Markscheme

November 2020

Computer science

Higher level

Paper 2

Subject details: **Computer science HL paper 2 markscheme**

Mark allocation

Candidates are required to answer **all** questions for **one** of the four options. Total 65 marks.
Maximum 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	- In the case of an "identify" question, read all answers and mark positively up to the maximum marks. Disregard incorrect answers. - In the case of a "describe" question, which asks for a certain number of facts *eg* "describe two kinds", mark the first two correct answers. This could include two descriptions, one description and one identification, or two identifications. - In the case of an "explain" question, which asks for a specified number of explanations *eg* "explain two reasons …", mark the first two correct answers. This could include two full explanations, one explanation, one partial explanation *etc*.

Option A —Databases

1.   (a)   *Award [2 max].*
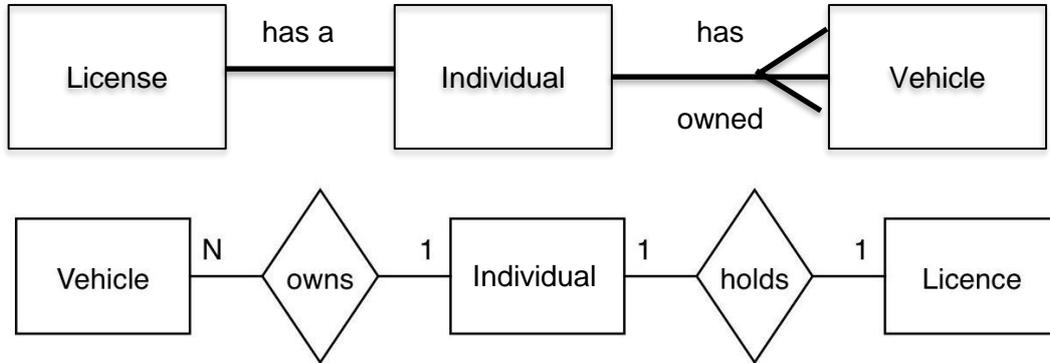
Accept either diagram. License, Individual, and Vehicle correctly positioned
Award [1] for 1 to 1.
Award [1] for 1 to m.

(b)   *Award [3 max].*
The order that a date is written varies by country, *eg* dd/mm/yyyy in UK, mm/dd/yyyy in US;
Some dates, such as 01 June 2018 (01/06/2018) could be confused with 06 January 2018 (06/01/2018);
leading to the incorrect information being stored and consequent decisions, *eg* license expired at the wrong time due to incorrect age;

The Date/String datatype will make analysing slower;
integers can be analysed more quickly than Date/String datatypes;
because functions will be used to split/extract the data;

The Date datatype will make validating data more difficult.
because it is easier to validate separate integer fields;
rather than one entire data that will need to be split by functions;

(c)   *Award [6 max].*

Normalisation
Medical information is only stored if the person has a medical condition / some people will have no medical information stored / some people may have more than one medical conditions;
this requires an additional table / normalisation requires a separate table / one to many relationship;
leaving in the person table will increase storage capacity / empty fields take up space;

Privacy
All employees will need access to the person table and that would include sensitive medical information;
Protects people's privacy / data needs to comply with the Data Protection Act / private medical data seen by non-authorised personnel may cause harm to the licence applicant / potentially result in legal action;
having confidential data in a separate table allows that table to be only available to people with a certain permission level;

Updates
Person table is also likely to be less permanent and need updating more often;
medical information rarely gets updated so is more permanent;
the person table may be useful in other applications, as it offers a way of identifying citizens/the medical data is not likely to be used by other applications;

Mark as [3] + [3]

(d)   *Award [4 max].*
May lead to update/deletion anomalies;
Address changes may result in duplicate addresses;
so letters / fines may go to the wrong address;
Storing data multiple times wastes storage space;
and may slow down data retrieval / data entry;

(e)   *Award [4 max].*
Accept any suitable example

Police access
It is a legal requirement for the VDLA to give access to the police;
eg a speeding vehicle / an accident / linked to a police investigation;
so the police will need to look up the details of the person that owns the car;
facial scanning software may be cross-referenced with the car owner's driving license photo;
so that police have a way to check the identity of the person driving the car;

Insurance company access
A person taking out car insurance signs a consent form to give VDLA access to their records;
this would allow the insurance company to build up a profile *eg*, check for driving offences, see how long that person had owned a car;
thus, a quote could be created quickly / minimal effort;
provides proof of driver eligibility (*ie*, not serving a driving ban);

Medical access
Doctors may require access to an accident victim's records;
To check blood type etc;

Mark as [2] + [2]

2. (a) *Award **[1 max]**.*
A logical unit of work that is executed in full or not at all;
A transaction is a single logical operation that comprises of a sequence of database operations;
A transation satisfies the ACID (Atomicity, Consistency, Isolation, Durability) properties;

(b) *Award **[2 max]**.*
The database would store details of concerts, tickets sold, customers;
An information system would access the database to present data it in a way that aids informs managers / aids decision making (*eg* charts of ticket sales);

Accept any other reasonable example.

