

Markscheme

Mock examination — Set 2

Computer science

Higher 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 HL Paper 1 markscheme (Mock examination — Set 2)

Mark allocation

Section A: Candidates are required to answer **all** questions. Total 56 marks.

Section B: Candidates are required to answer **all** questions. Total 24 marks.

Maximum total = 80 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 eg “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 eg “explain two reasons ...”, mark the first two correct answers. This could include two full explanations, one explanation, one partial explanation etc.

Section A

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

Decodes each instruction (held in the instruction register/IR);
Sends control signals to coordinate the other components (e.g. the ALU, registers and memory);
Controls/manages the fetch–decode–execute cycle (e.g. using the clock to time each step);
Controls the flow of data between the CPU, memory and input/output devices (along the buses);

Note: Do not accept “carries out calculations” or “performs logical operations” (these are functions of the ALU).

(b) (i) Award **[1]** max

180;

Note: $B4 = 11 \times 16 + 4 = 180$. Do not accept 60 (the red intensity) or 230 (the blue intensity).

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

C8B4E6;

*Note: Accept lower case (c8b4e6) and a leading # (#C8B4E6). Award **[0]** if the green or blue digits have been changed.*

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

Award **[1]** for each correct column (P, R and Q)

A	B	C	P	R	Q
0	0	0	1	0	0
0	0	1	1	1	1
0	1	0	0	1	1
0	1	1	0	0	0
1	0	0	0	1	1
1	0	1	0	0	1
1	1	0	0	1	1
1	1	1	0	0	1

*Note: The eight combinations of A, B and C may be listed in any order, but each must appear exactly once; award **[2]** max if a combination is missing or repeated. Follow through: award the R mark for an R column that is correct for the candidate’s P column, and the Q mark for a Q column that is correct for the candidate’s R column.*

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

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

The OS uses a device driver for each make/model of tablet;
...the driver translates between the tablet's own signals/commands and a standard form used by the OS;
The OS gives every application the same standard interface for tablet input (e.g. pen position and pressure), whichever tablet is connected (device independence);
...so a drawing application does not need to know the details of each manufacturer's hardware (abstraction);
The OS handles the input/interrupts from the tablet and passes the input to the application that is currently being used;

*Note: Award **[1]** max if the answer does not explain how the OS hides the differences between the tablets (e.g. through device drivers or a standard interface).*

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

*Award **[1]** for each value*

Without pipelining: $10 \times 4 = 40$ (clock cycles);
With pipelining: $4 + (10 - 1) = 13$ (clock cycles);

Note: Accept a correct timing diagram showing that the tenth instruction finishes in cycle 13. Do not award the second mark for 14 ($4 + 10$) or 10.

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

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

The CPU runs the main program/controls the task, e.g. reading and decoding the video file and managing the timeline and user interface;
...these tasks involve sequential steps and decisions (branching), which suit the CPU's few, powerful cores;
The CPU sends the frame data and the instructions for the effect to the GPU (copying the frames into the GPU's memory, unless the CPU and GPU share unified memory);
The GPU has thousands of simpler cores that carry out the same operation on many pixels at the same time (in parallel);
...each pixel can be processed independently, so millions of pixels in a frame can be changed at once;
...the GPU's high-bandwidth memory allows the large frames to be read and written quickly;
The processed frames are returned to the CPU/main memory, which then saves/encodes the video;
The CPU can prepare the next frames while the GPU processes the current ones, so both work at the same time;

*Note: Award **[2]** max if the answer only describes the architecture of the CPU and the GPU without explaining which parts of this task each one carries out.*

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

Internet layer;

Note: Accept network layer. Do not accept network interface layer, transport layer or application layer.

