# Computer Science

## *Timezone 1*

*To protect the integrity of the assessments, increasing use is being made of examination variants. By using variants of the same examination, candidates in one part of the world will not always be responding to the same examination content as candidates in other parts of the world. A rigorous process is applied to ensure that the content across all variants is comparable in terms of difficulty and syllabus coverage. In addition, measures are taken during the standardisation and grade awarding processes to ensure that the final grade awarded to candidates is comparable.*

# Contents

# Grade boundaries

## Higher level overall

| Grade: | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Mark range: | 0 - 14 | 15 - 29 | 30 - 40 | 41 - 51 | 52 - 61 | 62 - 72 | 73 - 100 |

## Standard level overall

| Grade: | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Mark range: | 0 - 14 | 15 - 29 | 30 - 43 | 44 - 53 | 54 - 63 | 64 - 73 | 74 - 100 |

## HL/SL solution

| Grade: | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Mark range: | 0 - 4 | 5 - 9 | 10 - 14 | 15 - 18 | 19 - 22 | 23 - 26 | 27 - 34 |

## Higher level paper one

| Grade: | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Mark range: | 0 - 16 | 17 - 32 | 33 - 45 | 46 - 55 | 56 - 66 | 67 - 76 | 77 - 100 |

## Standard level paper one

| Grade: | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Mark range: | 0 - 10 | 11 - 20 | 21 - 31 | 32 - 39 | 40 - 46 | 47 - 54 | 55 - 70 |

## Higher level paper two

| Grade: | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Mark range: | 0 - 10 | 11 - 20 | 21 - 24 | 25 - 31 | 32 - 37 | 38 - 44 | 45 - 65 |

## Standard level paper two

| Grade: | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Mark range: | 0 - 7 | 8 - 15 | 16 - 20 | 21 - 22 | 23 - 25 | 26 - 27 | 28 - 45 |

## Higher level paper three

| Grade: | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Mark range: | 0 - 3 | 4 - 7 | 8 - 10 | 11 - 13 | 14 - 16 | 17 - 19 | 20 - 30 |

# HL/SL solution

## The range and suitability of the work submitted

Like last year, many solutions consisted of a front-end coded in either Java, Python or HTML, which connected to a variety of back-end databases. There was a wide range in the quality of projects submitted. Most candidates showed a clear understanding of the assessment criteria, but many chose to address them in a superficial manner and few candidates put effort into producing high quality work.

There were fewer issues with inappropriate scenarios and inadequate products (e.g., basic spreadsheets) but some solutions did not address a real-life need for a real client and would be better suited as a classroom exercise in coding.

Some candidates made worthwhile efforts to create control systems that were typically developed as a prototype and not as a fully functional solution.

## Candidate performance against each criterion

### Criterion A

This was generally done well with appropriate evidence of client interaction. Unfortunately, many consultations were less meaningful as candidates decided on a product before the interview, instead of first finding out from the client what the actual problem is. The rationale for the solution was rarely described explicitly and the rationale for the software often remained generic and not product related. Criteria for Success (CfS) need to be sufficient to fully describe a functional solution. They need to be specific and testable. Inadequate CfS constitute a lack of focus and will have an impact on achievement in Criterion A, Criterion B, Criterion D and Criterion E.

Please note that examiners cannot be expected to listen to (potentially lengthy) audio/video files. Evidence of client interaction must be in writing.

### Criterion B

Examiners witnessed a continuing decline in effort to document the design. Many candidates submitted a few, incomplete design components that showed limited algorithmic thinking. Most Record of Task documents did not reflect the need to address a real problem with a real solution and lacked a proper implementation stage during which the client should use the final product for its intended purpose over a period of time.

### Criterion C

Most projects showed moderate to high complexity at standard level, but ingenuity was often impacted by a poor design. Reasonably good attempts were made to describe some tools and techniques, but only a few candidates achieved at the highest band by explaining and justifying. Stronger projects tended to provide detail by annotating code excerpts, while weaker projects provided a narrative without actual explanations.

### Criterion D

Most videos did not show the full functionality of the product as a solution. Issues included using limited test data, talking about functionality without actually showing it, not showing persistence of data changes by restarting the program, only showing inadequate / incomplete Criteria for Success. Similarly, most

candidates should provide better documentation in Criterion B and Criterion C, add appropriate comments in the code in order to make extensibility straight forward. Note that a trivial or underdeveloped solution cannot achieve marks for Extensibility.

### Criterion E

Generally, candidates attached in an appendix the evidence of client feedback after testing, but few candidates use this feedback in their own evaluation of the Criteria for Success. Recommendations for improvement should address the current solution. Unfulfilled Criteria for Success are typically considered trivial improvements, as are generic suggestions for 'more data' or 'more functionality'.

## Recommendations and guidance for the teaching of future candidates

Some candidates embarked on overly ambitious projects like a fully functional library system or a time-tabling system for their school. These project ideas are beyond the scope of the Computer Science IA and should be discouraged.