(c) *Award **[3 max]**.*
Durability ensures that transactions are saved permanently and do not accidentally disappear or get erased;
Billetmania does not need to worry the transaction being lost even in the event of power loss, crashes, or errors;
A seat in a theatre is guaranteed even if the database crashes;

(d) *Award **[4 max]**.*
Concurrency allows multiple users try to book tickets at the same time / complete a transaction at the same time;
Concurrency prevents access by more than one user to the same row/record;
Concurrency uses row locking;
Select seats are locked for a short period of time to ensure that they aren't double booked/overwritten;

3.   (a)   *Award [1 max].*
A record is one row in a table;
All of the fields relating to item of information;

   (b)   *Award [3 max].*
Fields and Table: Vineyard and NameOfWine FROM Wine;
Criteria: StockQty >= 25 AND StockQty <= 35;   *note: allow >25, <35*
Description: Fruit is extracted from the field.
(Do not award the mark if Description = Fruity)

SELECT Vineyard, NameOfWine FROM Wine
WHERE (StockQty >= 25 AND StockQty <= 35)
AND Description LIKE '*Fruity*';

Accept other versions including:

SELECT Vineyard, NameOfWine FROM Wine
WHERE (StockQty BETWEEN 25 AND 35)
AND Description LIKE '*Fruity*';

   (c)   (i)   *Award [2 max].*
All three points needed for [2]. Two points needed for [1].
UnitPrice * StockQty;
Totalprice named;
WINE table named;

Accept versions similar to the one shown below:

SELECT UnitPrice * StockQty AS totalprice
FROM Wine;

   (c)   (ii)   *Award [2 max].*
A derived field is created / does not exist in the table/ is temporary;
So the rules of normalization do not apply / is not affected by duplication / redundancy;
No new dependencies are created;

(d)   *Award **[6 max]**.*

Example of solution (accept different names for tables)

Vineyard (<u>Vineyard</u>, Region)
Wine (<u>WineID</u>, Wine, Vineyard(fk), Year, Flavour, APV, UnitPrice)
Stock (<u>StoreID</u>, <u>WineID</u>(fk), StockQty)

Vineyard table *eg* Vineyard (<u>Vineyard</u>, Region)
Award **[1]** for primary key. Either Vineyard or VineyardID.

Wine table *eg* Wine (<u>WineID</u>, Wine, <u>Vineyard</u>(fk), Year, Flavour, APV, UnitPrice)
Award **[1]** for primary key WineID or composite Wine/Vineyard
Award **[1]** for identifying the foreign key (Vineyard)
Award **[1]** for splitting the three description fields

Stock table *eg* Stock (<u>StoreID</u>, <u>WineID</u>(fk), StockQty)
Award **[1]** for composite key StoreID/WineID or Store/Wine/Vineyard
Award **[1]** for WineID or equivalent key shown as foreign key

Note: *Should candidates provide other reasonable solutions, please contact your team leader.*

(e)   *Award **[2 max]**.*
Relationship 1-1 might not exist;
Which means a single-field PK might not uniquely identify a record;
So a composite key is needed made up of 2 or more fields;
Allow any suitable example;

4. (a) *Award [2 max].*
   Data warehouse involves historical processing (*eg* trends) of information / timestamped whereas databases involves the day-to-day processing / data from one region;
   OLAP systems in a data warehouse are used by analysts and senior police officers whereas OLAP systems in a database are used by clerks and police officers;
   Data warehouse focuses on *information out* whereas a database focuses on *data in*.
   Data warehouses stores considerably more data than a database / TBs compared to GBs of data / greater variety of data / data from a variety of sources;
   Data warehouses tend to be unnormalised. Databases are normalised;

   (b) *Award [4 max].*

   Decision making tools
   Single manageable structure to support decision making;
   With advanced OLAP tools to assist in data searching / data mining tools;

   Increased amount of data
   Draws upon data stored in multiple databases / cross reference data from multiple police departments;
   Allow complex queries/searches across a number of areas (e.g. cross reference with traffic violations or non-criminal data);

   Speed to access data.
   Real time access to data / Improves speed of data retrieval;
   Instead of contacting other departments for data, PRPD in any department can perform the searches themselves;

   Time Variance
   Ability to view data over a period of time;
   Which can highlight trends;

   (c) *Award [4 max].*
   Data to be extracted would need to be identified;
   Data would need to be converted into a standardised format / ETL;
   Encryption of data would need to be setup / certification to approve uplinks;
   Imported data would need to be time-stamped;
   Fast internet connection at each of the database points;