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

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

Each file/message is divided into small packets;

Each packet has a header containing the source and destination (IP) addresses and a sequence number;

The connection is not reserved for one transfer (as it would be with circuit switching);

...packets from different computers/transfers are interleaved and sent one after another along the shared connection;

...so a large video upload does not block the connection, because packets for web pages and emails can be sent between its packets;

Routers forward each packet independently, using the destination address in its header;

Packets arriving for the agency are passed to the correct computer using the addresses (and port numbers) in their headers (e.g. by NAT on the router);

At the destination, the packets of each transfer are put back in order using their sequence numbers;

*Note: Award **[2]** max if the answer only describes how a file is split into packets and reassembled, without explaining how the single connection is shared.*

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

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

In the client–server system, every billboard downloads the whole file from the one web server, so the server must send 60 copies (about 240 GB) over the agency's single connection;

...the server's upload bandwidth is shared between all the downloads, so it is a bottleneck and each billboard receives the advert slowly;

In the peer-to-peer system, the file is divided into pieces and each billboard can download different pieces from several other billboards at the same time;

...every billboard that has received a piece also uploads it, so the total upload capacity grows as more billboards take part (scalability);

...the agency's server needs to send each piece only about once, so its connection is no longer the bottleneck;

A slow or unavailable billboard does not hold up the others, because each piece can be obtained from several other billboards;

*Note: Award **[2]** max if the answer does not explain both why delivery from the single web server is limited and how the peer-to-peer system avoids this limit.*

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

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

The attacker posts a comment that contains a malicious script (e.g. JavaScript) instead of ordinary text;

The web server stores the comment and includes it in the page without checking/filtering or encoding it;

When other visitors open the page, their browsers run the script as if it were part of the trusted website;

...the script can then, for example, steal the visitors' session cookies/login details, redirect them to a fake site or change what the page shows;

(e) *Award [2] max*

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

The web server sends the CSP to the browser with each page (e.g. in an HTTP header);

The CSP tells the browser the sources (domains) from which scripts and other resources may be loaded;

The browser refuses to run scripts that do not come from an allowed source/(by default) scripts written directly into the page;

...so a script injected into a comment is not run in the visitors' browsers;

The CSP can also limit where the page can send data, so stolen data cannot easily be sent to the attacker;

Note: Award [1] max if the answer only describes what a CSP is, without stating that a script injected into a comment would not be run. Do not credit descriptions of input validation or of filtering the comments; the question asks about a CSP.

(f) *Award [3] max*

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

The web server accepts connections from anyone on the internet, so it is the device most likely to be attacked/compromised;

If an attacker takes control of the web server, the firewall blocks/filters traffic from the web server's segment to the staff segment;

...so the attacker cannot easily spread (e.g. malware) to the staff computers or reach the client files on the file server;

The damage from a successful attack is limited to one segment/the rest of the network keeps working;

The traffic in each segment can be monitored separately, so unusual traffic from the web server is easier to detect;

Note: Credit answers that call the separate segment a demilitarized zone (DMZ) or that refer to subnets or VLANs. Award [2] max if the answer does not explain what happens if the web server is compromised.

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

Award **[1]** for the UPDATE statement and **[1]** for the ALTER TABLE statement

The UPDATE statement is DML/it changes the data in an existing record (the budget of P103), so it is allowed and is carried out;

The ALTER TABLE statement is DDL/it changes the structure of the table (it adds a new field), so it is refused;

Note: Do not award a mark for stating which statement is carried out unless the answer classifies it as DML or DDL or explains that it changes the data or the structure.

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

Award **[1]** for selecting Title, Budget and ClientName from CAMPAIGN and CLIENT with a correct join condition (CAMPAIGN.ClientID = CLIENT.ClientID)

Award **[1]** for WHERE Sector = 'Retail' AND Channel <> 'Social' (or NOT Channel = 'Social')

Award **[1]** for ORDER BY Budget DESC

```
SELECT CAMPAIGN.Title, CAMPAIGN.Budget, CLIENT.ClientName
FROM CAMPAIGN
INNER JOIN CLIENT ON CAMPAIGN.ClientID = CLIENT.ClientID
WHERE CLIENT.Sector = 'Retail' AND CAMPAIGN.Channel <> 'Social'
ORDER BY CAMPAIGN.Budget DESC;
```

