

© International Baccalaureate Organization 2023

All rights reserved. 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 the prior written permission from the IB. Additionally, the license tied with this product prohibits 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, whether fee-covered or not, is prohibited and is a criminal offense.

More information on how to request written permission in the form of a license can be obtained from <https://ibo.org/become-an-ib-school/ib-publishing/licensing/applying-for-a-license/>.

© Organisation du Baccalauréat International 2023

Tous droits réservés. 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 préalable de l'IB. De plus, la licence associée à ce produit interdit toute utilisation 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, moyennant paiement ou non, est interdite et constitue une infraction pénale.

Pour plus d'informations sur la procédure à suivre pour obtenir une autorisation écrite sous la forme d'une licence, rendez-vous à l'adresse <https://ibo.org/become-an-ib-school/ib-publishing/licensing/applying-for-a-license/>.

© Organización del Bachillerato Internacional, 2023

Todos los derechos reservados. 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 la previa autorización por escrito del IB. Además, la licencia vinculada a este producto prohíbe el uso 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—, ya sea incluido en tasas o no, está prohibido y constituye un delito.

En este enlace encontrará más información sobre cómo solicitar una autorización por escrito en forma de licencia: <https://ibo.org/become-an-ib-school/ib-publishing/licensing/applying-for-a-license/>.

# Computer science

## Higher level

### Paper 2

8 May 2023

**Zone A** morning | **Zone B** afternoon | **Zone C** morning

1 hour 20 minutes

---

#### Instructions to candidates

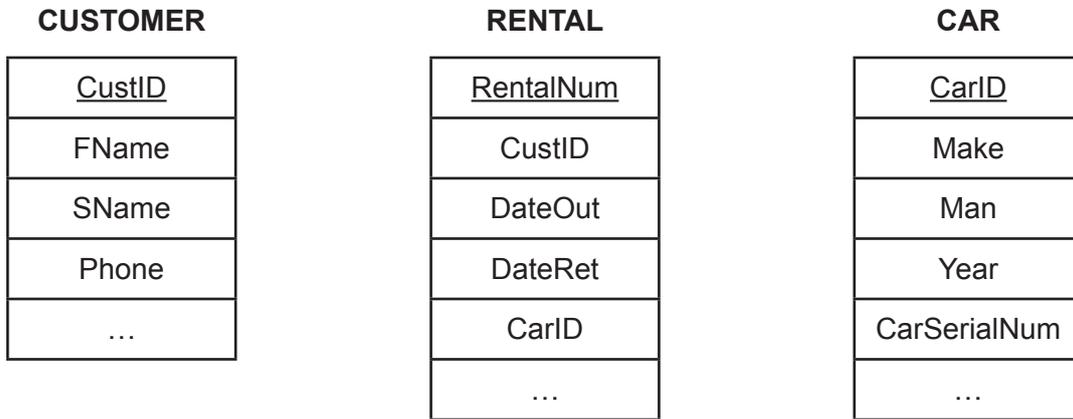
- Do not open this examination paper until instructed to do so.
- Answer all of the questions from one of the options.
- The maximum mark for this examination paper is **[65 marks]**.

Option	Questions
Option A — Databases	1 – 4
Option B — Modelling and simulation	5 – 8
Option C — Web science	9 – 12
Option D — Object-oriented programming	13 – 16

**Option A — Databases**

1. *Quick Rent-a-Car* is a popular car rental company in Chicago. The company stores information about the cars, customers and rental agreements.

The diagram below shows a part of the entity-relationship diagram (ERD) for the *Quick Rent-a-Car* database.



- (a) (i) State the type of relationship between the CUSTOMER table and the RENTAL table. [1]
- (ii) State the name of **one** primary key. [1]
- (iii) State the name of the table that contains foreign keys. [1]

Some of the information held in the three tables in the *Quick Rent-a-Car* database is shown below:

**CUSTOMER**

<u>CustID</u>	FName	SName	Phone	...
4098	Rod	Lever	480-9225-9180	...
6543	Cador	Travolta	708-4567-1012	...
8265	Ella	Hanks	605-6543-1056	...

