

No part of this product may be reproduced in any form or by any electronic or mechanical means, including information storage and retrieval systems, without written permission from the IB.

Additionally, the license tied with this product prohibits commercial use of any selected files or extracts from this product. Use by third parties, including but not limited to publishers, private teachers, tutoring or study services, preparatory schools, vendors operating curriculum mapping services or teacher resource digital platforms and app developers, is not permitted and is subject to the IB's prior written consent via a license. More information on how to request a license can be obtained from <http://www.ibo.org/contact-the-ib/media-inquiries/for-publishers/guidance-for-third-party-publishers-and-providers/how-to-apply-for-a-license>.

Aucune partie de ce produit ne peut être reproduite sous quelque forme ni par quelque moyen que ce soit, électronique ou mécanique, y compris des systèmes de stockage et de récupération d'informations, sans l'autorisation écrite de l'IB.

De plus, la licence associée à ce produit interdit toute utilisation commerciale de tout fichier ou extrait sélectionné dans ce produit. L'utilisation par des tiers, y compris, sans toutefois s'y limiter, des éditeurs, des professeurs particuliers, des services de tutorat ou d'aide aux études, des établissements de préparation à l'enseignement supérieur, des fournisseurs de services de planification des programmes d'études, des gestionnaires de plateformes pédagogiques en ligne, et des développeurs d'applications, n'est pas autorisée et est soumise au consentement écrit préalable de l'IB par l'intermédiaire d'une licence. Pour plus d'informations sur la procédure à suivre pour demander une licence, rendez-vous à l'adresse <http://www.ibo.org/fr/contact-the-ib/media-inquiries/for-publishers/guidance-for-third-party-publishers-and-providers/how-to-apply-for-a-license>.

No se podrá reproducir ninguna parte de este producto de ninguna forma ni por ningún medio electrónico o mecánico, incluidos los sistemas de almacenamiento y recuperación de información, sin que medie la autorización escrita del IB.

Además, la licencia vinculada a este producto prohíbe el uso con fines comerciales de todo archivo o fragmento seleccionado de este producto. El uso por parte de terceros —lo que incluye, a título enunciativo, editoriales, profesores particulares, servicios de apoyo académico o ayuda para el estudio, colegios preparatorios, desarrolladores de aplicaciones y entidades que presten servicios de planificación curricular u ofrezcan recursos para docentes mediante plataformas digitales— no está permitido y estará sujeto al otorgamiento previo de una licencia escrita por parte del IB. En este enlace encontrará más información sobre cómo solicitar una licencia: <http://www.ibo.org/es/contact-the-ib/media-inquiries/for-publishers/guidance-for-third-party-publishers-and-providers/how-to-apply-for-a-license>.

Computer science
Higher level
Paper 2

Monday 11 November 2019 (morning)

1 hour 20 minutes

Instructions to candidates

- Do not open this examination paper until instructed to do so.
- Answer all of the questions from one of the options.
- The maximum mark for this examination paper is **[65 marks]**.

Option	Questions
Option A — Databases	1 – 4
Option B — Modelling and simulation	5 – 8
Option C — Web science	9 – 13
Option D — Object-oriented programming	14 – 17

Option A — Databases

- 1. *One-to-7* is an international organization that works with teachers and other educators. One department within this organization provides an online forum for teachers to discuss ideas for lessons and to share resources.

In order to access this forum teachers are required to submit the following information, which will be stored in a table in the database.

Figure 1: Online form to register personal details to the *One-to-7* forum

Personal Details Fields marked with an * are required

Personal Details

Title*	First Name*	Last Name*
<input type="text" value="None"/>	<input type="text" value="First Name"/>	<input type="text" value="Last Name"/>
Email*	School Email	Phone*
<input type="text" value="yourname@example.com"/>	<input type="text" value="yourname@example.com"/>	<input type="text" value="(33) 123456789"/>
Date of Birth (dd/mm/yyyy)*	Country of Residence*	Country of Birth*
<input type="text"/>	<input type="text"/>	<input type="text"/>
Place of Birth*		
<input type="text"/>	<input type="submit" value="Submit"/>	

- (a) Identify **one** reason why the teacher’s name has been split into two fields. [1]
- (b) Outline **one** reason why there may be concerns about the amount of personal information that is requested. [2]

Once the submit button on the online form has been selected, the personal data is input into the database.

