



# **MARKSCHEME**

## **SPECIMEN**

### **COMPUTER SCIENCE**

#### **Higher Level**

#### **Paper 2**

## General Marking Instructions

These instructions should be read in conjunction with the examiner instructions on IBIS.

Once approved for marking, you can download up to 20 scripts into your worklist.

1. Follow the markscheme provided, do **not** use decimals or fractions and mark only in **RED**.
2. Where a mark is awarded, a tick (✓) should be placed in the text at the **precise point** where it becomes clear that the candidate deserves the mark.
3. For extended responses where markbands are used, it is helpful to write a brief comment indicating why the level was awarded.
4. Sometimes, careful consideration is required to decide whether or not to award a mark. In these cases write a brief annotation in the **left-hand margin** to explain your decision. You are encouraged to write comments where it helps clarity.
5. Unexplained symbols or personal codes/notations on their own are unacceptable.
6. Where an answer to a part question is worth no marks, put a zero in the mark entry box.
7. Every page and every question must have an indication that you have marked it. Do this by **writing “seen”** on each page where you have made no other mark.
8. Examiners should be aware that in some cases candidates may take a different approach, which if appropriate should be rewarded. If in doubt check with your Team Leader.

## Subject Details: Computer Science HL Paper 2 Markscheme

### Mark Allocation

For the option chosen candidates are required to answer **all** questions. Total 65 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:

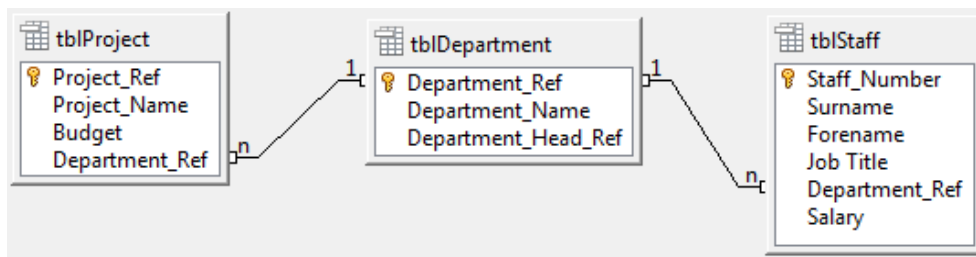
- Each statement worth one point has a separate line and the end is signified by means of a semi-colon (;).
- An alternative answer or wording is indicated in the markscheme by a “/”; either wording can be accepted.
- Words in ( ... ) in the markscheme are not necessary to gain the mark.
- If the candidate’s 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	<ul style="list-style-type: none"> <li>• In the case of an “identify” question read all answers and mark positively up to the maximum marks. Disregard incorrect answers.</li> <li>• In the case of a “describe” question, which asks for a certain number of facts <i>e.g.</i> “describe two kinds”, mark the first two correct answers. This could include two descriptions, one description and one identification, or two identifications.</li> <li>• In the case of an “explain” question, which asks for a specified number of explanations <i>e.g.</i> “explain two reasons ...”, mark the first two correct answers. This could include two full explanations, one explanation, one partial explanation <i>etc.</i></li> </ul>

## Option A — Databases

- A1.** (a) (i) Organized data store; **[1 mark]**
- (ii) Flat file – all data is stored in one table;  
Relational – data is stored in separate related tables; **[2 marks]**
- (b) (i) *Award up to [2 marks max].*  
Repetition of data;  
May lead to inconsistencies;  
May lead to partial updates;  
Unnecessary duplication; **[2 marks]**
- (ii) *Award up to [2 marks max].*  
Line managers are entered multiple times;  
Likelihood of inconsistencies;  
May lead to wrong connections between subordinates and managers; **[2 marks]**
- (c) (i) *Award marks as follows up to [4 marks max].*  
*Award [1 mark] for creating separate tables;*  
*Award [1 mark] for showing three suitable tables;*  
*Award [1 mark] for showing links;*  
*Award [1 mark] for showing the nature of the relationships;*  
*Award [1 mark] for correct ref to primary key;*



**[4 marks]**

*continued ...*

*Question A1 continued*

- (ii) *Award marks as follows up to [4 marks max].*  
*Award [1 mark] for all relevant tables selected;*  
*Award [1 mark] for all relevant fields selected;*  
*Award [1 mark] for correct condition;*  
*Award [1 mark] for correct link between tables (WHERE clause);*  
*Award [1 mark] for correct use of AND;*

**QBE View**

Field	Surname	Salary	Project_Name
Alias			
Table	tblStaff	tblStaff	tblProject
Sort			
Visible	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Function			
Criterion		>= 50000	

**SQL View**

SELECT “tblStaff”.”Surname”, “tblStaff”.”Salary”,  
 “tblProject”.”Project\_Name” FROM “tblProject”, “tblDepartment”,  
 “tblStaff” WHERE “tblProject”.”Department\_Ref” =  
 “tblDepartment”.”Department\_Ref” AND “tblStaff”.”Department\_Ref” =  
 “tblDepartment”.”Department\_Ref” AND “tblStaff”.”Salary” >= 50000