Alternative solution using an implicit join:

```
SELECT Title, Budget, ClientName
FROM CAMPAIGN, CLIENT
WHERE CAMPAIGN.ClientID = CLIENT.ClientID
AND Sector = 'Retail' AND NOT Channel = 'Social'
ORDER BY Budget DESC;
```

Note: Accept either quote style, the string literals in any letter case, the fields in any order, and != in place of <>. Ignore minor syntax errors. Table prefixes are only required where a field name is ambiguous (ClientID). Accept other correct ways of combining the tables, e.g. JOIN ... USING (ClientID), or a subquery on CLIENT. Do not award the first mark if both tables are listed without a join condition (this returns every combination of campaign and client). Do not award the second mark if OR is used instead of AND, if either condition is missing, if Channel = 'Video' is used in place of the Social condition (this misses Ride to Work), or if hard-coded IDs (e.g. ClientID IN ('C02', 'C04')) are used in place of Sector = 'Retail'.

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

Award **[1]** for an advantage and **[1]** for a disadvantage

Advantage: the results of the join are stored, so the view opens more quickly/the query does not have to be run every time the view is used;

Advantage: there is less processing load on the database when many account managers use the view at the same time;

Disadvantage: the stored results can become out of date until the view is refreshed (e.g. a changed budget or a new campaign does not appear immediately);

Disadvantage: the stored results take up extra storage space;

Disadvantage: refreshing the view takes processing time, especially if the tables change often;

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

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

Without isolation, both applications could read 31 500 and both write 36 500, so one increase is lost (lost update);

Isolation means that transactions running at the same time give the same result as if they had been carried out one after the other;

...e.g. the DBMS locks the P107 record when the first transaction reads/updates it, so the second transaction must wait until the first has committed;

...so the second transaction reads 36 500 and writes 41 500;

Note: Award **[1]** max if the answer only defines isolation without applying it to the two updates of P107.

(e) Award **[5]** max

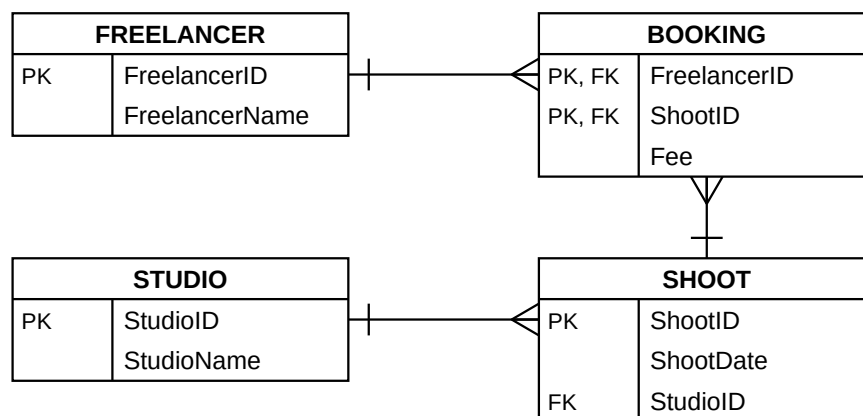
Award **[1]** for a *FREELANCER* table with *FreelancerID* as its primary key and *FreelancerName*;

Award **[1]** for a *STUDIO* table with *StudioID* as its primary key and *StudioName*;

Award **[1]** for a *SHOOT* table with *ShootID* as its primary key, *ShootDate*, and *StudioID* as a foreign key (and no *StudioName*);

Award **[1]** for a *BOOKING* table that contains *FreelancerID* and *ShootID* as foreign keys, and *Fee*;

Award **[1]** for (*FreelancerID*, *ShootID*) as the composite primary key of the *BOOKING* table;



Note: Accept any clear way of showing keys (e.g. underlining, PK/FK labels or asterisks) and any sensible table names; relationship lines are not required. For the fifth mark, accept a surrogate primary key (e.g. *BookingID*) only if *FreelancerID* and *ShootID* are both included as foreign keys. Do not award the first mark for a *FREELANCER* table that contains *Fee*, the third mark for a *SHOOT* table that contains *StudioName*, or the fourth mark for a *BOOKING* table that contains *FreelancerName*, *ShootDate*, *StudioID* or *StudioName*. Award **[3]** max if *StudioName* has not been separated from *SHOOT*.