   (d) *Award [5 max].*
   Classification uses predictive methods whereas clustering uses descriptive methods;
   Typically in classification you have a set of predefined classes and want to know which class the object belongs whereas in clustering it tries to group a set of objects based on some relationship between them;
   Classification uses supervised learning and clustering uses unsupervised learning;
   Classification uses labelled samples and clustering uses unlabelled samples;
   An advantage of clustering over classification is that it is adaptable to changes and helps single out useful features that distinguish different groups/makes no assumptions about the number and form of any groups that emerge;
   Both cluster analysis and classification typically output results in the form of diagrammatic representation of the objects being analysed (*eg* series of plots);
   The results of both methods can be used to make predictions and to decide on the importance of different data variables in making the predictions;
   Classification requires prior knowledge but cluster analysis does not;

(e)   *Award [5 max].*
Link analysis is used to highlight the relationships between databases/objects, which in this context might be different crimes;
It begins by deciding what constitutes a link (association) between two objects;
For example, the police department might link individuals based on the number of times their name appears in the same crime investigation, or link crimes based on their location, type, nature of the victim;
Link analysis discovers new relations / new patterns of interest / checks the similarities in the data sets / finds anomalies where know patterns are violated;
Similarities between crimes may be identified even if they initially appear unrelated (*eg* victims of a particular crime type were all born on the same day);
Studying the links can reveal patterns, including groups and outliers. So the police department might be able to identify a "gang" of associates (groups) or a crime that was thought to be related but wasn't (outlier);
Link analysis is based on a branch of mathematics called graph theory, which represents relationships between different objects as edges in a graph;

Option B — Modelling and simulation

5. (a) *Award [3 max].*
House_Type: char/string/Boolean;
House_Num: integer;
Profit: fixed-point decimal/currency/float/double;

(b) *Award [6 max].*
Award [1] for including HOUSE_TYPE *and* HOUSE_NUM *as input*
Award [1] for calculating total revenue;
Award [1] for calculating the number of days for each project
Award [1] for calculating the labour costs
Award [1] for calculating material costs
Award [1] for calculating the profit

REVENUE: *Award [1] for Sales_price * House_Num based on the If statement*
NUM_DAYS: *Award [1] for the correct number of days * house_num.*
LABOUR_COSTS: *Award [1] for NUM_DAYS * 2500*
MATERIAL_COSTS: *Award [1] for checking the HOUSE_TYPE and then multiplying HOUSE_NUM by the correct material cost.*
PROFIT: *Award [1] for deducting all of the costs from the revenue.*

Example answer:

	A	B	C
2	**HOUSE_TYPE**	**HOUSE_NUM**	**REVENUE**
3	s	10	=IF(A3="s",220000*B3,IF(A3="l",400000*B3, 0))

	D	E	F
2	**LAND_COST**	**NUM_DAYS**	**LABOUR_COSTS**
3	500000	=IF(A3="s",17*B3,IF(A3="l",23*B3,0))	=2500*E3

	G	H
2	**MATERIAL_COSTS**	**PROFIT**
3	=IF(A3="s",100000 *B3,IF(A3="l",190000 * B3,0))	=C3-D3-F3-G3

Alternative solution:

	A	B
1	House_Type is small	TRUE
2	House_Num	10
3		$/1000
4	Total_revenue	=IF(B1,220,400)*B2
5	Land_cost	500
6	Total_labour_cost	=IF(B1,17*2.5,23*2.5)*B2
7	Total_material_cost	=IF(B1,100,190)*B2
8	Total_profit	=B4-B5-B6-B7

Note: accept cell references such as B3 instead of variables.

(c)   *Award [4 max].*

Award [2 max] for each variable tested.
Award [1] for the type of test and [1] for example test data.

Type test for HOUSE_TYPE
Normal: s or l / small or large; Abnormal: m / medium.
If Boolean has been used Normal: True or False; Abnormal: "small"

Range test for NUM_HOUSES
Normal HOUSE_NUM = 5 / Abnormal HOUSE_NUM = 12 / Extreme HOUSE_NUM = 6 or
HOUSE_NUM = 10

Accept any testing that relates to:
- *normal data // Note: an example will get this mark*
- *extreme data*
- *abnormal data.*

(d)   *Award [5 max].*
Award [1] for reading the profit, number of days from csv file / cell range.
Award [1] for adding the number of days to the date.
Award [1] for extracting the month from the date.
Award [1] for calculating interest and rates.
Award [1] for deducting the costs from the profit and displaying result.

Example pseudocode:
```
import FILE.CSV As SS
read PROFIT, NUM_DAYS from SS
END_DATE = #01/01/2021 + NUM_DAYS
NUM_MONTHS = MONTH(END_DATE)
COSTS = (NUM_MONTHS * 500) + (NUM_MONTHS * (400000 * .01))
output (PROFIT – COSTS)
```

Alternative pseudocode:
```
input Profit, NumDays from spreadsheet
endDate = date(dateval("1/1/21") + NumDays)
monthsTaken = month(endDate)
interest = 400000 * 0.01 * monthsTaken
output "Net profit = $", (Profit – interest)
```

6.   (a)   (i)   *Award **[2 max].***
A real-time simulation runs at the same rate as the actual physical system;
The change in the glacier size occurs at the same rate as the real change / a virtual glacier will be the same size as the real glacier after the same period of time;