**[4 marks]**

- A2.** (a) (i) *Award up to [2 marks max].*  
 A unit of work / logical action;  
 Performed on a database;  
 Performed by the DBMS;  
 Independent of other transactions;

**[2 marks]**

- (ii) *Award [1 mark] for an answer that only identifies either the nature of a database state or a database schema. Award [2 marks] for an answer that covers both the ideas in the points below.*  
 State is the completed database with data;  
 Schema is the plans for the database;

**[2 marks]**

- (iii) *Award [1 mark] that only identifies that conflicting results may occur. Award an additional [1 mark] for the development of the initial point up to [2 marks max].*  
 Queries may produce conflicting results;  
 Because the total amount of money in all the accounts should be the same as before the transaction took place;

**[2 marks]**

*continued ...*

*Question A2 continued*

- (b) *Award [1 mark] for each step identified in the correct order, up to [4 marks max].*

Begin the transaction;  
Copy original data to a log;  
Perform the debit operation;  
Perform the credit operation;  
Error checks;  
If no errors occur then commit the transaction and end it;  
If errors occur then roll back the transaction and end it;

**[4 marks]**

- (c) *Award marks as follows up to a maximum of [6 marks max].*

*Award [2 marks] for identifying a cause of the conflict;*  
*Award [2 marks] for explaining the consequence;*  
*Award [2 marks] for the solution;*

*Answers may include:*

Different systems may attempt to access the same data at the same time;  
Both wanting to make updates;

If one starts and then the second finished before the first is saved;  
This could potentially lead to inconsistent updates;

Solution is to isolate the transactions;  
When one system is accessing the data, lock the transaction;  
Release it when transaction is committed;

**[6 marks]**

- A3.** (a) (i) *Award up to [2 marks max].*

Controls the creation of a database;  
Controls the maintenance of a database;  
Controls the use of a database;  
Mediates between the data handling applications and the operating system;

**[2 marks]**

- (ii) *Award up to [2 marks max].*

A means to query the data;  
Forms for displaying the data;  
Reports for producing output;  
Charts for displaying data;  
A DDL or equivalent for constructing/amending the schema;

**[2 marks]**

*continued ...*

*Question A3 continued*

- (b) *Award marks as follows up to a maximum of [4 marks max].  
Award [1 mark] for each clear use of data sharing/different views;  
Award [1 mark] for an explanation of why the feature is important;  
Answers may include:*

Data sharing allows use of the data by different users;  
and by different applications;  
Multiple views present different subsets of the data to different users;  
Data can be presented in different ways according to the user's needs;

**[4 marks]**

- (c) *Award marks as follows up to a maximum of [6 marks max].  
Award [1 mark] for an implication of changes to data structures;  
Award [1 mark] for an implication of changes to data itself;  
Award [1 mark] for relating each of these to applications, up to [2 marks max];  
Award [2 marks] for a clear, detailed and balanced explanation;  
Answers may include:*

The organization uses a variety of applications for manipulating data;  
It may need to amend or add to these;  
Changes to applications may conflict with existing data structures;  
Applications may conflict with each other;  
Applications must therefore act through the DBMS;  
This imposes consistency on the way that the data is manipulated;  
Data integrity is maintained by the DBMS rather than the applications;  
Data structures do not necessarily need to be altered in order to accommodate new applications;

**[6 marks]**

- A4.** (a) (i) *Award up to [2 marks max].*  
 A repository of stored data;  
 Related to a specific subject;  
 Time variant and non-volatile data;  
 Includes tools to extract, transform and load data into the repository;  
 Tools to manage and retrieve metadata; **[2 marks]**
- (ii) *Award up to [4 marks max].*  
 Large volumes of data;  
 Very different types of data;  
 Data is probably poorly integrated;  
 Legacy systems working independently;  
 Difficulty in answering strategic questions; **[4 marks]**
- (iii) *Award up to [4 marks max].*  
 Selecting only certain columns to load;  
 Translating coded values (*e.g.* if the source system stores 1 for male and 2 for female, but the warehouse stores M for male and F for female), this calls for automated data cleansing;  
 Encoding free-form values (*e.g.* mapping “Male” to “1” and “Mr” to M);  
 Deriving a new calculated value (*e.g.*  $\text{sale\_amount} = \text{qty} * \text{unit\_price}$ );  
 Sorting;  
 Joining data from multiple sources (*e.g.* lookup, merge);  
 Aggregation;  
 Transposing or pivoting;  
 Splitting a column into multiple columns;  
 Disaggregation of repeating columns into a separate detail table;  
 Lookup and validate the relevant data;  
 Applying any form of simple or complex data validation; **[4 marks]**

*continued ...*



*Question A4 continued*

- (b) *Award marks as follows up to [4 marks max].  
Award [1 mark] for an outline of association;  
Award [1 mark] for an outline of cluster analysis;  
Award [2 marks] for a contrast between the two;  
The contrast should be awarded if the outlines between the two show clear difference. Accept examples for “Meubles de France” to show contrast.  
Answers may include:*

**Association**

Looks at how entities/events are connected, and finds where one or more events may lead to another;  
May use examples linked to Meubles de France such as purchasing patterns of customers to assist in decision making;

**Cluster analysis**

Finds patterns in customer behavior by grouping and analyzing variables that connect them. Can find previously unknown links which help in decision making.  
May use examples linked to Meubles de France such as age of customers, amount spent on each visit, days of the week of visits, *etc.* to assist in decision making;

The difference is that association looks to see if a set of events leads to another whereas cluster analysis looks for unknown patterns of behavior.