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

Policy;

Note: Do not accept “reward”, “value” or “strategy” on its own.

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

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

Each time a page is loaded, the agent observes the state (e.g. the page/time of day), chooses one version to show (action) and receives a reward (1 or 0);
It uses the rewards to update its estimate of the value/expected reward of each version (e.g. its click rate so far);
Its policy is updated so that versions with a higher estimated reward are chosen more often;
...so that over many page loads it maximizes the cumulative reward (the total number of clicks);
It still sometimes shows other versions (exploration), so that it does not miss a better version, while mostly showing the best version so far (exploitation);

*Note: Award **[1]** max if the answer does not describe how the rewards change the agent’s future choices.*

(b) (i) Award **[1]** max

$1500 \div 2000 = 0.75/75\%$;

Note: The correct value earns the mark without working. Do not accept 3.75% (the support of the rule, $1500 \div 40\,000$) or 30% ($1500 \div 5000$).

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

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

A lift greater than 1 shows that buying a tent makes buying a sleeping bag more likely/the two items are bought together more often than if they were independent;
Customers who buy a tent are 6 times as likely to buy a sleeping bag as customers in general (75% compared with 12.5% of all orders);
...so the rule shows a strong association, not just that sleeping bags are popular;

*Note: Award **[1]** max if the answer does not interpret the value 6 (e.g. six times as likely as in orders in general).*

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

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

Clustering places visitors in the same group when the distance between their feature values is small/they are similar;
With 300 features, the distances between all pairs of visitors become very similar, so near and far visitors are hard to tell apart/the clusters are not well separated;
Many of the features are likely to be irrelevant or noisy, which hides the features that really distinguish groups of visitors;

The data is sparse (most visitors have no values for most pages/products), so there is too little data to find reliable groups;
With fewer, relevant features, the groups are better separated and easier for the agency to interpret and target;
Fewer features also reduce the processing time and memory needed;

Note: Award [1] max unless a reason is given why many features make the groups worse (e.g. distances become similar, irrelevant/noisy features, sparse data); faster processing or less memory is not such a reason.

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

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

Each convolutional layer slides small filters (kernels) across the image to produce feature maps;
The values (weights) in the kernels are learned during training, by adjusting them to reduce the loss/errors;
Early layers detect simple features such as edges, corners and colours;
Later layers combine these into more complex features, such as letter shapes and parts of the logo, and finally the whole logo (a spatial hierarchy);
The same kernel is applied at every position, so a feature is detected wherever it appears in the image;
Pooling layers reduce the size of the feature maps, so small changes in the logo's position or size do not matter;
The fully connected layer(s) use the high-level features to decide whether the logo is present;

Note: Award [1] max if the layers are only named, without describing how the features are built up.

(e) **Award [5] max**

The following points might be considered:

- Privacy: the model infers a sensitive fact (financial difficulty) from browsing history, location and online times; visitors have not chosen to share this information
- Informed consent: visitors to the partner websites are unlikely to know, or to have agreed, that their data will be used in this way; accepting cookies is not informed consent to financial profiling
- Exploitation and societal impact: showing loan adverts to people at a vulnerable time may lead them into more debt, harming individuals and increasing problem debt in society
- Algorithmic fairness and bias: location and online times can act as proxies for income, ethnicity or shift work, so some groups may be targeted more than others; people wrongly predicted to be in difficulty are also labelled unfairly
- Transparency: visitors cannot see why they are being shown the adverts or how the predictions are made
- Accountability: it is unclear who is responsible if people are harmed – the client, the agency, the partner websites or the developers of the model
- Security: data that identifies people in financial difficulty would be valuable to criminals, so a data breach could lead to fraud or scams
- Benefits: people who urgently need money may find a cheaper loan than the alternatives; the adverts are more relevant than untargeted adverts; the agency earns income and keeps its client