- (c) Outline why the transaction needs to be atomic in the context of this scenario. [2]

Once the teacher is registered they can post comments on the forum.

- (d) Explain how transactions are managed to ensure isolation when registered teachers add comments to a discussion thread on the forum. [3]

The *One-to-7* database is managed by the database administrator (DBA).

- (e) Identify **two** tasks that are carried out by the database administrator (DBA). [2]

The DBA is considering using the email address as the primary key, but is concerned that many of the 250 000 educators who are registering for this online forum may have more than one email address. Users may create duplicate accounts, deliberately or accidentally, by using different email addresses as usernames.

- (f) Explain the factors that would need to be considered in using a composite primary key instead of only using the email address. [4]

(Option A continues on the following page)

(Option A continued)

2. Taxonomy is the part of science that focuses on naming and classifying or grouping organisms. Swedish naturalist, Carl Linnaeus, developed the current hierarchical classification system of taxonomic classification in the 1700s. This classification is hierarchical so one class has many orders, and one order has many families.

Birds can be classified under the Linnaean hierarchical taxonomy as “Aves”. In one system of classification, Aves has 23 orders, one of which is the order Ciconiiformes, which holds six families within it.

- (a) Construct an entity-relationship-diagram that shows the relationships between the class, order and family. [2]

Spoonbills constitute one of the families within the order Ciconiiformes. Data can be collected about spoonbills, such as length of the bill or beak, and is included in the database.

- (b) (i) Outline, using an example, how data validation can ensure that data is entered into the birds database correctly. [2]

- (ii) Outline, using an example, how data verification can ensure that data is entered into the birds database correctly. [2]

- (c) Distinguish between data and information when entering the length of a spoonbill’s beak in the database. [2]

- (d) Identify **two** characteristics of a conceptual schema. [2]

- (e) Explain why the use of data modelling is critical to the success of a database, such as the one used in this scenario. [3]

- (f) Explain, using the example described in the scenario, why referential integrity is important in databases. [3]

3. *Trucking On*, a logistics and haulage company based in Marseille, France, keeps track of its lorries and operations in a database.

Lorries are made up of the truck and a trailer. Trucks are usually coupled with the same trailer but sometimes trailers are moved to a different truck for part of the year.

Each driver only uses one truck.

- (a) Identify **two** characteristics that make a database unnormalized. [2]

(Option A continues on the following page)

(Option A, question 3 continued)

The following LORRY table shows information about the lorries.

Truck ID	Truck make	Energy source	Driver	Tel	Trailer ID	Trailer space	Coupled_ from	Coupled_ to
TK1	Fodel	Diesel	Pierre	010238397	TR1	60	01/01/18	31/12/18
TK2	Fodel	Diesel	Elizabeth	010228293	TR2 TR9 TR2	60 30 60	01/01/18 19/06/18 01/08/18	18/06/18 31/07/18 31/12/18
TK3	Viene	Diesel	Jinyeong	010298491	TR3	60	01/01/18	31/12/18
TK4	Viene	Electric	Thierry	010238399	TR4 TR9	80 30	01/01/18 01/08/18	31/07/18 31/12/18
TK5	Fodel	Diesel	Chiho	010238398	TR5	60	01/01/18	31/12/18
TK6	Fodel	Diesel	Akwete	010233392	TR6	60	01/01/18	31/12/18
TK7	Viene	Diesel	Chryssy	010238341	TR7	40	01/01/18	31/12/18
TK8	Viene	Electric	Jacque	010966396	TR9 TR8	30 60	01/01/18 31/05/18	30/05/18 31/12/18

Trailer space is measured in cubic metres.