**[4 marks]**

- (c) *Award marks as follows up to [6 marks max].  
Award [3 marks] for each clear explanation of a consequence of data mining which could lead to an infringement of civil liberties.  
Answers may include:*

Data mining can reveal connections in personal data;  
Which could be misinterpreted/give a misleading data profile (shadow);  
Used for reasons that the individual does not know/would not want;

Once the data is available, it may be hard to restrict access to it;  
There may be pressure from third parties to share the data for example from insurance companies or government agencies;  
Which is unknown/not wanted by the individual;

The data may appear benign, but in the hands of a third party lead to unforeseen issues or when analysed provide information that may not be accessible from a standard query. Not all customers will be able to foresee the extent to which their data may be used;

There may be mission creep and the reasons for the original data collection may be superseded meaning that the security of the data originally collected cannot be guaranteed;

**[6 marks]**

**Total: [65 marks]**

## Option B — Modelling and simulation

- B1.** (a) All the variables acting in a situation/system;  
Inputs and outputs;  
Reduces the system to a mathematical representation;  
Using functions/formulae/equations;  
*Award [1 mark] for each relevant point up to a maximum of [2 marks].* **[2 marks]**
- (b) Equation only has to be entered once;  
Values of Z, C, I and R can be changed;  
For each of the different buildings entered on the sheet;  
Value of V calculated automatically;  
*Award [1 mark] for each relevant point up to a maximum of [2 marks].* **[2 marks]**
- (c) All variables included – descriptions do not need to be given;  
Column for V contains formula for at least one building;  
Formula is correct;  
Copy down indicated for more than one building;  
**[4 marks]**
- (d) *More than one method may be used:*
- Method 1 (not efficient but worth **[3 marks]**)  
Lookup values held within the spreadsheet;  
On separate named sheets;  
User goes to each table as required;
- Method 2  
Drop down list for columns holding C and R;  
Linked to the lookup values;  
Held in lists;  
For example when R input different materials presented to choose from;  
Appropriate number inserted in cell;
- Other methods may be used. If necessary consult with your Team Leader.*  
*Award [1 mark] for each relevant point up to a maximum of [4 marks].* **[4 marks]**
- (e) Z input once as it is constant for all buildings, and known for the city;  
I known for each building – (purpose of library, school. etc.);  
Input manually;  
Each building would have to have C and R classified;  
Either by experts visiting or records held;  
Entered by drop down box for each building;  
V calculated;  
Those with V below a specified limit identified as in danger;  
*Award [1 mark] for each relevant point up to a maximum of [6 marks].* **[6 marks]**

- B2.** (a) Computer simulation changes the values of variables in the model;  
To see the effect that the changes have on the model (or its outcomes);  
*Award [1 mark] for the initial point and [1 mark] for any appropriate subsequent development of it up to a maximum of [2 marks].*

**[2 marks]**

- (b) *Award marks as follows up to [4 marks max].*  
*Award [1 mark] for a suitable simulation identified;*  
*Award [1 mark] for clear rules embedded outlined;*  
*Award [1 mark] for type of data and rules inaccuracy;*  
*Award [1 mark] for an effect of inaccuracy;*

*For example:*

In a traffic flow simulation a town planner may wish to restrict the traffic in a particular zone and sets up a simulation;

Using a map of the town, traffic flows according to the known traffic density and direction;

An out of date road plan is used and traffic flow simulated;

(for example it is ignored that a new road is under construction that goes around the town). The results could cause the planner to set up expensive and unnecessary measures to restrict traffic in the town causing local unrest;

In particular if the position of residents housing in the town was left out of the simulation these residents could find themselves unable to reach their own houses by car;

**[4 marks]**

- (c) *Award marks as follows up to [4 marks max].*  
*Award [2 marks] for a relevant advantage discussed;*  
*Award [2 marks] for a relevant social consequence discussed;*

*Answers may include:*

**Advantages** – knowing the weather in advance makes it possible to plan events such as harvesting of crops, deciding to hold the school fête indoors instead of out of doors, etc.

Predictions of drought or floods give organizations time to prepare to help with food aid or plan for evacuation.