- Legal implications: data protection laws (e.g. GDPR) restrict profiling and require a lawful basis for processing; adverts for loans are regulated in many countries (e.g. required cost and risk warnings)
- Reputation: if the targeting becomes public, the agency and its other clients could be criticized
- Safeguards or alternatives: target the adverts by the content of the page rather than by predicted difficulty; exclude visitors predicted to be in severe difficulty; include clear cost warnings and links to free debt advice; human review of the campaign
- The need to reassess the agency's ethical guidelines regularly as targeting models become more powerful

Marks	Level descriptor
No marks	No knowledge or understanding of the relevant issues and concepts.
1–2 marks	Limited knowledge and understanding of targeting adverts at financially vulnerable people. The answer may be little more than a list of points. (General)
3–4 marks	Detailed knowledge that links targeting adverts at financially vulnerable people to the scenario; Fairly balanced analysis. (In context)
5 marks	An appropriate conclusion, linked to the analysis.

Section B

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

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

The existing model learned its patterns from the original (possibly copyrighted) photographs;

...so the synthetic images may closely reproduce parts of copyrighted works, e.g. characters, logos or an artist's style, especially if the model has memorized some training images;

The intellectual property problem is moved rather than removed;

...because the synthetic images are derived from the original photographs, the studio may still be infringing/may still need permission for the original use of the photographs to train the existing model;

The existing model reproduces the biases in the original photographs;

...so groups that were under-represented/stereotyped in the photographs are also under-represented/stereotyped in the synthetic dataset;

Bias may be amplified;

...because generative models tend to reproduce the most common patterns and lose rare ones (e.g. model collapse when models are trained on generated data), so under-represented groups appear even less often;

The synthetic dataset adds no new information about groups or situations that were missing from the original photographs;

...so the new model still cannot learn to represent them fairly/accurately;

Note: Answers that only state that synthetic images are “not real” or “may be low quality”, without linking this to intellectual property or bias, should be awarded [0].

Note: Accept other reasonable reasons, e.g. the origin of each synthetic image cannot be traced, so the studio cannot show clients which works its model learned from.

(ii) *Award [2] max*

Award [1] for why more steps take longer and [1] for why more steps give finer detail

Each step is a full pass of the whole image through the denoiser (CNN), involving a very large number of calculations, so 50 steps need about five times as much computation as 10 steps;

Each step uses the result of the previous step, so the steps must be carried out one after another and cannot be carried out in parallel to save time;

Each step removes only part of the noise, so with more steps the image is refined more gradually and fine features such as edges and textures can be reconstructed;

With fewer steps, more noise must be removed at each step, so the denoiser's predictions are less accurate, giving blurred/less detailed images or artefacts;

The denoiser is trained to remove noise gradually (as formalized in the DDPM), so more steps follow this gradual process more closely;

Note: Do not award the first mark for only stating that 50 steps is more than 10 steps, or the second mark for only stating that more steps give “better quality”.

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

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

The VAE's encoder maps the approved design to a point in the latent space;
...points close to it are sampled/the point is moved slightly and then decoded, giving new images that are similar to the approved design;

In the latent space, similar images are placed close together;
...so small changes give controlled variations (e.g. of colour or pose) rather than unrelated images;

The studio can interpolate between the latent points of two approved designs;
...and decode the points in between to give a series of images that blend features of both;

Particular directions in the latent space can correspond to attributes;
...so one attribute (e.g. lighting or expression) can be changed while the rest of the design is kept;

Note: Do not award marks for only repeating that VAEs are useful for flexibility or experimentation, without outlining how.