A real time simulation operations it can mean that it happens without a delay/action carried out before the next input;
Changes to the simulation inputs will be displayed immediately;

   (a)   (ii)   *Award **[2 max].***
Abstraction removes specific detail that will not the accuracy of the simulation / climate change variables that have little influence;
Abstractions helps create the model / reduces the complexity of the model so that it works effectively;
Symbols may be used to represent patterns (*eg* temperature change);

   (b)   *Award **[4 max].***
Sea levels are based on so many factors (*eg* CO_2 emissions, cattles); that it is difficult to accurate predict all of these variables;

Global events (*eg* nuclear war, volcanic erruptions);
may cause for an under-prediction;

Unknown variables (confounding factors) may be affecting rising sea levels;
Population growth may increase or decrease and this may impact upon emissions;

Accept any reasonable argument

   (c)   *Award **[6 max].***
*Award **[3 max]** for positive reasons.*
*Award **[3 max]** for negative reasons.*
*Award **[1 max]** for a reasonable conclusion.*

Positive
Making it free means more people may use it;
Educating the public may encourage more people to be environmental friendly (*eg* boycott plastics);
This may lead to pressure on governments to reduce emissions;
May provide valuable data for scientists who don't have access to NASA resources;

Negative
Users may not fully understand the simulation results and may misinterpret findings;
May change the perception of the accuracy of the simulation software / Users may not value predictions when the simulator seems so simple;
Predictions that don't reflect the real change may do damage to public perception of climate change;

7. (a) (i) *Award [1 max]*
A visualisation is graphical software technique that creates an image / diagram / graph to communicate data;

(a) (ii) *Award [1 max].*
2D map of a coastline;
standard flat map;
cross-section of the land;

(b) *Award [5 max].*
Visualizations provides a quick way of seeing where the sea levels are now and where they will be if the glaciers melt;
People can understand a visual image more easily than raw data;
No need to interpret figures so time is saved;
A visual of land being claimed by the water has greater emotional impact that raw data;
Public / politicians are more likely to be influenced by a visualization than raw data;

Accept other reasonable reasons.

(c) *Award [6 max].*
Wireframe model is created with polygons;
The 3D rendering process automatically converts a 3D wireframe model into 2D images on a computer;
2D images are combined to create a 3D image;
3D renders may, or may not, use photorealistic effects;
It is likely that scanline rendering would be used because it uses less processing / system resources than ray tracing;
it is unlikely that ray tracing would be used because of the time taken for rendering;

8. (a) *Award [4 max].*
No need to anticipate and code in advance the reactions that NPCs need to display;
since the supervised training will ensure the characters can adapt their responses;

Save programmers development time;
because NPCs can be used in similar RPG games;

Provides greater game realism;
NPC characters would evolve to behave more like real people;

NPCs would remember previous encounters with players;
thus, altering their gameplay rather than running through a loop;

Mark as 2 + 2

(b) *Award up to [6 max].*
Start with a set of "units" (here, NPCs) with some predetermined code/behaviour with added randomness / genetic algorithms work successively towards a solution from a starting point;
Apply a process of selection to the units, so that the probability of survival depends on their success against some predetermined criteria;
Supervised learning has training data with a known set of responses to a known set of problems;
The fitness function (defined by the programmers) limits the adaptability of a given population of NPCs to the skills or tasks defined by the fitness function;
A (*fitness) function* is used to measure the effctiveness of a solution / offspring (respawned NPCs) are evaluated using a fitness function;
The best solutions are retained / worse solutions are discarded / the most adaptive NPCs will remain in the population and others removed (die);
The retained solutions are mutated (randomness) to generate another set of solutions / retained NPCs are mutated and respawned / recloned;
This process is repeated until the best solution is identified / repeated until the most adaptable NPC model is identified;

Accept other reasonable reasons or examples.

(c) *Award [6 max].*
The NPC makes a random set of moves;
Records the location/identity of objects relative to its current position;
Building up a map of identifyable surrounding objects;
This is repeated until all objects in the space have been placed in distance and direction from a starting point;
The NPC would learn how to react to the object (*eg* swim in a river / walk on a path);
NPC would learn where it could walk and where it could not (*eg* river would need to be avoided);
NPCs that die could be regenerated from ones that had avoided dying up to that point in the game;
NPC would adapt to a changing environment *eg* a house built in a game;

(d) *Award [4 max].*
Chatbots are designed to output language in response to human language;
Most NPC conversations are scriptive speeches / repetitive…
so using chabots will add more realism to the game;
NPCs could verbally communicate in script or sound form;
There would still need to be some hardcoded responses that relate to the game storyline;
Natural language processing (NLP) could be applied to chatbot conversations;

Option C — Web science

9. (a) (i) *Award [2 max].*
Award [1] for the text within the <p> tags (My First Page large font/Welcome to the Faculty in a smaller font;
Award [1] for the text box and correctly labelled button;

My First Page

Welcome to the Faculty

[   ] clickme

OR

Award [2 max].
Initially it displays:
My First Page – in large font;
Welcome to the Faculty – in normal font;
A text box + a button labelled "clickme";
A Then if you type in the text box and then press the button, it displays what is typed between "Welcome to the" and "Faculty" – e.g. type "Science" and it will display;
"Welcome to the Science Faculty";

Note:Students either show the output on the web-browser or describe what would be displayed on the web-browser.