The table can also be represented as:

LORRY (TruckID, Truckmake, Energysource, Driver, Tel, TrailerID, Trailerspace, Coupled_from, Coupled_to)

(b) Outline why data redundancy may be a problem in the LORRY table. [2]

The table has been split into two tables following the rules of normalization. The resulting two tables are shown with the primary key underlined:

LORRY (TruckID, Truckmake, Energysource, Driver, Tel)

TRAILER (TrailerID, Trailerspace, Coupled_from, Coupled_to, TruckID)

(c) Identify the steps to create a query to find the names and telephone numbers of the drivers who drive lorries that have more than 60 cubic metres of trailer space. [4]

(d) Normalize the database to 3NF. Use the same format as shown, ensuring that primary keys are clearly indicated by underlining them. [7]

(Option A continues on the following page)

(Option A continued)

4. The collection, storage and sharing of data is becoming increasingly important for organizations who have a choice about which type of database to use to store their data. Two examples of database types are relational and object-oriented.

- (a) Explain **two** advantages of using a relational database rather than an object-oriented database. [4]

The 2016 US presidential election was seen to be a victory for data analytics. Companies that specialize in analytics use data warehouses.

- (b) State **two** characteristics of a data warehouse. [2]
- (c) Outline why data needs to be transformed before it can be loaded into the data warehouse. [2]
- (d) Outline why opinion poll data and other election data are timestamped when added to the data warehouse. [2]
- (e) Outline why analytics companies use link analysis. [2]
- (f) Outline why analytics companies use deviation detection. [2]

Once data has been loaded into a data warehouse it can be mined. The use of data analytics is believed to have been important to the outcome of the US election campaign.

- (g) Discuss whether the advantages of data mining techniques in this scenario outweigh the disadvantages. [6]

End of Option A

Option B — Modelling and simulation

5. Investment planning is something that many people need to be aware of. Planning for your future requires research because bad choices can be very costly.

Interest paid on money invested is usually in the form of compound interest. The formula to calculate compound interest is:

$$T = P * (1 + r)^n$$

T is the total value of the investment,

P is the principal sum invested,

r is the interest rate per time period converted to a decimal (for example, 5% is 0.05),

n is the number of time periods.

- (a) Calculate the total value of the investment after two years if the principal sum of \$30 000 is invested. The yearly interest rate is 10% and this rate is compounded at the end of each year. [2]

An additional \$1000 is added to a principal amount of \$30 000 at the end of each month. The monthly interest rate is 0.5% and this rate is compounded at the end of each month.

- (b) Outline, using a diagram or otherwise, a method of calculating the total value of the investment after 12 months. [5]

Each month, tax is calculated on the monthly profit at a rate of 25% when the investment total (T) is \$40 000, or below. However, when the investment total (T) is above \$40 000, the tax rate is 40%.

The tax is calculated at the end of each month after interest has been added. A running total of the tax is kept and only deducted from the investment total (T) at the end of the year.

- (c) Construct an algorithm to calculate the fund value at the end of each month. This algorithm should also calculate the total value of the investment after the tax has been deducted after 12 months. [6]

Many investment companies offer alternative investment schemes and use modelling to set the rates of interest.

- (d) Explain why the investment company would use modelling when setting the rates of interest. [4]

(Option B continues on the following page)

(Option B continued)

6. A supermarket chain has identified four potential locations to build a distribution centre. The distribution centre will store goods and distribute them to a large number of supermarkets in the region.

A computer simulation will be run to determine the best location to build the distribution centre.

- (a) Outline **two** reasons for using a computer simulation in this scenario. [4]
- (b) (i) Identify **two** data inputs for this computer simulation model. [2]
- (ii) Identify **two** criteria that may be used to determine the best location to build the distribution centre. [2]
- (iii) Discuss how the simulation model will use the data inputs in (b)(i) and the criteria identified in (b)(ii) to generate recommendations. [6]
- (c) Outline why the use of a computer simulation may not be beneficial. [2]

(Option B continues on the following page)

