensed, public domain or its own images and keep records of their sources, reducing the risk of infringing intellectual property.
- Bias mitigation: The studio can audit its dataset for under-represented groups and rebalance it before training, rather than relying on an external provider's unknown dataset.
- Style adherence and consistency: A model trained on the studio's own artwork can match its brand identity and keep characters consistent across campaigns.
- Confidentiality: Client briefs and unreleased designs do not have to be sent to an external service.
- Optimization: The studio can choose the number of denoising steps and the image resolution to balance quality against computational demands for each project.
- Transparency: Knowing exactly what the model was trained on makes it easier to disclose the use of AI to clients and audiences.

Challenges of training its own diffusion model

- Computational demands: Training requires many GPUs/TPUs running for weeks; for example, the model card for the original Stable Diffusion model reports about 150 000 GPU-hours of training. Candidates should be aware that the studio's current infrastructure may not be able to support this.
- Generation costs: Each image still requires many iterative denoising steps, so powerful GPUs are needed to deploy the model as well as to train it.
- Dataset size: Diffusion models are usually trained on hundreds of millions of image-text pairs (e.g. subsets of the LAION-5B dataset). A small, licensed dataset may produce lower-quality images or lead the model to memorize and reproduce its training images.
- Cost and expertise: Specialist staff, maintenance, energy costs and environmental impact (high energy use and carbon emissions).
- Remaining bias: A curated dataset can still reflect the biases of the people who selected and labelled the images.

Using an existing text-to-image generation service

- Immediate access to high-quality models without investing in infrastructure; the provider manages the computational demands (e.g. through cloud-based platforms/software as a service).
- The training data of many services is not disclosed, so the studio may be exposed to IP risks. Legal actions such as Getty Images' case against Stability AI and lawsuits brought by artists show both the risk and the legal uncertainty; credit candidates who note the outcome of the case (in November 2023 the UK High Court dismissed Getty's remaining copyright claim, with only limited trade mark findings).
- Some providers train only on licensed images and offer legal protection (indemnity) to business customers (e.g. Adobe Firefly), which reduces the IP risk.
- Bias in the provider's model can only be partly controlled, e.g. through careful prompts and human review of outputs; analyses of image generators (e.g. Bloomberg, 2023) found that generated images amplified gender and racial stereotypes.

- Less control over style, consistency and future changes to the service, and ongoing per-image or subscription costs.

Compromise

- Fine-tuning an existing open-source diffusion model on the studio's own licensed images (transfer learning) needs far less computation than training from scratch, but IP and bias issues in the original training data remain.

Conclusion

A reasoned conclusion that expresses an opinion and is supported by evidence.

Marks	Level descriptor
No marks	• No knowledge or understanding of the relevant issues and concepts. • No use of appropriate terminology.
Basic 1–2 marks	• Minimal knowledge and understanding of the relevant issues. • Minimal use of appropriate terminology. • The answer may be little more than a list. • No reference is made to the case study or independent research.
Adequate 3–4 marks	• A descriptive response with limited knowledge and/or understanding. • A limited use of appropriate terminology. • There is some analysis. • There is evidence of some research.
Competent 5–6 marks	• Knowledge and understanding of the related issues and/or concepts. • Uses terminology appropriately in places. • There is evidence of analysis. • There is evidence of research. • There is a conclusion.