(a) (ii) *Award [3 max].*
The button's onclick method is called, i.e. setVar();
This first sets the variable subject to the value of the HTML element with id kippers, i.e. the text box;
and then sets the inner text of the element with id subj, i.e. the span, to "Pugh", so that the display is "Welcome to the Pugh Faculty";

(b) *Award [4 max].*
Award [1] for identifying correct validation check.
Award [1] for describing the validation check.

Presence check:
It happens in JavaScript in the web page on the **client**;
because this will save the server handling many transactions/to avoid sending a null value that will need to be handled by the server;

Lookup check
It happens on the **server**;
because the database is stored on the server rather than on the client's computer;

Mark as [2] + [2]

(c)   *Award [3 Max].*
        The website certificate is used to initiate sessions between the user's browser and the
        server;
        The web server sends the browser a copy of its SSL certificate;
        If the browser trusts the certificate, it sends a message to the web server;
        The web server sends back a digitally signed acknowledgement to start an SSL encrypted
        session;

(d)   *Award [4 Max].*
        The Common Gateway Interface (CGI) is used to provide interactivity to web
        applications/enable forms to be submitted;
        It uses a standard protocol that acts as an intermediary between the CGI program and the
        web server;
        The CGI allows the web server to pass a user's request to an application program;
        And then the forwarded data is received to the user's browser;

(e)   *Award [4 Max].*
        Cloud computing is based on sharing computing resources over the internet;
        Cloud computing can be offered as a service to individuals and companies (SaaS) which
        means they do not have to maintain the infrastructure;
        Cloud computing and applications are more scalable than client-server architectures;
        Client-server architecture does not need to be operating over the internet, but instead can be
        limited to a local network;
        Client-server architectures may require more maintenance of the infrastructure by the
        organization who are using it;
        cloud computing is typically more resilient than client server, because the services offered
        are more widely distributed, rather than being concentrated on one server or on a small
        number of alternative servers;

10. (a)   *Award **[2 Max]**.*
The internet is a global network of interconnected computers / a network of networks;

The World Wide Web is software / a service that runs on the hardware of the internet and provides access to content / a collection of pages that can be accessed through hyperlinks / a way of accessing and sharing the information that is held on the internet in webpages;

The World Wide Web uses the http protocol. This is only one of the many protocols used by the internet;

E-mail, File Transfer Protocol (FTP), and instant messaging services are part of the internet but not of the web;

(b)   *Award **[4 Max]**.*
Open standards provide a publicly available specification for a specified task;
This is an agreed set of parameters that enable interoperability and/or compatibility to occur;

Using Open standards means that you are not subject to a governing body with its own agenda/self-interest;
Thus, you can be confident that you won't be subject to fees/bias;

Open standards promote interoperability;
This enables the various devices to communicate with each other;

Open standards advocates also argue that openness encourages better and more secure systems;
this is because more people are able to analyse the standards and resulting software and no-one has a proprietary interest in suppressing knowledge of problems to keep sales up.

Mark as [2] + [2].

(c)   *Award **[2 Max]**.*
The HITS Algorithm ranks the page based on a combination of its importance as a hub and an authority;
The PageRank Algorithm ranks the page by counting the number and quality of links to a page to determine the relative importance of the website;

(d)   *Award **[3 Max]**.*
Meta tags are included in the header of a web-page which are available to a web-crawler and give information about the page that it could make use of;
When the web-page is crawled, a copy of the HTML is replicated in the search engine database;
When a user enters text into a search the search engine retrieves the data indexed from the web-page;
And the search engine ranks and displays the content (in order of relevance);

11. (a) *Award [2 Max].*
The surface web is the part of the web that can be reached by a search engine whereas the deep web cannot;
The deep web may consist of dynamic content such as the result of database queries or may be protected proprietary content;

(b) *Award [3 Max].*
The dark web uses a layered encryption system;
Data is routed through a large number of intermediate servers;
Which means it is almost impossible to decrypt the information layer by layer;
With the result that the user's details are practically untraceable, and their anonymity can be maintained;

(c) *Award [3 Max].*
A port is used to facilitate the communication between a computer and an application;
Certain ports such as Port 21 (FTP), 23 (Telnet) and 80 (HTTP) are reserved;
Every time a port is opened on a computer it provides access to that computer;
This means that the security of that computer may be potentially compromised every time a new port is opened;
Or port conflicts may occur when more than one application tries to use a specified port;