**RENTAL**

<u>RentalNum</u>	CustID	DateOut	DateRet	CarID	...
00000001	4098	22/01/2020	25/01/2020	WRE2345	...
00000002	6543	22/01/2020	23/01/2020	ELA3489	...
00000003	8265	22/01/2020	26/01/2020	ACC2345	...
00000004	6543	24/01/2020	29/01/2020	WEM6789	...

(Option A continues on the following page)

(Option A, question 1 continued)

**CAR**

<u>CarID</u>	Make	Man	Year	CarSerialNum	...
WRE2345	Pacifica	Chrysler	2018	5Y2SP670X9Z459140	...
DEF4567	UX	Lexus	2017	2CNDL73F456219488	...
ELA3489	ES350	Lexus	2015	1FTFX28L7VNB18489	...
WEM6789	Pacifica	Chrysler	2018	8YTBN54K8CSD2879	...
ACC2345	UX	Lexus	2017	4DFBS43L0MNB3156	...
DHE7623	EF475	Lexus	2016	2GHKU98P1SWY456	...

(b) State the result from the following query: [1]

```
SELECT Make
FROM CAR
WHERE Man = "Lexus"
AND Year = 2016;
```

(c) Outline **two** possible validation checks for the CarID attribute. You may assume that the CarID will always be in the format shown. [4]

(d) Identify the steps to create a query to find the surname (SName) of the customer who rented the car from 22 January 2020 (22/01/2020) until 26 January 2020 (26/01/2020). [4]

(e) Explain why queries would be used to create views of the *Quick Rent-a-Car* database. [3]

Views of the database can also be created by using a query language.

(f) Explain how a data definition language can be used to implement a data model such as the *Quick Rent-a-Car* database. [3]

(g) Explain why maintaining data consistency is important in the *Quick Rent-a-Car* database. [3]

(Option A continues on the following page)

**(Option A continued)**

2. Database recovery is performed when disasters, hardware failures or application problems occur. Database administrators have to make decisions about which type of database recovery strategy to use.
- (a) Identify **one** factor that a database administrator could consider when choosing a database recovery strategy. [1]
  - (b) Describe **two** methods of database recovery that a database administrator can carry out. [4]

Many organizations are concerned about staff having access to sensitive or inappropriate data.

- (c) Describe **two** methods organizations could use to ensure that staff do not have access to sensitive or inappropriate data. [4]
3. Artisan College has the following data about students and the courses they have chosen.

**STUDENT table**

<u>StudentID</u>	FName	SName	MajorID	Major	CourseNum	CourseName
2907	Jacob	Smith	MAT	Mathematics	MAT0011 MAT0027 EGL0010	Discrete Maths Calculus I Classics I
4019	Jane	Paterson	PHI	Philosophy	PHI0010 CS00100	Philosophy Programming 1
5145	Norris	Neeld	EGL	English	SOC0102	Ascent of man
6132	Xavier	Morrison	MUS	Music	MUS0002 SOC0102	Origin of jazz Ascent of man
8966	Samantha	Juarez	EGL	English	EGL0010 EGL0101	Classics I Shakespeare II

The table can also be represented in the following form.

STUDENT (StudentID, FName, SName, MajorID, Major, CourseNum, CourseName)

- (a) Outline why this table is not in 1st Normal Form (1NF). [2]
- (b) Construct the 3rd Normal Form (3NF) of the unnormalized relation shown above. [8]
- (c) Explain the importance of data modelling in the design of a database. [5]

**(Option A continues on the following page)**

**(Option A continued)**

4. *Lacto Dairy Products* has data stored in a data warehouse.

- (a) Define the term *data warehouse*. [2]

*Lacto Dairy Products* uses a network data model.

- (b) Describe **two** characteristics of a network data model. [4]

Data is obtained from a number of different sources. This data needs to be extracted, transformed and loaded into the data warehouse.

- (c) Outline **two** ways Extract, Transform, Load (ETL) processes can be used to clean up data for the data warehouse. [4]

*Lacto Dairy Products* supplies dairy products directly to both wholesalers<sup>1</sup> and retailers<sup>2</sup>.  
*Lacto Dairy Products* have different pricing models for wholesalers and retailers.

- (d) Discuss the advantages and disadvantages of *Lacto Dairy Products* using data segmentation. [5]

Data analytics, such as the analysis of data held in data warehouses, can assist the managers at *Lacto Dairy Products* in their future decision making.