**Social consequences** – having life more easily planned for disasters makes for a secure environment.

Lives can be saved if measures to cope with disasters are in place before they occur. Too much dependence could lead to misreading some of the known “weather symptoms”. The predictions are not perfect and a false sense of security could result in not being able to adapt to unexpected and unpredicted weather events.

**[4 marks]**

- B3.** (a) Data collected (from satellite or observation) continually;  
Recent path stored and updated as data received;  
Predicted path calculated;  
Using model based on current situation (and behaviour of previous cyclones);  
Predicted (and past) path displayed on simple 2D map of area;  
Coloured (shaded) to show force of cyclone;  
*Award [1 mark] for each relevant point up to a maximum of [4 marks].* **[4 marks]**
- (b) *Award marks as follows up to [4 marks max].*  
*Award [1 mark] for identifying a technical difficulty of data collection;*  
*Award [1 mark] for identifying a technical difficulty of processing;*  
*Award [1 mark] for each technical difficulty that is explained, up to [2 marks max];*  
*If only difficulties relating either to data collection **or** processing are addressed, award [3 marks max].*
- The following points are likely to be included:*  
Continual data collection from centre of cyclone difficult;  
Strength of cyclone needs to be recorded and transmitted which may not show on satellite picture;  
Transmission could be broken (unclear/interrupted) by the effect of the cyclone;  
Inaccurate data fed into the model would give false results;  
Real-time processing needs fast powerful computers;  
Which may not be readily available at the start of the cyclone;  
Computer resources need to be distributed to avoid losing the information in case of failure; **[4 marks]**
- (c) *Award marks as follows up to [6 marks max].*  
*Award [1 mark] for each advantage identified, up to [3 marks max];*  
*Award [1 mark] for each advantage explained, up to [3 marks max];*
- Visualization gives a quick way of seeing where the dangers are now and where they are likely to be;  
No need to interpret figures so time is saved;  
Location where rescue/support services are needed now and in near future are shown;  
Emergency services hence more effective and dangers to people can be minimized;  
Images distributed easily to many places – aid from across the world; **[6 marks]**

*continued ...*

*Question B3 continued*

- (d) 2D visualization is more rapid than 3D as the calculations are less complex and speed is necessary in this situation;

3D also requires more powerful computers which could be difficult to find in a situation that requires a distribution of computers;

Some aspects on the effect of the cyclone such as in a mountainous area would be visualized in 3D but not in 2D visualization;

Potential difficulty in arriving at an area due to damage could be seen in 3D but not 2D;

*Award [1 mark] for each relevant point up to a maximum of [3 marks].*

**[3 marks]**

- B4.** (a) In supervised learning the goal, or pattern, is known in advance;  
For example in optical character recognition where the set of characters is known in advance;

Unsupervised learning involves finding unknown solutions, patterns or conditions;

For example in data mining where unknown patterns can identify trends;

*Award [2 marks] for a clear outline of the difference and for each case [1 mark] for an identified application.*

**[4 marks]**

- (b) *Award [1 mark] for each relevant point, up to [6 marks max].*

Genetic algorithms work successively towards a solution from a starting point;

Which may even be a random set of solutions;

A (*fitness*) *function* is used to measure the ranking of a solution;

The set of solutions is examined against a fitness function;

The best solutions (generally the 50 % best) are retained and;

The retained solutions are mutated to generate another set of solutions;

This process is repeated until the best possible fit solution is identified;

**[6 marks]**

*continued ...*

*Question B4 continued*

- (c) *Award marks as follows up to [4 marks max].*  
*Award [2 marks] for an outline of the use of sensors by the robot;*  
*Award [2 marks] for the way in which the model can be built from the data collected;*

*The following points should be included:*

The robot makes a random set of moves;  
Records distances and direction of objects reached;  
In relation to itself;  
Building up a map of surrounding objects;  
This is repeated until all objects in the space have been placed in distance and direction from a starting point;

**[4 marks]**

- (d) *Award marks as follows up to [6 marks max].*  
*Award [1 mark] for a feasible link between the robot and human;*  
*Award [2 marks] for an outline of method and format of information sent from robot to human;*  
*Award [2 marks] for a method of human communication/control of robot;*  
*Award [2 marks] for a description of the way in which the rescue could be carried out;*

*Answers may include:*

A suggested link from robot to rescue workers;  
The format of the information sent (visual *etc.*);  
A description of the two-way communication between robot and rescue workers;  
A description of the way the above could be used to make a successful rescue;

**[6 marks]**

**Total: [65 marks]**

**Option C — Web Science**

- C1.** (a) Cloud computing services are provided for a particular group with a limited number of users;

**[1 mark]**

- (b) *Award marks as follows up to [4 marks max].*

*Award [2 marks] for the features of traditional client-server;*

*Award [2 marks] for features of cloud computing **which distinguish the two**;*

Traditional client-server model:

Servers on the one of the premises of the company;

Connected to a Local Area Network (LAN);

Maintained by IT team of that organization;

Cloud computing:

Servers outsourced to third party;

Maintained by third party technical support team;

Based on Internet connectivity;