Teachers should encourage the candidates to follow the Systems Development Life Cycle and document the various stages as they work through them.

There is a clear correlation between strong Criteria for Success and successful projects. It is therefore recommended that teachers spend time on guiding candidates to produce sufficient, specific and measurable Criteria for Success.

Candidates should be guided to create a proper Design Overview which documents their algorithmic thinking in flowcharts and/or pseudocode, and which explains the data structures to be used in their solution. Screenshots of the actual product have no place in design and will be discounted.

Please limit the screencast videos to 7 minutes. Do not allow students to speed-up their videos as sped-up videos tend to become unintelligible. These videos should only show evidence for the full functionality of the solution. Code design or extensibility should not be included in the screencast.

Candidates should be guided to write a proper evaluation against the Criteria for Success, which incorporates and refers to meaningful client feedback to be added in an appendix. Recommendations for improvement should go beyond unfulfilled CfS (which are considered trivial). They should avoid the word ‘more’ (more content, more images, more levels, etc.) and instead focus on realistic improvements to the existing product (like solving problems that only became apparent after the client started using the product).

The majority of projects followed the guidelines for submission and included the appropriate documents. Some candidates submitted links to either product or video. This is not acceptable and linked components will be ignored by the moderator. Both the product and the video need to be uploaded as part of the submission. Only one 7-minute video should be included to show evidence of the functionality of the solution.

## Further comments

Please check your submission after uploading it to IBIS. Schools will no longer be consulted if components are missing or corrupted.

# Higher level paper one

## General comments

The syllabus coverage in most schools is very good.

A few schools, however, need to look carefully at their courses to ensure that both theory and constructing algorithms are each given sufficient attention.

Candidates achieved a wide range of marks on this paper.

## The areas of the programme and examination which appeared difficult for the candidates

Some candidates constructed excellent algorithms.

Some candidates appeared to be well prepared in the areas covering some of the theoretical aspects of computer science, such as those in Section A and those in questions 10 and 11 of Section B

## The areas of the programme and examination in which candidates appeared well prepared

Some candidates lost marks by not reading question thoroughly.

Candidates often lost marks by not fully answering questions or by not providing sufficient depth in their responses.

Whilst the majority of candidates showed an improved facility in the construction of algorithms, it is still the case that some students are unable to tackle this aspect of the course with confidence.

Some candidates who constructed excellent algorithms did not appear to be prepared well in the areas covering some of the theoretical aspects of computer science. They wrote their answers relying on guesswork and luck. It is evident that these students did not learn definitions of computer science terms such as “polling”, “virtual memory”, “legacy system”, “feedback”. They answered questions 3, 4, 9, 10b and 12b by guess or by golly, hoping to earn marks for these answers.

## The strengths and weaknesses of the candidates in the treatment of individual questions

### Question 1

Most candidates were able to outline the function of a web browser.

### Question 2

Most candidates were able to identify two applications of queues in computing.

### Question 3

A good range of answers was seen for this question. Many answers were excellent. A small number of candidates confused Unicode with a programming language.

### Question 4

(a) Many good and comprehensive answers were seen.

(b) Many candidates scored full or nearly full marks for this question. Some candidates confused the term 'polling' in computing with focus groups in which people are polled about the river pollution.

### Question 5

Many candidates scored full or nearly full marks for constructing a truth table for the logic expression

(A NAND B) NOR C

### Question 6

Many candidates provided vague or too general answers.

### Question 7

A well answered question. Many candidates were able to distinguish between RAM and ROM.

### Question 8

(a) It is evident that some candidates do not know the meaning of the term 'leaf node'.

(b) A well answered question.

(c) Some candidates did not delete the root node.

### Question 9

Many candidates correctly answered this question

Some candidates confused "virtual memory" with "cloud storage".

### Question 10

Many responses to (a) and (b) were verbose while missing crucial information that would award the mark.

(c) A well answered question.

(d) Problems that may occur during data migration were well explained.

(e) A well answered question. Candidates seemed to be very familiar with the direct changeover and parallel running.

(f) All candidates attempted it and many got it right.

### Question 11

Many responses to (a) and (b) were verbose while missing crucial information that would award the mark.

(c) A well answered question.

(d) Problems that may occur during data migration were well explained.

(e) A well answered question. Candidates seemed to be very familiar with the direct changeover and parallel running.

(f) All candidates attempted it and many got it right.

### Question 12

(a) and (b) The majority of students attempted it and many got it right. However, some responses were vague enough not to describe the situation with reference to the basic technical terminology of the devices.

(c) and (d) Many candidates provided vague or too general answers.

### Question 13

(a) Most candidates achieved some marks for this question, full marks were not rare.

(b) was reasonably well answered with many gaining at least two of the marking points.

(c) Most candidates outlined well the purpose of the stack access method isEmpty().

(d) There were excellent responses to this question. Most candidates achieved some marks for the algorithm. However, some candidates did not score any marks for Part (d).