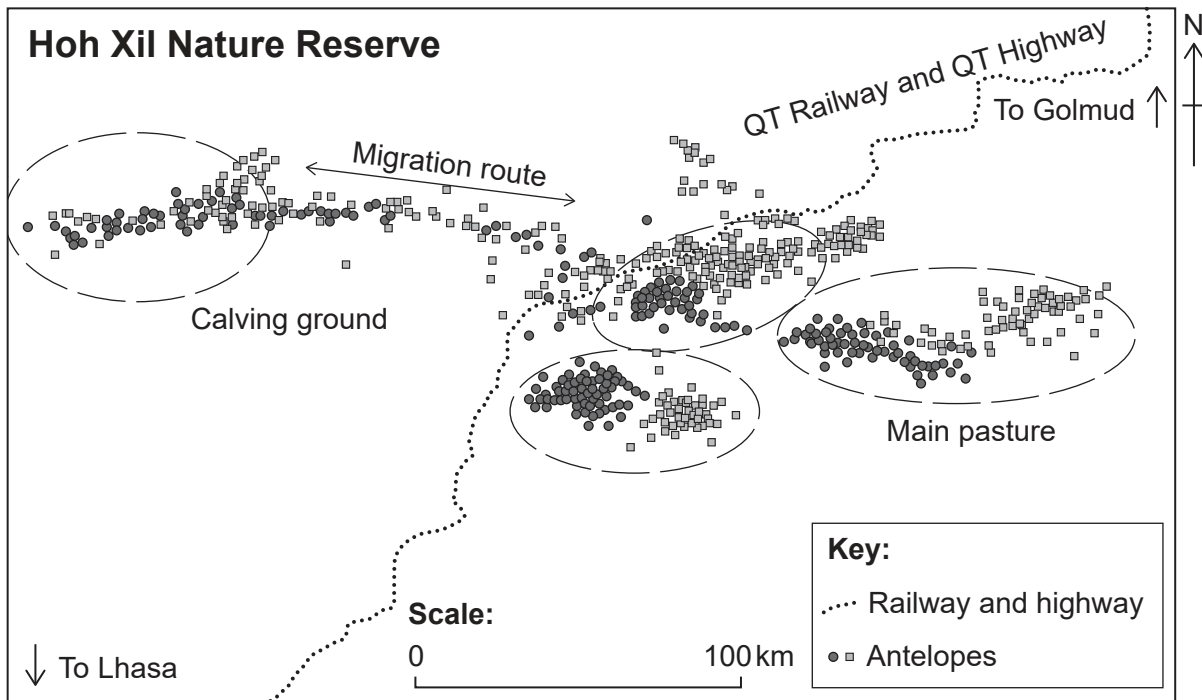
Turn over

(Option B continued)

- 7. The QT Railway and QT Highway have recently been built to connect Qinghai with Tibet. This has made the mountainous region of Tibet accessible to the rest of China.

The endangered Tibetan antelope migrates over 300 km each year to give birth to a single calf. The QT Railway and QT Highway cross this migration route, see **Figure 2**.

Figure 2: Map of the study site and habitat of Tibetan antelope



[Source: adapted (simplified redraw of the nature reserve data map) from Karina Manayeva, Buho Hoshino, Hiromasa Igota, Takashi Nakazawa and Ganzorig Sumiya, 2017. Seasonal migration and home ranges of Tibetan antelopes (*Pantholops hodgsonii*) based on satellite tracking. *Int. J. Zool. Res.*, **13**: 26–37. © 2017 Karina Manayeva et al.]

The building of the QT Railway and QT Highway presents potential dangers to both the antelopes and the public travelling through this nature reserve.

To understand more about the antelopes' migration patterns, some antelopes have tracking devices attached. Using the information from these tracking devices, a 2D visualization model can be developed and updated in real time.

- (a) Identify **four** items of data that would be included to create the 2D visualization model. [4]
- (b) Explain why a 2D visualization model would be used rather than a 3D visualization model in this scenario. [4]
- (c) Explain why the development of a visualization model was necessary in this scenario. [4]

(Option B continues on the following page)

(Option B continued)

8. Optical character recognition (OCR) is a method where printed text or handwritten text is converted to machine-encoded text.

(a) Outline **one** problem that may lead to printed text characters not being detected correctly. [2]

Recognizing handwritten characters presents more of a problem because people have different handwriting styles.

For example, the digitized handwritten letter X in **Figure 3** does not exactly match the digitized letter X in **Figure 4**.