Connected to a WAN;

*If feature are given, but are not different for each award [2 marks max].*

**[4 marks]**

- (c) *Award marks as follows up to [4 marks max].*

*Award [1 mark] for each benefit identified, up to [2 marks max];*

*Award [1 mark] for a reason for each benefit, up to [2 marks max];*

*Answers may include:*

Reduce costs as fewer technical staff will need to be employed;

Technical staff in third party may have greater expertise as they may be able to specialize;

May reduce costs as third party may benefit from economies of scale for purchase of storage, hardware, etc;

Expertise may be held by more people, so staff turnover may have less effect;

The effects of unexpected hazards may be reduced as data may be distributed across a number of locations;

**[4 marks]**

*continued ...*

*Question C1 continued*

- (d) *Award marks as follows up to [4 marks max].  
Award [2 marks] for a privacy issue identified and elaborated;  
Award [2 marks] for a security issue identified and elaborated;*

**Privacy**

Sensitive data is accessible to a third party;  
If outsourcing occurs, potential exposure of data is increased;

**Security**

How secure is the data?;  
Can it be guaranteed that this data will not be inadvertently passed to another company?;

*[4 marks]*

- C2.** (a) Can be rendered by all internet browsers;  
*Other answers relating to the language itself (e.g. use of tags to delimit statements) would also be acceptable.*

*[1 mark]*

- (b) Server retrieves event for days to be displayed from a database server;  
Takes results and generates HTML to display them in a table;  
Embeds HTML in page;  
Page sent to browser;  
*Use of bullet points and a list is acceptable for this type of question.*

*[4 marks]*

- (c) (i) A client-side script may be made available to a browser by the use of script tags that are embedded in the HTML code;  
By the use of an external file;
- (ii) A client-side script will not require access to a remote server so that any processing that is done will be done more quickly and use less bandwidth;  
This will reduce the load on the server;

*[2 marks]*

*[2 marks]*

*continued ...*



*Question C2 continued*

- (d) *Award marks as follows up to [4 marks max].*  
*Award [2 marks] for outlining the way meta-tags are used;*  
*Award [2 marks] for alternative methods;*

*Answers may include:*

Meta tags contain keywords/descriptions related to the web page's content;

They are embedded in HTML code can be read by search engines;

This can help their ranking;

Search engines have become more sophisticated;

They no longer place a high value on the content of meta tags;

As they are not always a reliable guide to the web page's content;

Search engines now use other parameters;

*e.g. Some will place more value on the number of links pointing to a web page;* **[4 marks]**

- (e) *Award marks as follows up to [6 marks max].*  
*Award [2 marks] for an explanation of the creation of server side scripts;*  
*Award [2 marks] for an explanation of the client-side interaction;*  
*Award [2 marks] for relating to the box office website;*

*Answers may include:*

Before the web page is generated scripts will be run on the server-side;

Examples of these scripts are;

The handling of inputs;

The retrieval of information from databases;

The performing of calculations;

Scripts will be written in languages such as PHP;

Server-side scripts are hidden from users and therefore secure;

The server sends data to the browser (client) in HTML;

This could also include JavaScript code which will be interpreted by the client's browser;

Allowing (in this case) booking details to be entered by the user on the client side;

Server driven information can be delivered on the fly using software such as Ajax;

**[6 marks]**

- C3.** (a) (i) It does not contain a fixed set of tags, therefore new ones can be added; **[1 mark]**
- (ii) A set of rules and procedures that both sender and receiver must adhere to in order to allow coherent data transfer; **[1 mark]**

- (b) *Award up to [3 marks max].*  
 The use of open standards implies that anyone can use them;  
 They are standards that are agreed from the beginning;  
 therefore ensuring interoperability;  
 For example, the Internet backbone relies on the IP protocol which is an agreed standard, allowing the transfer of information to occur; **[3 marks]**

<http://fsfe.org/projects/os/def.html> (accessed 16 Aug 2011)

- (c) *Award marks as follows up to [6 marks max].*  
*Award [2 marks] for the difference between lossy and lossless compression;*  
*Award [2 marks] for a relative example for each;*  
*Award [2 marks] for the reason for using lossy or lossless in the example;*

*Answers may include:*

Lossless compression is used when loss of data is unacceptable when transferring files such as audio files;

Lossy compression may not significantly affect the final version of the file when it is decompressed;

Lossy compression will reduce file size;

Reduced file size may be an important requirement such as in the use of MP3 music files;

Lossy compression results in faster file transfer;

Which is important when Internet connections are slow or files are large;

If lossy compression is used the original file cannot be reinstated;

*Also characteristics of the data itself (e.g. frequencies too high for human hearing) could be introduced.* **[6 marks]**

*continued ...*

*Question C3 continued*

- (d) *Award up to [2 marks max].*  
 Easier to set up;  
 Less time will need to be spent in configuring the network;  
*Other advantages could deal with the increased range of available files and the lower (or even zero) costs involved (depending upon the network).* **[2 marks]**