(ii) *Award [2] max*

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

The flow-based model learns an exact, reversible mapping between random noise and images;

...so the final image can be mapped back to the exact noise input that produced it, showing how it was formed;

The mapping can be reversed without loss/is deterministic;
...so by storing the noise input (and settings) the studio can regenerate exactly the same image to prove how it was produced;

Each transformation between the noise and the image can be traced/recorded;
...giving a transparent record for the client or an audit, which supports AI disclosure;

Note: Do not award marks for only repeating that flow-based models are useful for transparency or reproducibility, without outlining how.

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

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

The generator learns only from the discriminator's feedback/its weights are updated using the gradient of the discriminator's output for the generated images;
When the discriminator is (almost) certain that every generated image is fake, its output saturates, so the gradient passed back to the generator is (almost) zero/vanishes;
The generator therefore receives no useful information about which changes would make its images look more real, so its weights hardly change;
The discriminator gives the same verdict to slightly improved images as to poor ones, so improvements are not rewarded;

Note: Do not award marks for only restating the stimulus (e.g. “the discriminator is too strong”).

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

*Award **[1]** for identifying a way and **[1]** for outlining how it restores the balance*

Lower the discriminator’s learning rate/update the discriminator less often than the generator;

...so the discriminator does not get too far ahead and the generator can catch up;

Pause training of the discriminator when its loss falls below a threshold;

...so only the generator is updated until the discriminator’s feedback becomes useful again;

Reduce the discriminator’s capacity, e.g. fewer layers/neurons, or add dropout;

...so it cannot separate real and fake images perfectly/memorize the real images, and its feedback stays informative;

Add random noise to the images given to the discriminator/use (one-sided) label smoothing, e.g. a target of 0.9 instead of 1 for real images;

...so the discriminator is less confident and the gradients do not vanish;

Train the generator with a non-saturating loss (maximize the probability that the discriminator labels its images as real);

...which still gives strong gradients when the discriminator rejects the generated images;

Use a different loss or constraint for the discriminator, e.g. Wasserstein loss with gradient penalty/spectral normalization;

...so the discriminator gives a smooth score of how far the generated images are from the real ones instead of a saturated real/fake verdict;

Augment the images shown to the discriminator (e.g. adaptive discriminator augmentation);

...so it cannot overfit a small training dataset and become too strong;

*Note: Award **[1]** max if the answer only states that the discriminator should be made weaker or the generator stronger, without outlining how.*

Note: Credit other valid techniques that are described (e.g. feature matching, a two-timescale update rule).

(d) Award **[12]** max

Candidates should evaluate GANs and diffusion models against Visionary Studios’ needs (output quality, computational efficiency, training stability, flexibility and scalability, and consistency, lines 76–95), drawing on challenges 1 and 3, and express an opinion on which model would serve the studio better. The best responses will demonstrate research and cite real-world models, studies and/or experts in the field. Named models, studies and figures below are examples only; candidates are not required to know them.

Output quality and diversity

Diffusion models: Produce photorealistic, highly detailed images and cover the variety of the training data well, because the training objective (predicting the noise added to every training image) rewards modelling all of the data, whereas a GAN is rewarded only for fooling the discriminator. Research such as Dhariwal and Nichol (2021), “Diffusion Models Beat GANs on Image Synthesis”, found that diffusion models could match or exceed the best GANs on image quality while covering more of the data distribution.

GANs: Generate very sharp images and work very well in narrow domains with consistent data (e.g. StyleGAN for faces). However, mode collapse (lines 56–58) can reduce variety, which matters when clients range from fashion brands to video game developers. Candidates should be aware that GANs have been harder than diffusion models to scale to large, varied datasets.

Computational efficiency and speed

GANs: Generate each image in a single pass through the generator, so generation is very fast; for example, GigaGAN (Kang et al., 2023) generates a 512-pixel image in about 0.13 seconds. This suits thousands of images a week and quick iteration with clients.