### Question 14

Responses to question 14 ranged from poor to excellent.

Many candidates constructed excellent algorithms. However, constructing algorithms proved difficult for some candidates. They made some common errors while using arrays and loops, such as overlooking boundaries, two-dimensional array was not indexed by two subscripts - one for the row and one for the column, initializing the highest row total incorrectly, outputting the highest total many times, etc.

## Recommendations and guidance for the teaching of future candidates

- Teachers should ensure that all basic terms, given in the syllabus, are familiar to the candidates.
- Candidates should be familiar with computer science terms and write answers that contain computer science, not general observations.
- Candidates should be taught to understand the command terms (which are linked to a specific level of assessment object) and a question's mark total in order to gauge the appropriate length of a response, and consequently better manage their time during the examination.
- When an application or context is given, it is important to keep this in mind and answer appropriately.
- Candidates should be discouraged from writing long repetitive paragraphs for questions which have only one or two marks allocated to them.
- Candidates should be exposed to programming concepts and pseudocode. They need to develop their confidence in understanding and writing algorithms.
- Classes should spend time on problem solving, giving candidates the experience of breaking down unfamiliar problems into algorithmic steps and pseudocode.

# Standard level paper one

## General comments

Candidates achieved a good range of marks on this paper with candidates generally attempting all of the questions on the paper.

## The areas of the programme and examination which appeared difficult for the candidates

Candidates who fully answered the questions or gave sufficient depth in their responses, in relation to the number of marks available, generally achieved higher marks. Candidates who gave responses appropriate to the keyword used in the question generally achieved higher marks than those who did not. For example, if the question states 'outline one issue', the response should begin by stating the issue, followed by an expansion of that issue. Responses that begin with a descriptive statement, often do not match the marking points and rarely match both marking points that would be given for the issue and its expansion. Practical questions requiring algorithms to be written in pseudocode also sometimes caused difficulty, with candidates not always fully following the rules and concepts of programming.

## The areas of the programme and examination in which candidates appeared well prepared

Candidates appeared to be well prepared in the areas covering some of the theoretical aspects of computer science requiring shorter, more specific responses, such as those in Section A, but also, some of the Section B sub-parts. Some good algorithm responses were also seen in this session.

## The strengths and weaknesses of the candidates in the treatment of individual questions

### Question 1

Candidates who were able to outline the function of a web browser by relating its use to the World Wide Web or its interaction with a web server achieved the best marks.

### Question 2

Candidates who stated the role of the memory address register in relation to the address of the current instruction achieved the mark.

### Question 3

Candidates whose responses outlined that Unicode is an established standard for data representation, or who were able to outline the relative higher number of characters capable of being stored, compared with ASCII, achieved the highest marks.

### Question 4

(a) Most successful candidates identified the fact that RAM is volatile as a characteristic. Other similar responses were also seen.

(b) Most successful candidates were able to explain or partially explain that cache is a very fast type of memory used to store frequently used data and instructions, going on to indicate that this helped to speed up the operation of the whole system.

## Question 5

Most candidates were able to score some marks for drawing the correct truth table for the given logic expression. A wide range of marks were seen, with many high scoring candidates.

## Question 6

Candidates who were able to give a statement about the structure of a collection followed by how it may be accessed achieved the best marks. For example, a collection is a grouping of a variable number of data items into a single unit, that can be accessed using methods, such as hasNext() or getNext().

## Question 7

Most candidates were able to identify at least one layer of the OSI seven-layer model, with many identifying two.

## Question 8

(a) The vast majority of candidates were able draw a trace table with the correct number of columns, many of whom scored at least one or more additional marks for completing the correct trace. Candidates are reminded to makes sure that any variables that are initialised in the algorithm also need to be initialised in the trace table, and that inappropriate entries, such as dashes, should not appear mid table, once a value has been established for that variable.

(b) Most candidates recognised that the purpose of the algorithm is to find the average of the numbers in the array that are between 5 and 7 inclusive.

## Question 9

Many candidates achieved a mark for outlining the fact that abstraction removes unnecessary details from a problem, with those who also stated that this was to reduce complexity or increase efficiency achieving both marks.

## Question 10

(a) Candidates mostly identified the need for training as an issue concerning the roles of end users that must be considered in relation to the new system. Other responses were also seen, including fear of redundancy and end-user satisfaction.

(b) Candidates who identified legacy systems as old systems still in use and/or as systems that were generally incompatible with newer systems or hard to maintain, achieved the highest marks.

(c) Most candidates named surveys/questionnaires or interviews as a method of gathering requirements from end users.

(d) Candidates who identified data loss or data corruption and who then went on to explain how this could occur during data migration, achieved the best marks.

(e) Candidates often described the meaning of parallel running and/or direct changeover, which was not the focus of the question. However, candidates who explained an advantage of parallel running, such as it is a less risky method of changeover, along with reasons why, achieved good marks.