(d) *Award [6 Max].*
World Wide Web has enabled citizens to communicate easily and for 'ordinary' citizens to express their opinions;
Therefore, the ability to publish is not confined to certain 'privileged' groups such as broadcasters and journalists;
The World Wide Web has to a large degree given access to common resources to all citizens globally who have access to the Internet;
However, it can be argued that the evolution of the World Wide Web has led to a greater centralization of power, for example the digital oligarchs (Microsoft, Google, Amazon, Apple, Facebook);
This centralization has led to a reduction in democracy as the digital oligarchs have an increasing stranglehold over the lives of their 'digital subjects' through the aggregation, analysis and monetarization of their data;
There are still issues with a lack of digital democracy, for example, many citizens may not have access to the World Wide Web, either through income, geography or necessary skills;

12. (a) *Award [4 Max].*
In graph theory, a graph is a set of nodes (also called vertices) that can be connected through edges;
In a web graph each web page is represented by a vertex;
In a web graph each hyperlink is represented by a directed edge;
Therefore, the web graph provides a representation / model of the connectivity of the World Wide Web;
The direction of the edge is from the page with the hyperlink to the page that is the target of the hyperlink;

(b) *Award [4 Max].*
Although the growth in the number of web-pages on the World Wide Web has increased exponentially, the diameter of the World Wide Web increases linearly;
The increase in routing hubs may lead to a more efficient routing algorithm being used;
And in combination with the increasing speed of web-pages loading;
Means that despite the significant increase in the number of web-pages, despite there being an increase in the number of hops required to navigate from one web-page, there may be no discernable increase in time for this process to occur;

(c) *Award [4 Max].*
The semantic web is based on data and enables computer systems to understand the meaning of information and its relation to other pieces of information;
this increased usability is dependent on the ability of the different computers to process the intended meaning of content better;
however, giving information more expressive power may come at an expense of usability;
for example, a language such as RDF can be highly expressive power, but requires the user to specify relations of information to other information;
in some cases, it may be beneficial to create shared databases (i.e. increase the usability);
but, there may be other cases where the increased expressivity leads to a decrease in usability;

(d) *Award [2 Max].*
embedded, devices are integrated into their environment;
context aware, these technologies can adapt to the needs of the user;
machine learning, these technologies can learn from previous behaviors;

(e)   *Award [6 Max].*
Collective intelligence is the intellectual by product of a group of people working together;
This ability to harness the benefits of collective intelligence have been made possible by the development of the internet;
Collective intelligence can be used for a variety of purposes, for example linked to social media (Facebook and Twitter), gaming sites (Online Multi-player games), Scientific research (Large hadron Collider and Wikipedia;
Collective intelligence benefits from the open sharing of resources and may allow for a more rapid evolution of ideas than 'traditional' methods;
Collective intelligence may not have a central authority, so the user created information may be considered to be factual when it is not;
Collective intelligence may lead to a disruption in the way that knowledge is stored or classified, *ie* the development of folksonomies in place of ontologies;
The lack of a formal organizational structure may lead to the acquisition of the collective intelligence being potentially piecemeal or haphazard;
There may be no owner of this additional knowledge, so it may not be stored or maintained in the most efficient manner;

Option D — Object-oriented programming

13. (a) *Award [2 max]*.
Static means it is the same for all instances of the class;
because it is contained in the class rather than in an instance of the class/object / it is defined at the class level;
Static means that less memory is taken up;
as only 1 memory allocation it created for all instances rather than 1 per instance;

[2]

   (b) *Award [3 max]*.
Award *[1]* for aggregation (accept Composition also);
Award *[1]* for An instance of ParkingArea can contain Vehicles;
Award *[1]* for A single ParkingArea can contain between 0 and 300 instances of Vehicle;

[3]

   (c) *Award [2 max]*.
To distinguish between 2 variables with the same name;
Allows the local/instance variable to be set to the parameter/class variable;

Note: *Award [1] if the incorrect technical terms for variables are given but the idea of distinguishing between 2 types of variables with the same name is expressed.*

[2]

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

Award *[1]* for correctly declaring the variable as type Vehicle;
Award *[1]* for using the correct constructor;

eg
```
Vehicle v=new Vehicle("X1234567");
```

[2]

     (ii) *Award [2 max]*.
Award *[1]* for calling the correct method on the correct instance variable name;
Award *[1]* for using the constant correctly, ie. Vehicle.BLACK;

```
v.setColour(Vehicle.BLACK);
```

Note: *accept the constant prefixed by the name of the instance instead of the class*

[2]

14. (a)   *Award [6 max].*

Award [1] for any correct loop (starting at 0 and while <vehicles.length);
Award [1] for checking if vehicles[index]==null;
Award [1] for adding vehicle instance to correct index, eg. vehicles[i]=v;
Award [1] for stopping loop at the point when vehicle is added (ie. doing it efficiently);
Award [1] for returning the index of the array when vehicle is added;
Award [1] for returning -1 if not added;