- (e) Explain why data analytics can be used to assist the managers of *Lacto Dairy Products* in their future decision making. [5]

---

<sup>1</sup> wholesaler: a person or company that buys large quantities of goods from various producers or vendors, and resells to retailers

<sup>2</sup> retailer: a person or company that sells goods to the public

**End of Option A**

**Option B — Modelling and simulation**

5. A restaurant is concerned about the increase in its electricity bills as it uses electricity for cooking, refrigeration, heating, air conditioning, air extraction systems and lighting.

The electricity bill is paid every month and is based on a daily charge of \$0.75 with a charge of \$0.20 per unit of electricity used.

The table below shows the electricity usage from the previous year.

Month	Electricity use (Units)	Days in month
January	5800	31
February	5000	28
March	5200	31
April	4800	30
May	4750	31
June	4900	30
July	5500	31
August	6500	31
September	5600	30
October	5960	31
November	5850	30
December	6280	31

A tax of 10% is applied to the total bill.

The restaurant would like to create a computer model of its electricity usage using data from the previous year.

- (a) Define the term *computer model*. [1]
- (b) State **three** variables and their data type that could be used in the computer model. [3]
- (c) Construct the pseudocode that will input the units used in each month and the days in each month to calculate and output the electricity bill for the restaurant in any month **and** the quantity of electricity used per day. [5]
- (d) Outline **one** limitation of this model in helping the restaurant owners to save money on their electricity bills. [2]

**(Option B continues on the following page)**

**(Option B, question 5 continued)**

The model needs to determine the months with the highest and lowest electricity usage.

- (e) State **two** ways that this model could be implemented. [2]
- (f) Construct the pseudocode to provide the restaurant owner with this information. [8]

6. Governments use simulations to help them to decide how their resources will be allocated in the future. This enables them to predict how much income they will require from sources such as taxation.

The simulation needs to consider information such as the size of the population, the income from taxation and the expenditure on healthcare and education.

Information is collected every three years.

- (a) Describe the difference between a model and a simulation in the context of this scenario. [2]
- (b)
  - (i) Outline **one** advantage of a government using a simulation to predict how resources will be allocated in the future. [2]
  - (ii) Outline **one** disadvantage of a government using a simulation to predict how resources will be allocated in the future. [2]
- (c) Describe **two** ways that the collection of data could be improved to enhance the accuracy of this simulation. [4]
- (d) Discuss whether the ethical concerns expressed by citizens about the collection of this data are justified. [5]

7.
  - (a)
    - (i) State one example of 2D visualization. [1]
    - (ii) State one example of 3D visualization. [1]
  - (b) Describe the significance of a key frame in the process of completing 3D visualization. [2]

A computer game makes use of 3D animation and involves game characters that move throughout different scenes.

- (c) Explain the implications of 3D animation on computer resources during the creation of this game. [5]

**(Option B continues on the following page)**

**(Option B continued)**

8. Autonomous vehicles, or self-driving cars, make use of genetic algorithms.

- (a) Outline how a genetic algorithm is iterative. [2]
- (b) Outline the role of the fitness function in a genetic algorithm. [2]

Autonomous vehicles can be controlled by the “drivers” using verbal commands which can be “learned” by either supervised or unsupervised learning.

- (c) (i) Outline **one** benefit of using supervised learning in order for the autonomous vehicle to learn verbal commands. [2]
- (ii) Outline **one** benefit of using unsupervised learning in order for the autonomous vehicle to learn verbal commands. [2]
- (d) Explain the advances in natural language processing that will ensure the autonomous vehicle responds appropriately to verbal commands. [5]
- (e) Explain why there are differences between human and machine learning of a language. [5]
- (f) Describe the use of a chatbot. [2]

**End of Option B**

**Option C — Web Science**

9. Sonia is a student at an international school and is developing her skills in building websites. The computer science teacher at the school directed her to an online resource site that provides tutorials. She provided her with the uniform resource locator (URL) of this site.

(a) Define the term *uniform resource locator* (URL). [1]

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

The site uses the protocol HTTPS.

(c) Identify **two** characteristics of HTTPS. [2]

(d) Distinguish between a protocol and a standard. [2]

The internet protocol (IP) has three basic characteristics: connectionless, best effort and media independent.