(f) (i) A wide range of possible training methods were seen, with the vast majority of candidates achieving this mark.

(ii) Candidates who identified a correct method of training also achieved some or all of the marks for identifying advantages and disadvantages of their chosen method. Some candidates only gave either advantages or disadvantages, but those who gave both, generally achieved high marks.

## Question 11

(a) Virtually all candidates correctly stated one or two technologies required to provide a VPN from a VON tunnelling server, a VON aware router, an encryption protocol or VPN client software.

(b) Many correct answers were seen to identify factors that may affect speed of data transmission, including amount of traffic, size of data, physical distance to travel or characteristics of network equipment.

(c) Candidates mostly achieved the mark explaining that data compression reduces the size of the data to be transmitted. Some candidates achieved additional marks for noting that the reduced file size would reduce the transmission time or the amount of bandwidth consumed, which can significantly reduce costs.

(d) Candidates who stated that encryption scrambles readable text so that it can only be understood if it is decrypted using the correct key achieved both marks. Candidates should be aware that encrypted text can still be accessed or read, but that it cannot be understood.

(e) Candidates who described a firewall as monitoring incoming and outgoing network traffic and blocking traffic based on a pre-defined set of rules achieved both marks. Some candidates alternatively correctly identified the fact that a firewall can be used to prevent unauthorised access to a network.

(f) The vast majority of candidates achieved marks for discussing the social impacts of employees working from home. Candidates who discussed both positive and negative aspects achieved full marks.

## Question 12

(a) Candidates who achieved marks here generally did so for stating characteristics of an array including that it is fixed size, all the data in any given array must be of the same data type, or that items of data in an array may be accessed using an index.

(b) Candidates mostly identified linear or sequential search and binary search to achieve both marks.

(c) Some very good algorithms were seen for the sub-program check() with many high scoring candidates. However, candidates are reminded that if an algorithm is dealing with an array, as in this case, collection methods are not appropriate as access tools.

(d) A good range of responses were seen requiring a selection sort algorithm to sort the ROOMNUMS array into ascending order. The full range of marks were seen. Some common issues noted were the

incorrect use of bubble sort features, when a selection sort was required, or pseudocode that would result in descending order, when ascending order was required.

## Recommendations and guidance for the teaching of future candidates

- Students should be encouraged to develop their programming skills using the standard pseudocode as defined in the syllabus and supporting documents. This encourages them to demonstrate their understanding of the structure of algorithms in a manner that is transferrable to multiple programming languages and encourages computational thinking. Students should also be given sufficient time and practice in the use of pseudocode to construct standard and more complex algorithms, using a range of data structures, which can then be developed into unfamiliar and more comprehensive problems.
- Students should be encouraged to carefully read each question to ensure that they understand what is being asked and ensure that the response they give fully answers the question, and is written in the context of the question. This will give them the best chance to achieve high marks. Students should also pay attention to the number of points written in their responses and try to match them to the number of marks available for the question they are answering.
- Students should look at the command word in the question and make sure that their response is appropriate for that command word.
- Students should be encouraged to make use of past papers and published mark schemes to help develop their examination technique and ensure that they practice the writing of appropriate full responses to questions.

# Higher level paper two and standard level paper two

## Overall comments

This session saw the re-introduction of this paper after a four year hiatus. As the four options are very different in nature, each one has been commented on individually.

The Standard Level (SL) forms a complete subset of the Higher Level paper (HL), so this content has been merged. The SL question numbers have been added after, for example:

Question 9 / Question 7.

The percentages of candidates taking each option in the exam were:

| | SL | HL |
|---|---|---|
| Option A | 21 | 30 |
| Option B | 6 | 5 |
| Option C | 15 | 13 |
| Option D | 58 | 52 |
| TOTAL | 100 | 100 |

There were many cases where candidates from one school completed different options. In some, very rare cases, some candidates attempted to answer questions from more than one option.

## Option A

## General comments

The syllabus seems to be covered well by most schools. There are some outstanding candidates and only a few candidates whose performance was poor. Across the cohort there were some excellent responses to every question in this paper.

## The areas of the programme and examination which appeared difficult for the candidates

The normalization question was a challenge for some candidates, but majority got some marks and showed a partial understanding of the concept.

Database model questions lacked precision and it appeared most candidates did not understand this concept.

Many queries lacked a structured approach though a few candidates approached it methodically.

## The areas of the programme and examination in which candidates appeared well prepared

The questions on the relationships in ER diagram the descriptive answers in the HL section were done reasonably well by candidates.

## The strengths and weaknesses of the candidates in the treatment of individual questions

### Question 1

(a) (i) Understanding the relationship between entities was done correctly by most candidates though some did not appear to understand the question. Most knew the different types of relationship that existed.

(a) (ii) Most of the candidates could identify a primary key.

(a) (iii) Only some candidates could apply the foreign key concept to identify the table.

(b) Most candidates were able to identify the 'Make' was correctly based on the given query.