Figure 3: Digitized handwritten letter X

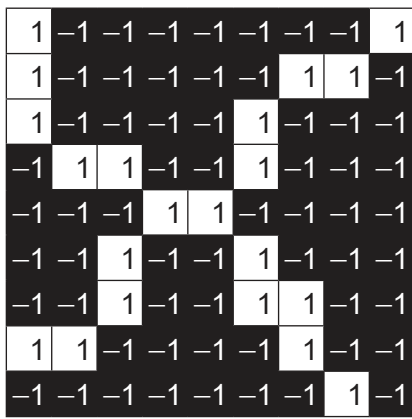
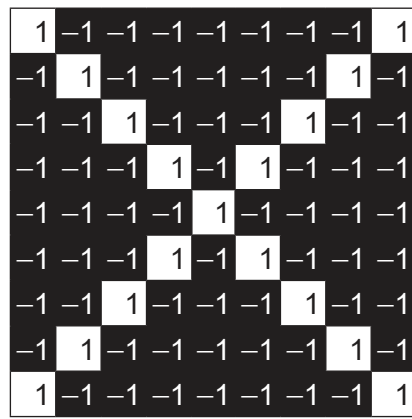


Figure 4: Digitized letter X



Artificial neural networks (ANNs) can be used to assist with recognizing handwritten characters that do not exactly match the expected pattern.

(b) Outline why an ANN can be used to overcome the challenges outlined in this scenario. [2]

(c) Explain how ANN pattern recognition techniques are applied to ensure that the handwritten letter X in **Figure 3** is recognized as a letter X. [4]

Predictive text, where the computer predicts the next word in the sentence, can be programmed to utilize a neural network.

(d) Identify **two** features that would be required by the ANN to predict the next word in the sentence. [2]

The sentence "The child is feeling" is entered into an application that uses predictive text and three options are suggested: *better*, *like*, *a*. Upon entering the two characters "hu" the word *hungry* is suggested.

(e) Explain how the application uses a neural network to suggest suitable words. [6]

(f) Outline **two** potential problems with training the ANN to suggest appropriate words. [4]

End of Option B

Turn over

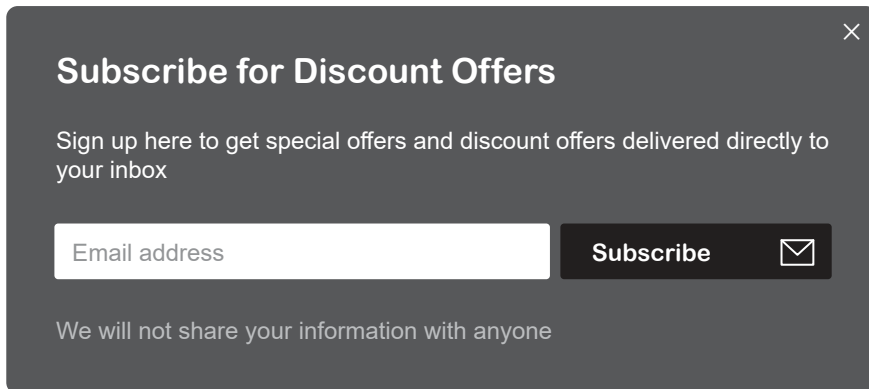
Option C — Web science

9. *OrderYourFoods* is an online company that allows customers to order food and get it delivered to their location. It operates in a number of cities in India.

On the home page of the company’s website www.orderyourfoods.com, the customer has to enter the city name and the address the food needs to be delivered to. Based on this information, the website displays a list of restaurants.

- (a) (i) State the protocol used in the company’s website. [1]
- (ii) Explain the importance of protocols on the World Wide Web. [4]
- (b) The website has a pop-up advert that encourages customers to subscribe for discount offers, see **Figure 5**.