(e) Outline what the term connectionless means in the context of the transmission of a page request. [2]

**(Option C continues on the following page)**

**(Option C continued)**

10. Below is a fragment of the code from a web page.



```
$sql = "INSERT INTO GuestsTable (FirstName,Surname,Email,Cellphone)
VALUES ('Sonia','Smith','sonia@smith.com','+61231456789')";

if ($conn->query($sql) === TRUE) {
    echo "New record created successfully";
} else {
    Echo "Error.".$sql."<br>".$conn->error;
}
```

- (a) (i) Describe the process outlined in the code fragment above. [3]
- (ii) Describe how the common gateway interface (CGI) enables the execution of scripts on a server. [3]

It is possible to increase the visibility of a website using search engine optimization.

- (b) (i) Identify **two** techniques used in search engine optimization. [2]
- (ii) Explain why incoming and outgoing links are significant factors in search results. [4]

Search engines frequently undertake web-indexing.

- (c) Describe how a web crawler would index a webpage. [3]

A limitation of commonly used search engines is that they can only access part of the web. They are unable to access the deep web.

- (d) Suggest why the problem of not being able to access the deep web could increasingly become a problem for search engines. [4]

**(Option C continues on the following page)**

**(Option C continued)**

**11.** Many people use online resources like peer-2-peer (P2P) networks to access music and movies from their mobile phones.

- (a) Identify **two** characteristics of a peer-2-peer network. [2]

A user downloads an album **and** an eBook about an artist from a file sharing site. The site uses lossy compression.

- (b) Evaluate the appropriateness of lossy compression for both the album **and** the eBook. [6]

An artist has been encouraged to join a cloud-based photosharing site.

- (c) Identify **two** characteristics of cloud computing. [2]

- (d) To protect their intellectual property the artist considers **two** methods of protecting their work.

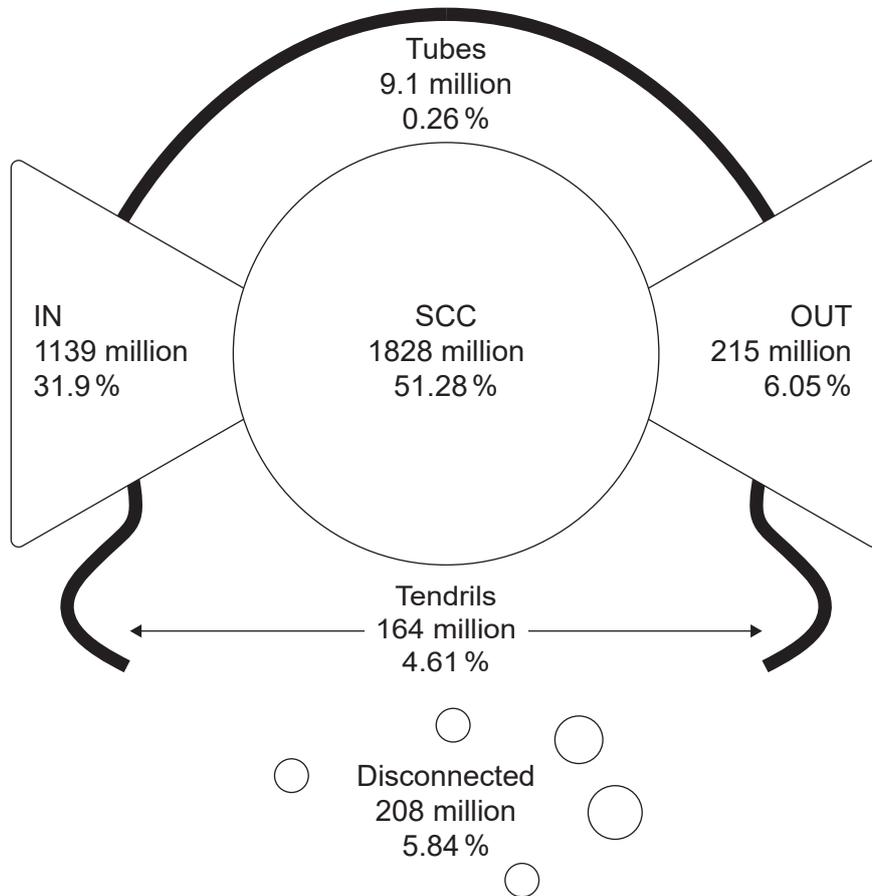
- Requiring the viewer to join the site and protecting the images behind the log in.
- Using a script to disable the “save image” option in the drop down menu.