(c) Though the question clearly hinted on the 'Format' some of the candidates could not identify that as a validation check.

(d) Most candidates were able to understand the basic structure of the query though some of them could not determine the relationship between the two tables correctly. Most of the candidates who got the basic structure correct failed to include the join between the CustID field for both tables.

(e) Most candidates gave generic answers and some of them were not able to precisely state the reason why queries are used to create views.

(f) Most candidates wrote answers on 'languages' and clearly drifted from the DDL concept. Some precise answers were given by some candidates.

(g) Most candidates gave generic answers and did not correctly write the most pertinent points in 'importance of data consistency'.

### Question 2

(a) Most candidates understood the 'time' and 'cost' factor. Some wrote 'size of the database' and that was not a factor to consider for database recovery strategy.

(b) Most candidates were able to do this question though some of them incorrectly suggested the hardcopy recovery in another place as the only solution.

(c) A majority of candidates were able to identify at least one way of staff not having access to sensitive information but some of them did incorrectly suggest password protection alone and encryption.

### Question 3

(a) Most candidates were clear about the problem with the given table and stated both 'need atomic values' and the example from the table given in the question. Some candidates did not use the terms 'atomic' or 'multivalued'.

(b) Most candidates were able to decompose the original table. While the primary keys were identified in most cases, only a few candidates identified the foreign keys. The. There were answers that showed that candidates did not understand the concept of normalization.

(c) Most candidates identified few of the elements of database modelling though some identified most of the points in importance of database modelling. There were some generic answers to this question that did not address the question.

### Question 4

(a) Most candidates were able to define the term data warehouse.

(b) Many candidates were unable to describe about the Network data model. They tended to drift into 'network data transfer' which was not relevant. Very few candidates achieved full marks for this question.

(c) Most candidates were able to describe how data from different sources could be cleaned before being loaded into a data warehouse.

(d) Many candidates were able to give advantages of database segmentation though the disadvantages were not addressed as well.

(e) Most candidates demonstrated a good understanding of the data analytics and were able to apply it to the given context.

### Conclusions

Some candidates demonstrated a lack of understanding of the concept of normalisation.

Some candidates did not use appropriate terminology in their responses.

## Recommendations and guidance for the teaching of future candidates

While candidates seemed to be able to understand the concepts such as normalisation and abstraction, they were not able to apply these concepts to a given scenario. Students should be taught these concepts using a range of scenarios and iteratively.

Terminology must be precisely used while writing answers. While some candidates appear to understand the concepts their answers often lacked precision and they lost marks as a result of this.

# Option B

## General comments

As every year, there was a wide range of achievement, but very few excellent candidates. Across the cohort, there were some good responses that were straightforward and based on conceptual understanding. However, many candidates showed superficial knowledge and understanding. Specifically, the extended response questions and the pseudocode construction questions were not well-answered.

## The areas of the programme and examination which appeared difficult for the candidates

Constructing algorithms remains difficult for many candidates who do this option. Extended response questions were often addressed in a superficial manner with limited writing that could not achieve full credit. Some candidates used generic statements lacking specific detail relevant to the questions.

## The areas of the programme and examination in which candidates appeared well prepared

Most candidates responded well to the questions related to models vs simulations, hardware implications of 3D rendering, and supervised vs unsupervised learning, showing fair theoretical understanding of these concepts. Questions that required extended answers were typically well done by those candidates who wrote extensive responses.

## The strengths and weaknesses of the candidates in the treatment of individual questions

### Question 5 / Question 4

(a) This question was answered well.

(b) This question was answered well.

(c) The question was answered well by most of the candidates. Some candidates showed limited ability to write pseudocode, and a few gave no response.

(d) Many candidates gave vague answers that did not address the question. Only a few realized that one needs to switch off or replace appliances to save electricity.

(e) Most candidates did not realize the basic intent of the question and wrote lengthy responses to describe algorithms.

(f) Many candidates could not produce the correct pseudocode to find and output the names of the required months.

### Question 6 / Question 5

(a) The vast majority of candidates only provided the definition of model versus simulation but did not relate their response to the given scenario.

(b) Advantages and disadvantages of using a simulation were clear but most candidates did not relate their response to the given scenario.

(c) The scenario was not well understood by many of the candidates and their responses were off course.

(d) Most candidates were able to identify a suitable ethical issue, but many had difficulty developing their response into a discussion about the trade-off between concerns and benefits.

### Question 7 / Question 6

(a) This question was answered well.

(b) Most candidates did not know the purpose of a keyframe in an animation.

(c) Most responses were superficial or unfocused. Candidates typically wrote how much processing is involved in rendering without relating this to specific hardware components.

Hardly any candidate referred to the given scenario of game design.

### Question 8

(a) Some candidates were able to provide specific reasons why a genetic algorithm is iterative. However, most responses were generic.

(b) A wide range of responses was seen, many being to the point, but equally many showing that the idea of a fitness function was unknown.