Diffusion models: Need many sequential denoising steps (lines 36–45), so each image needs far more computation and the studio’s existing infrastructure may not support the workload (lines 81–82). Candidates should be aware that faster samplers, distillation (e.g. SDXL Turbo, 1–4 steps) and denoising in a compressed latent space reduce this gap, so the speed advantage of GANs is shrinking.

Training: Both need GPUs/TPUs, cloud-based platforms or HPC; diffusion models generally need longer training and greater computational resources (lines 84–85), which increases cost and energy use.

Training stability

GANs: Training balances two competing networks and is highly sensitive (lines 52–57). If one network becomes too strong, training can stall, oscillate or collapse, so experiments may have to be repeated, wasting time and computation. Mitigation techniques (e.g. Wasserstein loss, spectral normalization) help but do not remove the problem (HL challenge 3, lines 114–115).

Diffusion models: Train with a simple, stable objective (as formalized in the DDPM, lines 46–47), so results are more predictable — the trade-off between repetitive or suboptimal outputs and reliable but longer training described in lines 83–85.

Flexibility, scalability and control

Diffusion models: Are easily conditioned on text prompts, sketches and segmentation maps (lines 20–29) and support image-to-image translation and editing of parts of an image, which suits advertising campaigns, concept art and branded assets (lines 86–88). Many pre-trained models can be fine-tuned (transfer learning) instead of being trained from scratch.

GANs: Can be conditioned (e.g. class-conditional or image-to-image GANs), but text-to-image GANs have generally lagged behind diffusion models in following detailed prompts, so style adherence (lines 94–95) may be weaker. A GAN trained for one type of image may have to be retrained for a new client, which limits scalability.

Consistency

GANs: The generator maps each noise vector to one image, and nearby noise vectors give similar images, so reusing or slightly changing a stored noise vector can give a consistent character with small variations (lines 91–93).

Diffusion models: Each image starts from new random noise (line 37), so the same prompt gives different images; consistency needs extra techniques such as embedding-based approaches (lines 92–93), fine-tuning on the character, reference-image conditioning or fixed random seeds.

Ethical and legal considerations

Training data: Neither model removes the need for dataset curation for intellectual property and bias (challenge 2, lines 111–112). Carlini et al. (2023) extracted training images from diffusion models and concluded that diffusion models are much less private than GANs, so memorized copyrighted images may be a greater IP risk. On the other hand, mode collapse in a GAN can drop less common types of image, so under-represented groups may appear even less often in generated advertising.

Environmental impact: Diffusion models use more energy per image because of the many denoising steps, although training either model has a significant energy and carbon cost.

Alternatives

The studio need not rely on one model: it could use diffusion models for varied, prompt-driven work and GANs for fast, narrow tasks (e.g. upscaling or generating faces), or adopt models that combine both (e.g. adversarial diffusion distillation), as in the hybrid models of lines 60–75. Credit accurate research on newer approaches, e.g. flow-matching and transformer-based denoisers (e.g. Stable Diffusion 3).

Conclusion

A reasoned conclusion that expresses an opinion supported by evidence.

Marks	Level descriptor
No marks	• No knowledge or understanding of the relevant issues and concepts. • No use of appropriate terminology.
Basic 1–3 marks	• Minimal knowledge and understanding of the relevant issues or concepts. • Minimal use of appropriate terminology. • The answer may be little more than a list. • No reference is made to the information in the case study or independent research.
Adequate 4–6 marks	• A descriptive response with limited knowledge and/or understanding of the relevant issues or concepts. • A limited use of appropriate terminology. • There is limited evidence of analysis. • There is evidence that limited research has been undertaken.
Competent 7–9 marks	• A response with knowledge and understanding of the related issues and/or concepts. • A response that uses terminology appropriately in places. • There is some evidence of analysis. • There is evidence that research has been undertaken.
Proficient 10–12 marks	• A response with a detailed knowledge and clear understanding of computer science. • A response that uses terminology appropriately throughout. • There is competent and balanced analysis. • Conclusions are drawn that are linked to the analysis. • There is clear evidence that extensive research has been undertaken.