Example 1:

```
public int addVehicle(Vehicle v) {
    boolean found = false;
    int i = -1;
    while (!found && i < vehicles.length-1) {
        i = 0;
        if (vehicles[i] == null) {
            vehicles[i] = v;
            found = true;
        }
        i = i + 1
    }
    if (found) {
        return i;
    } else {
        return - 1;
    }
}
```

Example 2:

```
public int addVehicle(Vehicle v) {
    for (int i = 0; i < vehicles.length; i++) {
        if (vehicles[i] == null) {
            vehicles[i] = v;
            return i;
        }
    }
    return - 1;
}
```

[6]

(b)   *Award [4 max].*
Inheritance is when the behaviour of a class is transferred to another class;
whereas aggregation is when a class is contained inside another and used;

Aggregation is useful for modelling a real life situation of containment aka. "has a" (*eg* contents of a box);
whereas Inheritance is useful for modelling real life situation that one entity is a particular (more specific) type of another entity, aka. "is a";

While both can achieve the same result, Aggregation (composition) can lead to more robust and safer code;
whereas inheritance is a more flexible solution but prone to errors due to the behaviour of subclasses being changed;

Inheritance allows for code reuse / lowers maintenance cost;
whereas aggregation does neither of the above;

Can have more than 1 aggregation;
but only 1 superclass (in Java);

Inheritance allows for code reuse as in the creation of sub-classes;
whereas with aggregation each class must be written out in full;

Mark as [2] + [2].

[4]

(c)   *Award [4 max].*
Award [1] for all four classes shown;
Award **[1]** for showing the links in the correct place (any arrow or line is acceptable);
Award **[1]** for distinguishing the relationship ParkingArea - Vehicle;
Award [1] for distinguishing the relationships Car/Motorbike -Vehicle;
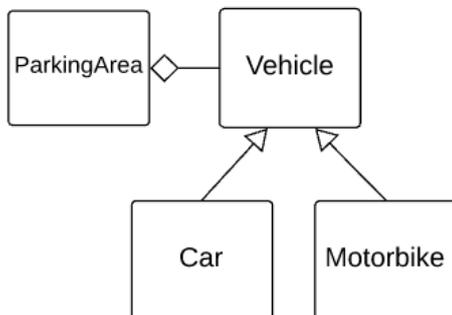
[4]

(d) (i) Award **[1]** for returning admin fee if <=5 hours;
         Award **[1]** for returning 0 if >5 hours;

```
public double pay(int hours) {
    if (hours <= 5) {
        return ADMIN_FEE;
    } else return 0;
}
```
[2]

   (ii) *Award [2 max].*
         Award [1] for for super.pay(hours);
         *Award [1] for returning result of correct formula (ie. adding on hourlyFee*hours);*

         Example:
```
public double pay(int hours) {
    return super.pay(hours)+hourlyFee*hours;
}
```
[2]

(e) *Award [2 max]*
     Cars and Motorbikes are also Vehicles as they extend the Vehicle class / Vehicle is the superclass and Car/Motorbike are subclasses;
     The subclasses can be cast as/considered to be/referred to as Vehicle which is its superclass;
     As Cars and Motorbikes inherit from Vehicles they are also Vehicles and can therefore be stored in an array of type Vehicle;
     Arrays must be declared as 1 type, and the state and behaviour of Vehicle is what both Car and Motorbike have in common. Therefore they can be referenced as Vehicles and stored in an array of Vehicles, even though they are subclasses;
[2]

15. (a) *Award [5 max]*
Award **[1]** for declaring both the variable for total cars and the variable for total motorbikes at class level (private);
Award **[1]** for initializing the variables, either by calling a mutator method, while declaring, or in the constructor;
Award **[1]** for deciding whether the car or motorbike is 50th or 60th in the daily additions;
Award **[1]** for printing the message if so;
Award **[1]** for incrementing the appropriate counter for the total cars / motorbikes added based on the value returned by `getKind();`

Note: Do not award marks for simply adding mutator or accessor methods for the motorbike / car totals.

[5]

(b) *Award [3 max]*

Car:
Award [1] for Car count mod 50 is 0, `coffee voucher;`
Award [1] for Car count mod 50 is n, n `coffee vouchers;`

Motorbike:
Award [1] for Motorbike count mod 60 is not 0, `No coffee vouchers;`
Award [1] for Motorbike count mod 60 is n, n `coffee voucher;`

Non Car/Motorbike
Award [1] for Aany amount of instances of the Vehicle class which are neither Car nor Motorbike, `No coffee voucher`

Note: do not award a mark for first test case above which is already give in question stem

Examples:

Test data	Vouchers printed
29th Car	`No coffee voucher`
50th car	`coffee voucher`
6th Motorbike	`No coffee voucher`
60th Motorbike	`coffee voucher`
Other Vehicle instance added (neither C nor M)	`No coffee voucher`

[3]

(c)   *Award [6 max].*
        Award [1] for correct loop;
        Award [1] for checking if not null;
        Award [1] for checking registration using the getRegistration().equals() method;
        Award [1] for setting vehicles[i] to null;
        Award [1] for returning the reference (not the null element of the array), by making another variable;
        Award [1] for returning null if it is not found;