(c) The majority of the candidates were able to provide a reasonable description of supervised and unsupervised learning. However, a few of them were unable to develop the responses beyond the generic points.

(d) The majority of the candidates were unable to provide a reasonable response for this question. The responses were mostly limited to a few generic points without referring to actual NLP techniques.

(e) Many candidates did not provide a structured response to this question. Most provided a mix of vaguely correct statements that meandered between human and computer learning without clearly focusing on differences. Subsequently, candidates who wrote lengthy responses would achieve some marks here and there.

(f) This question was generally answered well.

## Recommendations and guidance for the teaching of future candidates

Candidates should be exposed to programming concepts and pseudocode. They need to develop their confidence in understanding and writing algorithms. Candidates are also well-advised to not only learn the required topics by heart but to focus more on conceptual understanding. This will allow them to appropriately apply such concepts in a given scenario.

# Option C

## General comments

The syllabus seems to be covered by most schools. There are some outstanding candidates and only a few candidates who were very poor in their performance. Across the cohort, there were some excellent responses to every question in this paper.

## The areas of the programme and examination which appeared difficult for the candidates

The candidates found a few questions a little more challenging than usual such as 10(d) / 8(d), 11(d) / 9(d) and 12(f). The questions were based on search engines being unable to access the deep web, comparing two given methods to protect an artist's intellectual property and the significance of ontologies and folksonomies in the development of the semantic web. These questions expected candidates to apply the knowledge in a contextual way.

## The areas of the programme and examination in which candidates appeared well prepared

Domain name server, search engine optimization, protocols, search engine ranking, drawing a directed sub-graph, cloud computing and collective intelligence.

## The strengths and weaknesses of the candidates in the treatment of individual questions

### Question 9 / Question 7

(a) Many candidates reasonably answered this question. However, some candidates were unable to use the correct keywords to define the term.

(b) Most candidates reasonably answered this question.

(c) The question was answered correctly by most of the candidates.

(d) Many candidates found this challenging and could not use the appropriate key terms to differentiate protocol and standard.

(e) Most candidates didn't answer this question correctly and failed to use appropriate key terms to outline the meaning of connectionless in the context of transmission of page requests.

### Question 10 / Question 8

(a) (i) The question was reasonably answered by most of the candidates. However, there were few generic responses as well showing no understanding of the given code fragment.

(a) (ii) Many candidates reasonably answered this question. However, majority of the candidates failed to score full marks due to lacking specific details of how CGI enables the execution of scripts on the server.

(b) (i) Most candidates correctly answered this question.

(b) (ii) Many candidates reasonably answered this question. However, most of the candidates were unable to score full marks as they couldn't explain the significance of incoming and outgoing links in search results using valid points.

(c) The responses were mostly generic and lacked the technical details.

(d) Most candidates did not answer this question correctly. Most responses were generic and superficial in nature.

### Question 11 / Question 9

(a) The question was answered correctly by most of the candidates.

(b) Most candidates did not answer this correctly. The correct understanding of lossy and lossless compression lacked in majority of the responses. This resulted in most candidates scoring very few marks as they couldn't apply the knowledge correctly to explain the appropriateness of lossy compression for eBook and album.

(c) Most answers were generic in nature and related to cloud storage rather than cloud computing.

(d) Most candidates had difficulty in answering this question. Most responses were generic and superficial in nature and used limited points to compare both the methods given in the question to protect the artist's intellectual property.

### Question 12

(a) (i) Most candidates answered this question correctly.

(a) (ii) Most candidates answered this question correctly.

(b) Most candidates answered this question correctly, but some were unable to identify the correct characteristics.

(c) Most candidates were able to draw the graph. However, majority of the responses were unable to score full marks due to not correctly linking the nodes in the SCC.

(d) Most candidates provided a generic and superficial response for this question. Majority of candidates were unable to suggest the correct reasons for change in the percentages of IN and OUT portions.

(e) Majority of the responses lacked the appropriate characteristics of collective intelligence.

(f) Majority of the candidates showed a little understanding of the ontologies and folksonomies. This resulted into most candidates scoring a very few marks in this question. Majority of the candidates were unable to structure the response correctly.

### Conclusion

On the whole this option had candidates answer both application based questions and the content well though by a sizable number of candidates lost marks due to lack of details in answers.

## Recommendations and guidance for the teaching of future candidates

- Candidates are able to grasp concepts but when asked in context they are mostly unprepared to answer precisely. This can be addressed by discussing the various concepts with real-life examples and getting students to see the application clearly.

- Candidates should be encouraged to use appropriate key terms in their responses to score higher as a majority of the responses lacked the technical knowledge.
- Candidates should also be helped to develop a better understanding of the command terms used in the questions. This will help them to understand the question well and answer accordingly.

## Option D

### General comments

Most candidates showed good knowledge of the common SL/HL syllabus but relating the definitions and concepts to real-life situations was often difficult.