**C4.** (a) The web-pages are seen as vertices and the hyperlinks between them as edges; **[1 mark]**

- (b) (i) *Award up to [1 mark max].*  
*Answers may include:*  
 Lies at the centre of the bowtie structure between the IN and OUT web pages;  
 Is the part of the web where a web surfer is able to navigate to and from any particular web page; **[1 mark]**

- (ii) *Award up to [2 marks max].*  
*Answers may include:*  
 Have no links coming into them;  
 Tend to be newer, not had sufficient time to establish a number of in links;  
 Have links from them to other websites in the SCC, OUT (via tubes) or to other web sites (via tendrils);  
 Tend to be resource lists; **[2 marks]**

- (c) *Award [1 mark] for a basic statement and an additional [1 mark] for the development of it.*  
 The network diameter of the web growing no more than logarithmically with respect to the network size. Sources suggest there were 26 million pages in 1998, one billion pages by 2000 and 34 billion by 2011;  
 Therefore the diameter is not growing in a linear relationship with the number of web pages, so the ability of the web surfer to access all sites remains possible;  
 10-fold increase of web pages results in only 2 more additional “clicks”; **[2 marks]**

*continued ...*

*Question C4 continued*

(d) *Award up to [4 marks max].*

*Accept either an explanation of the HITS or an explanation of Google PageRank.*

*For example:*

Search engines use algorithms such as the Google PageRank or HITS to determine the ranking of any web page;

The Google PageRank calculates the rank as follows:

Rank is determined by number of votes for it. This is based on the number of “in” links and importance of pages voting for it;

Page rank uses a recursive algorithm;

However, some web masters use link farms to “artificially” raise the rank of the web page, some algorithms remove this information before calculating the rank;

HITS (hyperlink-induced topic search) is based on the following principles:

- Websites may be hubs (point to lots of authorities) or authorities (are pointed to by a number of hubs);

The HITS algorithm calculates the rank as follows:

- Determines a base set of web sites (a closed network)
- From this set a number of pages are located by a search engine to form a root
- Add to  $S$  all pages pointed to by any page in  $R$
- Add to  $S$  all pages that point to any page in  $R$
- Maintain for each page  $p$  in  $S$ :
  - Authority score:  $ap$  (vector  $a$ )
  - Hub score:  $hp$  (vector  $h$ )
- Calculate the authority weighting for each web page
- Calculate the hub weighting for each web page
- Normalizes the values;

**[4 marks]**

*continued ...*

*Question C4 continued*

- (e) *Award [1 mark] for a basic statement and an additional [1 mark] for the development of it.*

An ontology is the formal description of the concepts and relationships that exist within a specified domain of discourse.

“A folksonomy is a type of distributed classification system. It is usually created by a group of individuals, typically the resource users. Users add tags to online items, such as images, videos, bookmarks and text. These tags are then shared and sometimes refined.”

Folksonomies may be imprecise and informal, developing organically through social networking.

*[2 marks]*

- (f) *Award up to [2 marks max].*

Is a shared or group intelligence;

It can be measured;

May be based on independent decision making and a lack of conformity;

Based on consensual decision making;

Based on groups of people working openly;

Associated with distributed individual intelligences;

Is enabled by the Internet;

*[2 marks]*

*continued ...*

*Question C4 continued*

- (g) *Award marks as follows up to [6 marks max].*

*Award [2 marks] for advantages;*

*Award [2 marks] for disadvantages;*

*Award [2 marks] for a conclusion and/or judgments based on the given advantages and disadvantages (may be incorporated into the body of the answer);*

*Advantages of collective intelligence:*

Information can be easily shared, may prevent unnecessary duplication of data;  
Advancements may be rapid as large number of contributors may be involved, for example sites like Wikipedia may be considered as being developed through collective intelligence;

The total sum may be greater than the sum of the parts;

The project may develop further as many collaborators may freely add information therefore being more cost efficient than traditional methods of resolving problems;

*Disadvantages of collective intelligence:*

Too much information may be submitted;

The “group” may be informal so relevant information may not be processed as required;

The “group” may become unwieldy through too many contributors or lack motivation as there may be no designated project manager;

Knowledge may be developed without suitable checking mechanisms to ensure it is correct, decisions may be based on a consensus which may not be appropriate;

**[6 marks]**

**Total: [65 marks]**

## Option D — Object-oriented programming

**D1.** (a) Award **[1 mark]** for a definition, such as:

A variable/value that is passed to a method;

The value passed to a method in brackets/parentheses;

Award **[1 mark]** for a valid example from the code, such as:

`setRoute(int r)/r` is a parameter variable;

**[2 marks]**

(b) Award **[1 mark]** for a field and its data type, examples:

`String destination;`

`int/double/long length;`

Award **[1 mark]** for a corresponding data example:

“City centre” or “Bus terminal” etc.;

23 (km), 13 460 (m), etc.;

**[2 marks]**

(c) Award **[1 mark]** for any two correct outputs (including descriptors) and an additional mark for the third correct. Award **[1 mark]** for only the data items, if all correct.