Figure 5: An example of a pop-up advert



(Option C continues on the following page)

(Option C, question 9 continued)

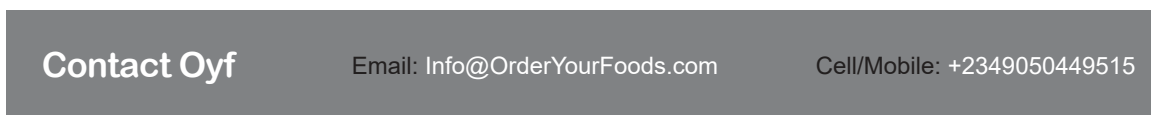
The following code is executed when the customer clicks on the subscribe button.

```
<? php
$to = 'info@orderyourfoods.com';
$from = '$_POST['email']';
$headers = "From: " . $from . "\r\n";
$subject = "New subscription at OrderYourFoods";
$body = "New user subscription: " . $_POST['email'];
if( filter_var($_POST['email'], FILTER_VALIDATE_EMAIL) )
{
    if (mail($to, $subject, $body, $headers, "-f " . $from))
    {
        echo 'Your e-mail (' . $_POST['email'] . ') has been added to our
        mailing list!';
    }
    else
    {
        echo 'There was a problem with your e-mail (' . $_POST['email'] . ')';
    }
}
?>
```

- (i) State the type of scripting used in the php code. [1]
- (ii) State **one** inbuilt function along with parameters used in the php code. [1]
- (iii) Describe the processing that occurs when the subscribe button is clicked. [3]

The website uses the following method to communicate with the potential customers, see **Figure 6**.

Figure 6: Banner for OrderYourFoods.com



OrderYourFoods currently displays an email address and a phone number to show customers how they can be contacted.

- (c) Identify **one** additional method of online interaction that can be incorporated in this website. [1]

OrderYourFoods is concerned that it never appears close to the top of the list of search results when the term “food delivery” is typed into a search engine.

- (d) *OrderYourFoods* has decided to use black hat search engine optimization techniques to be more visible in search engines.

Evaluate this decision. [5]

(Option C continues on the following page)

Turn over

(Option C continued)

10. The IB Coordinator of AB World Academy introduces the Extended Essay to the Grade 11 students in January by researching the difference between primary and secondary data on the internet.

Some of the students used the Google search engine (Google.com) and others used the Ask search engine (Ask.com). These search engines gave results.

- (a) Define the term *search engine*. [1]

Google uses the PageRank algorithm and Ask uses the HITS algorithm.

- (b) Distinguish between the principles of these two algorithms. [4]

The IB coordinator uploaded the assignments onto a cloud-based Learning Management Platform.

- (c) Describe the difference between cloud computing and local client-server architecture. [3]

As part of their research, students downloaded images from the internet. Most of the downloaded JPG images were compressed using lossy compression.

- (d) (i) State the alternative type of compression to lossy. [1]

- (ii) Evaluate the advantages and disadvantages for students of using compressed images in their IB Coursework. [4]

A PNG image uses open standards.

- (e) Distinguish between interoperability and open standards. [2]

11. The growth of the Internet has led to increased levels of collaboration between different IB World Schools.

- (a) Explain why a distributed network such as the World Wide Web may act as a catalyst to increase the collaboration between different IB World Schools. [3]

The World World Web consisted of over 1350 million web pages in 2017. Each web page is accessed by typing its uniform resource locator (URL) into the address bar of the browser.

- (b) Explain **one** function of a web browser. [3]

- (c) Describe how a domain name server (DNS) functions. [3]

- (d) Explain how the use of a distributed network, such as the World Wide Web, may lead to copyright and intellectual property issues. [5]

(Option C continues on the following page)

(Option C continued)

12. An international organization, OBI, has three offices in Europe, see Figure 7.

Figure 7: The location of OBI's offices in Europe



[Source: adapted (recoloured, cropped and annotated) Europe Map by Erin Dill 0, www.freevector.com. Under copyright and CC 4.0 licence (https://creativecommons.org/licenses/by-sa/4.0/)]

Office A and Office B report to Office C, the head office.

Office A and Office B collaborate on many projects. Office C is not involved in the collaboration between Office A and Office B.

- (a) Draw a directed graph based on the scenario. [2]
- (b) Describe the role of graph theory in determining the connectivity of the World Wide Web. [3]

OBI has projects that involve large numbers of contributors from all over the world. These projects are managed by OBI staff in Office A.

The contributors often include non-text based information.

- (c) Outline **two** issues that may arise from using a non-text based search. [4]
- (d) OBI has decided to gather information for one project by utilizing collective intelligence. Evaluate this decision. [5]

13. Southend Council in England is considering whether to use Pepper, a humanoid robot, to assist some elderly residents that live in care homes. These elderly residents suffer from memory loss and would rely on Pepper to direct them to and from the local shops. Pepper uses ambient intelligence and is connected to the Internet of Things (IoT).