Evaluate these **two** methods of protecting the artist’s intellectual property. [6]

**(Option C continues on the following page)**

**(Option C continued)**

12. The model below shows the bowtie graph representing the internet. It is based on data from a 2012 hyperlink graph crawl.



The crawl visited approximately 3500 million nodes and discovered 128 000 million edges.

- (a) (i) Define the term *node*. [1]

- (ii) Define the term *edge*. [1]

The strongly connected core (SCC) contains 51.8% of the nodes discovered in this crawl.

- (b) Identify **two** characteristics of a node in the SCC. [2]

**(Option C continues on the following page)**

**(Option C, question 12 continued)**

A directed sub-graph can be used to represent a portion of the World Wide Web (WWW).

- (c) Draw a directed sub-graph representing the structure described below. Use arrows to represent the direction of edges and circles to represent the nodes.

A, B and C are IN nodes.

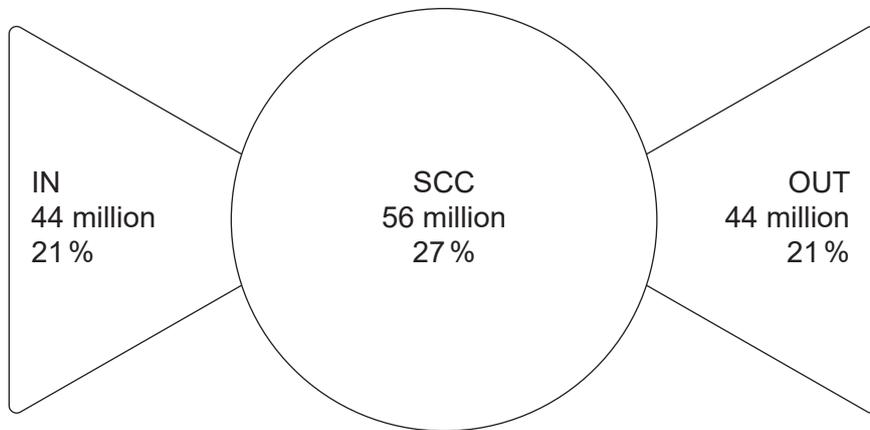
D, E, F, G are nodes located in the SCC.

H and I are OUT nodes.

In addition, node C connects directly to node I.

[4]

In earlier representations of the bowtie model the size of the IN and OUT portions are equal.



The percentage of websites in the IN portion has changed from 21 % in 2000 to approximately 32 % in 2012.

The percentage of websites in the OUT portion has changed from 21 % in 2000 to approximately 6 % in 2012.

- (d) Suggest why the percentages in the IN and OUT portions have changed between 2000 and 2012.

[4]

- (e) Identify **two** characteristics of collective intelligence.

[2]

- (f) Discuss the significance of ontologies **and** folksonomies in the development of the semantic web.