Too many HL candidates did not show a good understanding of the HL concepts, with some of them achieving hardly any marks on the linked list question.

Like every session, some candidates wrote outstanding code in their responses and some clearly had not studied the standard algorithms that are required.

### The areas of the programme and examination which appeared difficult for the candidates

Interpreting code remains difficult. Many vague responses were seen that lacked the specificity necessary to be credited.

Another issue is using the methods provided in class descriptions (given in the stem of a question) to write the code for missing methods. Many candidates instead defer to methods that they are familiar with, but that do not exist in the scenario provided.

Too many candidates were unable to correctly write the code to declare and instantiate an array of objects.

### The areas of the programme and examination in which candidates appeared well prepared

Candidates had a reasonable / good understanding of the Standard Level theory. Most were able to describe the impacts of encapsulation and inheritance.

### The strengths and weaknesses of the candidates in the treatment of individual questions

#### Question 13 / Question 10

(a) This was answered well, although some of the weaker candidates only provided examples.

(b) Any reasonable answer was accepted.

(c) Many candidates had not learned the meaning of the arrows in UML diagrams. Most guessed correctly, but “owner has a property” was incorrect.

(d) Candidates struggled to explicitly link the definition to the context.

(e) This question was answered well, since the 'static' concept has appeared on several past exam papers (fairly) recently.

(f) It appeared from the various guesses that the application of the 'static' modifier had not been covered in many schools.

(g) About half the candidates lost marks on this standard bit of code.

### Question 14 / Question 11

(a) Many candidates wanted this to be about data hiding (which was credited), but the essence of encapsulation really is the bundling of variables and methods (acting on those variables) into a single class.

(b) This question was answered well.

(c) This question was answered well, particularly because candidates wrote extended responses incorporating all that they learned about inheritance.

(d) Many candidates understood that this question addressed extended character sets, and many got at least one mark referring to Unicode in their answer. Generic answers about internationalization received no credit.

### Question 15 / Question 12

(a) This question was answered well.

(b) This question was not well-answered because it addressed specific (but basic) knowledge about a String being an immutable object. This means the original String variable can only be changed by a direct assignment.

(c) This question was less well-answered than expected. Many candidates gave vague answers that lacked the required specificity. "To find the houses in a city" versus "to output the addresses of all the houses in city x".

(d) Many candidates did not connect the hint in the question (unfilled positions) to the intended run-time error of a null pointer exception. Some found a syntax error instead (incorrect output statement) which was given partial credit (also for the solution in 15d ii).

(e) Candidates who had studied the standard sorting algorithms did very well, typically achieving full marks. On the other hand, many candidates had not studied these and gained very few marks.

(f) The hardest question in the SL paper required some understanding of array handling. The question allowed for a diversity of approaches and many good algorithms were seen, but also many incomplete attempts.

### Question 16

(a) Not many candidates were able to provide specific features, that were correct and concise. Most received some credit for stating that ADTs are dynamic even though its implementation (static or dynamic) is specifically not a feature of an ADT.

(b) Similar to question 15c), there were many imprecise responses that received either partial or no credit.

(c) This question was not well answered. Few candidates understood that this question required the exclusive use of the methods provided in the definitions of ClientNode and ClientList. Since the list was not of the type LinkedList or ArrayList, candidates cannot use methods from those classes.

(d) Many good algorithms were seen that outlined proper steps, meaning that candidates have learned how to sequence steps to perform a complex task.

(e) The Binary Search is a standard, recursive algorithm that candidates should be able to reproduce. It is however clear that recursion in general (many non-recursive responses given) and the Binary Search in particular are a weak aspect in the preparation of this year's HL candidates.

## Recommendations and guidance for the teaching of future candidates

Schools are increasingly introducing programming and computational thinking modules lower down the school. This strategy puts these schools at an advantage when offering computer science at IB Diploma programme, as their candidates will be better prepared for the practical aspects of this course. All schools are encouraged to do the same.

In addition to learning definitions of concepts by heart, candidates should be given the opportunity to experience and explore the application of those concepts when coding classroom exercises. Constructing the standard algorithms for searches, sorts and recursion can help with better understanding of array and linked list concepts. Candidates should be taught to provide precise answers when interpreting code excerpts. Candidates should be provided with opportunities to develop methods based on a given context (for example within a partially developed class).

## Further comments relating to all four options

Teachers should ensure candidates know which option they should be attempting in the exam. In most schools where candidates attempted more than options, the candidates did less well. This may suggest the candidates are answering questions on options they have not addressed in class.

Teachers should ensure candidates are aware they only need to answer questions from one option. Where candidates answered questions from more than one option, they tended to perform less well (self-penalised).

Teachers should teach the theory using examples wherever possible. The exam questions are scenario based, and this requires candidates to link the theory they have learned to the scenario. This higher level of thinking means that generic or rote learned answers are likely to only achieve partial marks.

Theory should be taught iteratively and in a range of contexts. The text on pages 12 – 15 inclusive that relates to the 'approaches to teaching and learning' and the text on Page 27 at the start of Topic 4 also applies to each of the options.