Examples:

Award **[2 marks]** for the following:

Bus id:1001 - N Prakesh: Route: 431 start: Klang

Award **[1 mark]** for the following:

1001 - N Prakesh 431 Klang

**[2 marks]**

(d) Award marks as follows up to **[6 marks max]**.

Award **[1 mark]** for any loop;

Award **[1 mark]** for correct loop end point (allow test for fixed number in array, i.e. `> 3` or `== 4`);

Award **[1 mark]** for running through all array elements;

Award **[1 mark]** for any test for bus number;

Award **[1 mark]** for correct access to the Route number (not the Route object);

Award **[1 mark]** for correct test;

Award **[1 mark]** for correct output of drivers;

Example answer:

```
private void showBuses(Bus[] b, int n)
{
    // Show buses with route numbers < n
    for (int x = 0; b[x] != null; x = x + 1)
    {
        if (b[x].getBusRoute().getRoute() <= n)
        {
            System.out.println(b[x].getDriver());
        }
    }
}
```

**[6 marks]**

*Question D1 continued*

- (e) *Award marks as follows up to [3 marks max].*  
*Award [1 mark] for a diagram with title, variable and method sections;*  
*Award [1 mark] for a variable section with three suitably defined variables;*  
*Award [1 mark] for a method section with suitably defined methods, corresponding to the defined variables;*

Driver
String: first String: last String/Integer: empNumber
setFirst(String: first) setLast(String: last) setEmpNumber(String/Integer n) String getFirst() String getLast() String/Integer getEmpNumber() String toString() // optional

*[3 marks]*

- D2.** (a) *Award [1 mark] for a suitable definition, for example:*  
 Encapsulation means having private variables;  
 Variables not accessible from outside the class;  
 Methods and variables are all included in the class definition;

*Award [1 mark] for relating to an example from the Bus class, such as:*  
 Class Bus/BusRoute has private (instance) variables;  
 Class Bus/BusRoute has int and start as private variables;  
 Class Bus/BusRoute has setter and getter methods to access the variables;  
 Class Bus/BusRoute has public methods to access the private variables;

*[2 marks]*

- (b) *Award [1 mark] for each possible disadvantage up to [2 marks max].*  
 It is a complex process, unsuited to small scale problems;  
 It is not suitable for problems involving direct access to hardware;  
 Design principles are very abstract/hard to implement for beginning programmers;  
 The world (of buses, for example) does not always divide up neatly into objects that can be programmed/encapsulated;

*[2 marks]*

*continued ...*



*Question D2 continued*

- (c) Award **[1 mark]** for each benefit and **[1 mark]** for a valid example relating to the benefit and a programming team up to **[4 marks max]**.

Work can be split up among programming teams (so that) programmers can work on classes (independently);  
Each team member does not need to know the internal details of a class in order to use/extend it;  
Any class already written can be re-used by any other member;  
*etc.*

**[4 marks]**

- (d) Award marks as follows up to **[3 marks max]**.

Award **[1 mark]** for three instance variables of the correct type (String/Logical/Numeric);  
Award **[1 mark]** for three setter methods with the correct and corresponding parameter types (String/Logical/Numeric);  
Award **[1 mark]** for three getter methods with the correct and corresponding return types (String/Logical/Numeric);

*Example answer:*

BusStop
String: name
Boolean: shelter
Integer/Real: distance
setName(String: first)
setShelter(Boolean shelter)
setDistance(Integer/Real distance)
String getName()
Boolean hasShelter()
Integer/Real getDistance()
String toString() // optional

**[3 marks]**

- (e) Award **[2 marks]** for a suitable data structure and a further **[2 marks]** for details, including sample data.

*Example answer:*

**Array**

The `BusRoute` class could include;

An array of `BusStop` instances;

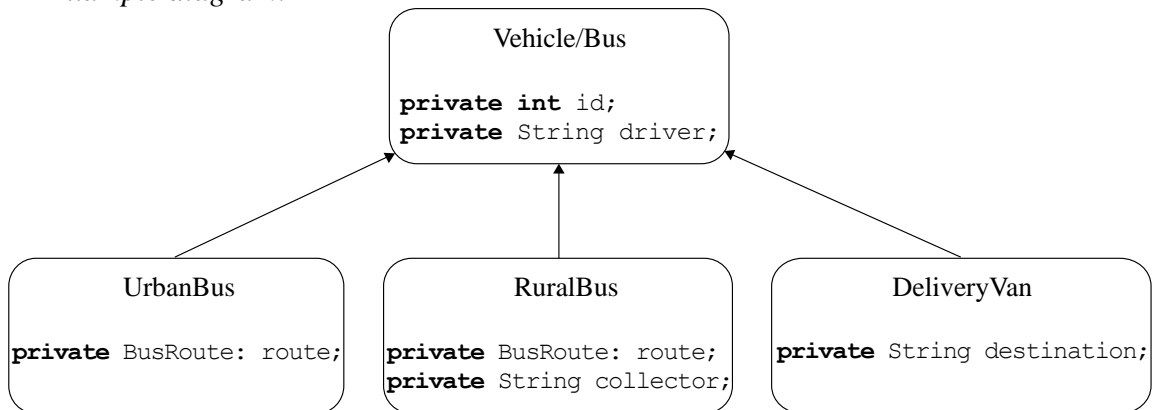
```
BusStop[] stops = new BusStop[10];
stops[0] = new BusStop("Musgrave Hill", true, 1.5);
```

For a code fragment such as the above, award **[1 mark]** for using the structure correctly and **[1 mark]** for suitable data samples.