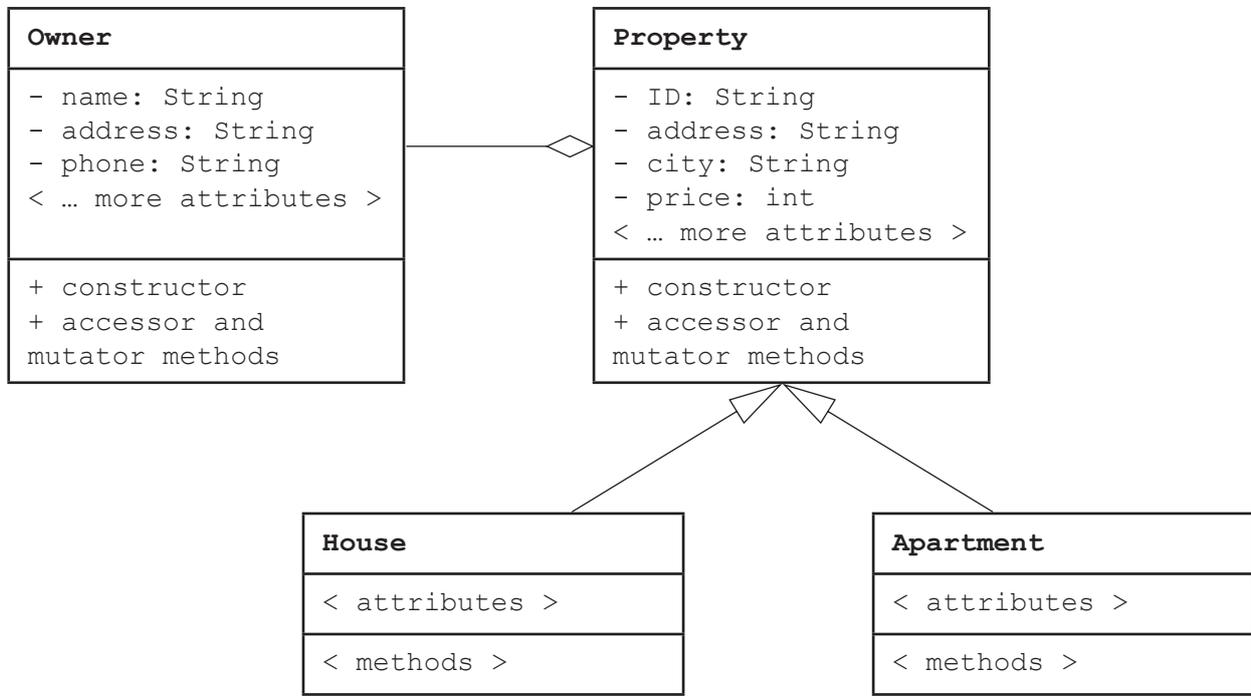
[6]

**End of Option C**

**Option D — Object-oriented programming**

13. A real estate business maintains an unsorted database of houses and apartments that it tries to sell for the property owners.

The following UML diagram describes the objects in the current system.



- (a) Define the term *primitive data type*. [1]
- (b) State an additional attribute in the class `Property` that would have data type
  - (i) Boolean [1]
  - (ii) integer. [1]
- (c) State the relationship between
  - (i) `Owner` and `Property` [1]
  - (ii) `House` and `Property`. [1]
- (d) Distinguish between a class and an instantiation in this scenario. [2]
- (e) Outline how the modifier `static` affects how a variable is used. [2]
- (f) Describe how the modifier `static` could be used to access the total number of both `House` and `Apartment` objects that have been created in this system. [2]

**(Option D continues on the following page)**

**(Option D, question 13 continued)**

The object-oriented software solution that implements this system for the real estate business allows a customer to select a maximum of 10 houses that he or she is interested in. These houses are stored in an array `wishList` of type `House`.

- (g) Construct the code needed to instantiate an array `wishList` that can store a maximum of 10 `House` objects. [3]

14. (a) Define the term *encapsulation*. [1]

- (b) Explain **one** benefit provided by *encapsulation*. [3]

The real estate business is planning to expand its database to include student housing. These “houses” are usually single rooms in a privately owned house.

- (c) Explain **one** benefit of using inheritance to create a new class `Room`. [3]

Many of the prospective students have an international background.

- (d) Explain **one** feature of modern programming languages that allows for internationalization. [3]

**(Option D continues on the following page)**

**(Option D continued)**

15. All the unsorted `House` objects in the database have been copied to a sufficiently large array `allHouses`. This array is not completely filled with `House` objects.

The array `allHouses` and all methods in this question are declared in the main program class. All methods can access the array `allHouses` directly.

Consider the following method.

```
public void unknown(String x)
{
    for (int i=0; i<allHouses.length; i++)
    {
        if (allHouses[i].getCity().equals(x))
        {
            System.out.println (allHouses[i].getAddress());
        }
    }
}
```

- (a) Define the term *method signature*. [2]
- (b) Describe how the original `String` variable, passed to a method as a parameter, can be assigned a new value by that method. [2]
- (c) State the intended purpose of the method `unknown`. [1]
- (d) (i) Outline the runtime error that is likely to occur if this method is called. [2]  
(ii) Outline how this error can be corrected. [2]
- (e) Construct the code for the method `houseSort` that will sort the array `allHouses` in ascending order of `price`. [5]

A method is needed to select from the original unsorted array `allHouses` the three most expensive houses below or equal to a given price.