Example 1:

```
public Vehicle removeVehicle(String registration) {
for (int i = 0; i < vehicles.length; i++) {
    if (vehicles[i] != null &&
vehicles[i].getRegistration().equals(registration)) { //question - why
do it in this order?
        Vehicle leaving = vehicles[i];
        vehicles[i] = null;
        return leaving;
    }
}
return null;
}
```

Example 2:

```
public Vehicle removeVehicle(String registration)
  { Vehicle v = null;
    boolean found = false;
    int i = 0;
    while (!found && i < Vehicles.length)
    { found = Vehicles[i] != null &&
            registration.equals(Vehicles[i].getRegistration());
      if (found)
      { v = Vehicles[i];
        Vehicles[i] = null;
      }
      else i = i + 1;
    }
    return v;
  }
```

[6]

16. (a) (i) *Award **[1 max].***
Faster to access (ie. O1 instead of On);
Arrays take less memory than lists (as they do not store pointers);
Mirrors real life situation;

[1]

(ii) *Award **[1 max].***
Static, therefore memory wasted when not being fully utilised;
Inability to grow, should further spaces be added, ie. If Parking area grows;

[1]

(b) *Award **[3 max].***
The array would need to be searched using a linear search which is inefficient O(n) whereas a binary search is much more efficient O(log n);
A linear search for a car in this parking area could require a maximum of 200 comparisons (ie. it would look at the first row of every column) whereas a binary search would only require 8 (ie log 200 base 2);
A linear search for a motorbike in this parking area would require a maximum of 400 comparisons (ie. it would need to look at both rows of every column) whereas a binary search would only require 7 (log 100 base 2);
In a binary search, each comparison would eliminate half of the remaining possible matches, whereas in a linear search it only eliminates 1;
A binary search has the registration plate as its key and therefore it doesn't require accessing the field of the vehicle object for each comparison, whereas in a linear search, each comparison would be slower because it would require getting the vehicle object's registration field using an accessor method;
A binary search is always faster than a linear search on a large data set. As the police may be searching for many cars and in many parking areas, their searches will be on large set of data;

[3]

(c) *Award **[5 max].***
Create a binary tree to store vehicle objects;
Loop around all of the elements in the 2 dimensional array (e.g. using inner and outer loop);
Skip empty locations;
Take the first vehicle found and add as the head of the binary tree;
Take all subsequent vehicles found and do the following;
Get the node's vehicle registration;
Find the appropriate insertion point in the binary tree by navigating from the root and comparing the registration to the key of the current node;
When a leaf node is reached compare the registration with the vehicle to be inserted and insert it as the left or right child node, using its registration as the key;
Continue until the end of the vehicle array has been reached;

[5]

17. (a) *Award **[6 max]***
        *Award **[1]** for looping around the main stack using isEmpty() method;*
        *Award **[1]** for using "while" loop in both loops (note: not a "do while"!);*
        *Award **[1]** for breaking out of outer loop when car is removed and other cars are replaced;*
        *Award **[1]** for correctly removing cars before target car is found;*
        *Award **[2]** for replacing cars in same order by popping off temporary stack;*

Example 2:

```
public void staffRemoveCar(Stack stack, String reg) {
Stack temp = new Stack();
while (!stack.isEmpty()) {
    Vehicle v = (Vehicle) stack.pop();
    if (v.getRegistration().equals(reg)) {
        // correct car is now removed
        // return others
        while (!temp.isEmpty()) {
            stack.push(temp.pop());
        }
        break; // don't keep repeating
    }
    temp.push(v);
}
}
```
                                                                              [6]

Example 2:

```
  public static void staffRemoveCar(Stack<Vehicle> stack, String reg)
  { Stack<Vehicle> temp = new Stack<Vehicle>();
    boolean found = false;
    while (!found && !stack.isEmpty())
    { Vehicle curr = stack.pop();
      found = curr.getRegistration().equals(reg);
      if (!found) temp.push(curr);
    }
    while (!temp.isEmpty()) stack.push(temp.pop());
  }
```

(b)   *Award **[4 max]**.*
      Award [1] for an initial valid comparison and [1] for a valid expansion.

Memory
Linked lists can deal with memory issues more efficiently than arrays;
Arrays are static so will have a maximum size and cannot grow beyond that;
Memory will be allocated for maximum size at instantiation so while it is not full memory is wasted Links are dynamic and can add/remove memory as required- no memory is wasted;

Implementation
No particular advantage is gained by the use of either structure;
As additions and removals (push and pop) work on one end of the stack, navigation speed (ie O1) is not very helpful; Linked lists can adjust the pointer at the end of the list (can use a tail pointer) so slow navigation is not an issue;

Errors
The use of arrays is more likely to lead to run-time errors;
As arrays can reach their maximum limit whilst linked lists are only limited in size by the specifications of the CPU;

[4]