Explain why Southend Council may have concerns about introducing Pepper for this purpose. [6]

End of Option C

Turn over

Option D — Object-oriented programming

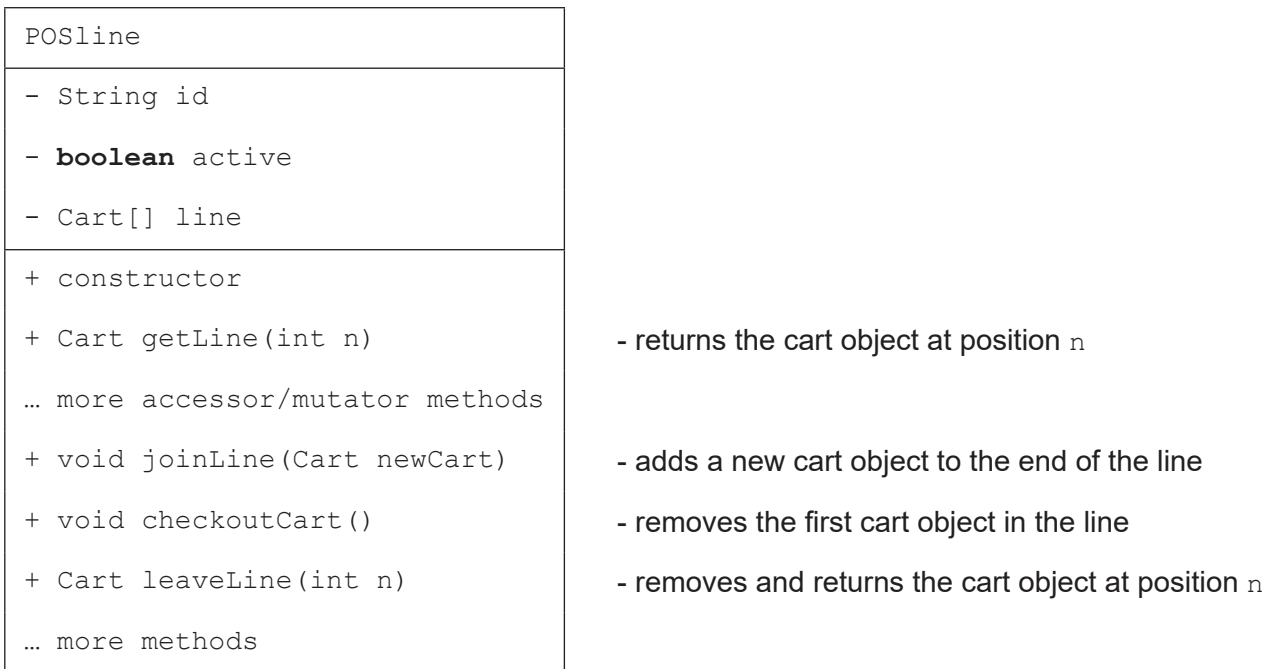
To improve customer satisfaction, a supermarket chain wants to create an object-oriented program (OOP) to simulate the lines of customers at the check-outs in their point of sale (POS) system.

This point of sale (POS) system consists of several check-out counters. After filling their shopping carts with items, customers line up at one of the check-out counters. In most cases, they wait in line until it is their turn to pay.

Three real-world objects are implemented using the following classes:

Cart	represents a customer together with the items they intend to purchase
POSline	represents an individual check-out counter and the line of customers with carts waiting at that checkout
POSsystem	the overall check-out system

The UML diagram for the class POSline is provided below.