- (f) Construct the code for the method `selectThree` that will take an integer parameter `budget`. It must return a sorted array of size 3 that contains the three most expensive `House` objects (in ascending order of price) with a price that is less than or equal to `budget`.

You may assume that the array `allHouses` contains at least three `House` objects with a price less than or equal to `budget`. As part of your answer you should use the method `houseSort()` as developed in part (e).

[7]

**(Option D continues on the following page)**

**(Option D continued)**

16. Once a month the real estate business organizes an open day for a particular house.

Interested clients can sign up by giving their name and phone number to the real estate business.

This information is stored in a list node `ClientNode` as follows.

```
public class ClientNode
{
    private String name;
    private String phone;
    private boolean attended;
    private ClientNode next;

    public ClientNode (String aName, String aPhone)
    {
        this.name = aName;
        this.phone = aPhone;
        this.attended = false;
        this.next = null;
    }
    public String getName()
    {
        return this.name;
    }
    public String getClient()
    {
        return this.name+" "+this.phone;
    }
    public void setAttended()
    {
        attended = true;
    }
    public ClientNode getNext()
    {
        return this.next;
    }
    public void setNext(ClientNode nextNode)
    {
        this.next = nextNode;
    }
}
```

- (a) Identify **two** features of an abstract data type (ADT).

[2]

**(Option D continues on the following page)**

**(Option D, question 16 continued)**

A new class `ClientList` is implemented as outlined below.

```

public class ClientList
{
    ClientNode root;

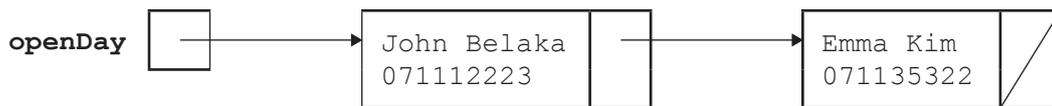
    // default constructor

    void enList(ClientNode visitor)
    {
        ... adds visitor to the end of the list - missing code ...
    }

    boolean isEmpty()
    {
        return (root==null);
    }
    ... more methods ...
}

```

A list `openDay` has been instantiated of the type `ClientList`.



(b) Describe the meaning of the following statement. [3]

```

openDay.enList(new ClientNode("Sophie Bella", "072456721"));

```

(c) By using object references, construct the method `enList` that allows a client to be added to the end of the list `openDay`. [5]

The attendance of potential buyers is tracked during the open day by changing the `attended` variable to `true` when a potential buyer arrives.

At the end of the day, the list `openDay` is processed so that every potential buyer that attended is added to new list `hasAttended` and every potential buyer that did not attend is added to a new list `notAttended`.

(d) Without writing code, outline the steps involved in splitting the list `openDay` into a list `hasAttended` and a list `notAttended`. [4]

**(Option D continues on the following page)**

**(Option D, question 16 continued)**

The real estate business keeps the data of all past owners in a file for future reference. This file is sorted by owner name. When needed, this file is read into a `LinkedList` called `contacts`, which has been instantiated as follows.

```
private LinkedList<Owner> contacts = new LinkedList<Owner>();
```

(e) Construct a recursive binary search method given as

```
public Owner binSearch(String name, int low, int high)
```

where `name` is the search term.

You may assume that `contacts` is accessible to `binSearch` and that it has been filled with many objects.

You may use the following standard `LinkedList` method.

```
.get(int index)
```

This returns the object located in the `LinkedList` at `index`.

[6]

**End of Option D**

---

**Disclaimer:**

Content used in IB assessments is taken from authentic, third-party sources. The views expressed within them belong to their individual authors and/or publishers and do not necessarily reflect the views of the IB.

**References:**

- 12. Image 1** Meusel, Robert, et al., 2014. Graph Structure in the Web - Revisited. *Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion* [e-journal]  
<http://dx.doi.org/10.1145/2567948.2576928>. Source adapted.

**All other texts, graphics and illustrations © International Baccalaureate Organization 2023**