**[4 marks]**

- D3. (a)** Award marks as follows up to **[8 marks max]**.  
Award **[1 mark]** for a common superclass;  
Award **[1 mark]** for a common id field;  
Award **[1 mark]** for a common driver field;  
Award **[2 marks]** for three sub-classes (arrows not critical, but some form of connection should be shown or labelled, e.g. “extends”) if clear, award **[1 mark]** only for a worthy attempt;  
Award **[2 marks]** for the additional collector field in the RuralBus;  
Award **[2 marks]** for the different destination field in the DeliveryVan;  
Award **[1 mark]** for BusRoute in the buses but not in the DeliveryVan class;

Example diagram:



**[8 marks]**

- (b) Award **[1 mark]** for identifying an advantage (may be implied) and **[1 mark]** for explaining what each advantage is. Award a further **[2 marks]** for a discussion of why this advantage is given by inheritance. Accept only the first reason where multiple advantages are presented.

Enhanced re-use;  
Common variables in the superclass;  
And therefore common validation methods/coding for these;  
No need to maintain code in two or more places;  
And therefore less chance of errors in code;

Maintainability;  
Since variables and related code are not replicated;  
Among different classes;  
If a change is required;  
Only one class need be re-written/maintained;  
And therefore less chance of errors in code;

**[4 marks]**

- (c) Award up to **[3 marks max]**.  
Each subclass of (vehicle);  
Implements this method;  
But returns a different value;  
(For example) the UrbanBus/RuralBus/DeliveryVan returns 0/1/2;

**[3 marks]**

**D4. (a)** *Award up to [3 marks max].*

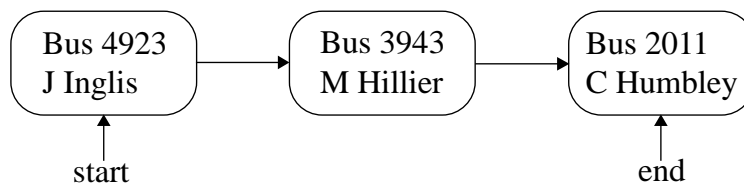
A queue is a first in first out structure;  
 Buses/objects should not be inserted in the middle of a queue;  
 Logically the first bus to leave/enter the queue will be the first to arrive/leave the queue;  
 Random access will not be required to a particular instance, making a queue more suitable than an array;

[3 marks]

**(b)** *Award marks as follows up to [3 marks max].*

*Award [1 mark] for three objects clearly representing a bus by some identifier (number or driver);*  
*Award [1 mark] for objects in the correct sequence as represented by arrows or otherwise;*  
*Award [1 mark] for labelled start and end of queue;*

*Example diagram:*



[3 marks]

**(c)** *Award marks as follows up to [4 marks max].*

*Award [1 mark] for correct return type of Boolean;*  
*Award [1 mark] for correct test for pos less than queue size;*  
*Award [1 mark] for correct test for pos > 0;*  
*Award [1 mark] for correct return value;*

*Example answer:*

```

private boolean removeBus(int pos)
{
    if ((pos < busQueue.size()) && (pos >= 0))
    {
        busQueue.remove(pos);
        return true;
    }
    else
    {
        return false;
    }
}
    
```

[4 marks]

**(d)** *Award up to [3 marks max].*

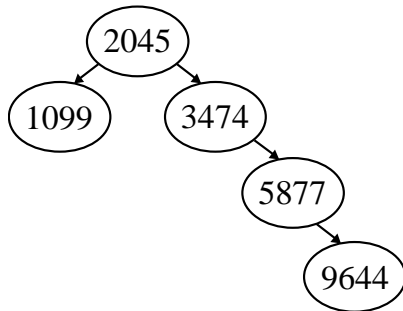
A binary tree has pointers to left and right nodes;  
 Nodes can be ordered;  
 Such that smaller values are placed to the left/right of a node;  
 Which reduces search time;  
 To  $O(\log(n))$ ;

[3 marks]

*continued ...*

*Question D4 continued*

- (e) Award **[1 mark]** for each correctly placed node.



**[5 marks]**

- (f) A recursive algorithm uses stack space which is finite an overflow error may occur;  
Recursive algorithms are often harder to write and understand;

**[2 marks]**

**Total: [65 marks]**

---