- 14. (a) State the relationship between the POSsystem and POSline objects. [1]
- (b) Draw a diagram to show the relationship between the objects POSsystem, POSline and Cart. You are not required to draw a complete UML diagram. [2]
- (c) Define the term *identifier*. [1]
- (d) Distinguish between a class and an instantiation. You must make reference to the UML provided. [3]

(Option D continues on the following page)

(Option D, question 14 continued)

- (e) State the code fragment that instantiates an array `line` of 20 `Cart` objects. [2]
- (f) Construct the method `public void joinLine(Cart newCart)` that adds a `newCart` to `line` in the next empty array location. You may assume that `line` is not full. [3]
- (g) Construct the method `public Cart leaveLine(int n)` that removes the cart at position `n` from `line` by shifting down all the array entries bigger than `n`, before returning the indicated `Cart` object. You may assume that a cart exists at position `n`. [4]

The supermarket chain wants to use this OOP simulation to experiment with different ways of organizing their check-out system. For example, it is possible to have different check-out counters such as cash-only or card only, or to have check-out counters for ten items or fewer.

15. (a) Define the term *inheritance*. [2]
- (b) Explain **one** advantage of the OOP feature “inheritance” with reference to this scenario. [3]
- (c) Describe **two** advantages of using libraries of classes. [4]

It is also possible to have one line that serves a number of check-out counters.

- (d) Outline **one** change needed to the original model to allow this possibility to be implemented. [2]
- (e) Outline **two** reasons why the use of multiple programming teams in different locations may be problematic when developing an integrated software solution. [4]

(Option D continues on the following page)

Turn over

(Option D continued)

16. When the number of customers in the supermarket is low, some check-out counters will close. When there are many customers waiting, a new check-out counter will open.

- (a) Outline why the variable `active` in the UML of the class `POSline` was defined as a boolean data type. [2]

When a new check-out counter opens, some customers from the nearest line will choose to move their carts to this check-out counter.

For this simulation, the assumption is made that every second cart in the old line will move to the new line and the other carts will remain in the original line.

- (b) Construct the code for a method `split` (part of the class `POSsystem`), that takes an existing, non-empty `POSline` as a parameter. It should copy every second cart from the original line into a new line. Afterwards it should delete every second cart from the original line.

An example call in `POSsystem` would be: `POSline number2 = split(number1)`, where `number1` is an existing `POSline`.

You may use any method declared or developed. [8]

As a result of the simulation, the company has decided to create a new POS system for their supermarkets.

There have been discussions about adapting existing open source code when developing modules of the new POS system.

- (c) Describe **two** ethical issues that may arise if modules of the new POS system are developed from open source code. [4]

(Option D continues on the following page)

(Option D continued)

17. The array `line` in `POSline` is now replaced by a singly linked list `list` of `CartNode` objects. The class `CartNode` has been defined as follows.

```
public class CartNode
{
    private Cart myCart;
    private CartNode next;

    public CartNode(Cart aCart)
    {
        this.myCart = aCart;
        this.next = null;
    }
    public Cart getCart(){ return this.myCart; }
    public CartNode getNext(){ return this.next; }
    public void setNext(CartNode nextNode)
    {
        this.next = nextNode;
    }
}
```

- (a) Define the term *object reference*. [2]

The class `POSlist` has been implemented with the standard list methods `addLast` and `removeFirst` that act on `list`.

- (b) Using object references, construct the method `public Cart removeFirst()` that removes the first `CartNode` object from `list`. The method must return the `Cart` object in that node or `null` if `list` is empty. [4]

- (c) Sketch the linked list `list` after running the following code fragment where `cart1`, `cart2`, `cart3`, `cart4` and `cart5` have been instantiated as objects of the class `Cart`.

```
POSlist queueList = new POSlist()
queueList.addLast(cart2);
queueList.addLast(cart1);
queueList.addLast(cart4);
queueList.removeFirst();
queueList.addLast(cart5);
queueList.addLast(cart3);
queueList.removeFirst();
```

- (d) Outline **one** feature of the abstract data structure *queue* that makes it unsuitable to implement customers waiting in line. [2]

(Option D continues on the following page)

Turn over

(Option D, question 17 continued)

A method `leaveList(int n)` is required in the class `POSlist`, similar to the method `leaveLine(int n)` that was added to the class `POSline`.

- (e) Using object references, construct the method `public Cart leaveList(int n)` that removes the `nth CartNode` object from `list`. The method must return the `Cart` object in that node.

You may assume that the `nth CartNode` object exists in the list.
You may use any method declared or developed.

[6]

- (f) Explain the importance of using coding style and naming conventions when programming. [4]

End of Option D