# Higher level paper three

## The areas of the programme and examination which appeared difficult for the candidates

Candidates found the 12-mark extended question challenging. It required a broad understanding of machine learning and recommender filtering systems. Many candidates did not have detailed computer science knowledge and wrote descriptive responses that reworded the case study. In some cases, parts of their answers were incorrect. Few candidates showed evaluative abilities, and concepts such as data sparsity, cold start, and popularity bias were listed rather than discussed.

Candidates also struggled to identify two cloud deployment models despite this being on the list of key terminology. Candidates struggled to identify two differences between PaaS and SaaS. Again these were listed as additional terminology.

## The areas of the programme and examination in which candidates appeared well prepared

Most candidates did not adequately prepare for the case study, and some seemed to have done little research. Candidates understood the concepts from a broad perspective only. Few were prepared thoroughly. The ethics question on the right to privacy was answered reasonably well, but this topic is in the computer science curriculum guide.

Candidates had an overview knowledge of machine learning evaluative metrics. They understood the importance of evaluating a recommender system, but most were unable to calculate F-score from the precision and recall data provided.

## The strengths and weaknesses of the candidates in the treatment of individual questions

### Question 1

(a) Those candidates that had researched cloud deployment models gained two marks. Many candidates could not identify a cloud deployment model. They were under the misapprehension that PasS, IaaS, and SaaS were correct answers when the case study quite clearly describes them as cloud delivery models and asks the candidate to research cloud deployment models. Some gave practical examples, such as Amazon Web Services (AWS) and were awarded a mark.

(b) The second identify question was on cloud delivery models to make a clear distinction from the first question. Candidates struggled to identify two differences between Platform as a Service (PaaS) and Software as a Service (SaaS). However, candidates understood the term SaaS from the curriculum guide they mostly guessed what 'Platform' meant. It seemed that few candidates had researched the section on cloud computing.

### Question 2

(a) The majority of candidates understood aspects of the right to privacy in the context of the case study. However, few candidates gained full marks despite the mark scheme allowing a range of answers.

(b) Most candidates understood the term overfitting, and many were able to identify at least one strategy to reduce overfitting. The question asked for two approaches, with two marks for each one. Candidates generally failed to expand on the first point. Many candidates suggested reducing the training time, which is incorrect. They failed to understand that early stopping is not the same as training less. Again, this highlights a lack of research

### Question 3

Most candidates defined the F-measure metric and demonstrated an understanding of its purpose. Candidates misread the question (Explain how the F-measure might be applied in this scenario) and ignored the "how" after the command term suggesting a lack of exam technique. Consequently, few candidates explained its operation and gave vague answers. The mark scheme awarded marks for calculating precision and recall, components of the F-measure. Some candidates were able to do this.

### Question 4

The extended response question in HLP3 links relevant computer science with a question requiring students to discuss or evaluate two or more challenges. The challenge was to discuss the advantages and disadvantages of the different approaches to building a recommendation system. This allowed candidates to consider content-based filtering, collaborative filtering and the implementation methods of K-NN, matrix factorisation, or any other machine learning model.

Most candidates focused on content-based and collaborative filtering describing their process with little evaluation. Some candidates stated challenges such as popularity bias, overfitting, and cold start but failed to explain strategies to deal with these issues.

The majority of candidates wrote an adequate response. Few referenced the case study, and their analysis was generic movie recommendations system. Often the examples were youTube related rather than NextStar. There were very few proficient responses, and it was rare to see journal articles referenced.

## Recommendations and guidance for the teaching of future candidates

The case study is released approximately a year before the examination. Schools should approach the case study from a research perspective because it helps to put the computer science in the case study into context. Treating the topic in the same way as an extended essay increases students' overall understanding of the various issues and is likely to increase engagement. A collaborative approach may also spark students' interest in this subject.

While the additional terminology list is helpful, the challenges faced should be the primary focus, or the extended response question will be impossible to answer in any depth. This higher-level paper requires more than factual knowledge.

Schools should encourage students to pay close attention to the wording of the questions and, in particular, the command terms used. There is a tendency to pick up the question's theme and write everything they know about it. Candidates should refer back to the question and phrase their responses according to the requirements of the command term.

Students who read journal articles and complete video courses on the case study topic have a broader understanding of the concepts. This additional reading allows candidates to employ references in the extended response question. Research that goes beyond the case study will be rewarded in the extended response.

Several unofficial online forums have provided questionable information and guidance. Some students, following an online tutorial, programmed a K-NN algorithm and other machine-learning models. They tended to understand better how hyperparameters can influence the accuracy of the model. Teachers need to inform students that they should research broadly and cross-check the information they find.

The quality of candidates' handwriting has declined and, in some cases, made understanding the extended responses very difficult. While this is understandable, it is still a written exam, and students should practise writing legibly under pressure in class before taking the exam.