

Computer science
Higher level
Paper 2

Monday 7 May 2018 (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 – 10
Option C — Web science	11 – 14
Option D — Object-oriented programming	15 – 19

Blank page

Option A — Databases

1. *Galaxy Bank* is a US based bank with many banks (branches) across the country. *Galaxy Bank* uses a relational database to support its operations.

Each **branch** has many **customers** and each **customer** may take out a number of **loans**.

- (a) Construct the entity-relationship diagram (ERD) that shows the relationship between the bank branch, the customers and their loans. [2]

Some of the data in the **LOANS** table is shown below.

The underlined attribute indicates the primary key.

LOANS

<u>Loan_ID</u>	Customer_ID	Amount	Type
XY080911	668848616	500 000	Home
RD050789	987887982	430 000 000	Venture
RG903540	124567552	231 900 000	Venture
XY348801	648782904	800 000	Home
ED569891	456783356	230 000	Education
...

- (b) With reference to the table **LOANS**, state the output of the following query:

```
SELECT LOANS.Loan_ID, LOANS.Customer_ID, LOANS.Amount, LOANS.Type
FROM LOANS
WHERE (LOANS.Amount > 600000) AND ((LOANS.Type = "Home") OR
(LOANS.Type = "Venture"));
```

[2]

(Option A continues on the following page)

(Option A, question 1 continued)

Two other tables in the database are identified below:

- The **ACCOUNTS** table contains the account details of the customers.
- The **CUSTOMERS** table contains the contact details of each customer.

The underlined attribute indicates the primary key in each table.

ACCOUNTS

<u>Acct_no</u>	Customer_ID	Balance	Account_Type
171345090	987887982	212 456.00	Deposit
171345087	987887982	3 450.00	Current
345904570	456783356	320 567.00	Deposit
345904571	456783356	590.00	Current
435362728	987635218	610.00	Current
356278236	533763373	4567.00	Current
...

CUSTOMERS

<u>Customer_ID</u>	Family_Name	Zip_Code	Phone
668848616	Smith	90906	3456042
987887982	Wallace	32001	3205600
124567552	James	23057	6783456
648782904	Evans	21026	4545542
456783356	Donnan	23058	3789036
453627286	Davies	24026	3678936
546373883	Walker	25098	3974673
...

- (c) Identify the steps to create a query to find the names of customers whose account balance is greater than \$300 000. [4]

Security is the top priority for *Galaxy Bank*.

- (d) Explain how controlling access rights contributes to the security of the *Galaxy Bank* database. [3]
- (e) Explain how row locking ensures the consistency of the data in the *Galaxy Bank* database. [3]

(Option A continues on the following page)

(Option A continued)

2. *ShowTime* is a local theatre that allows online booking for its various shows. All customers are able to access the database to book a movie of their choice.

However, some of the customers have received incorrect information when their booking has been made. This has been caused by update anomalies.

- (a) Identify **two** types of update anomaly. [2]

ShowTime have introduced a database recovery system in case the database becomes corrupted.

- (b) Outline **two** methods of database recovery that can be used to restore the system. [4]

- (c) Identify **three** tasks that are carried out by the database administrator (DBA). [3]

Customers can choose a credit card as their method of payment. However, some customers are concerned that their personal information could get shared with unauthorized third parties.

- (d) Explain **one** way that the DBA at *ShowTime* can ensure the anonymity of the customers is maintained. [3]

(Option A continues on the following page)

(Option A continued)

3. *Jackson City University* has a Music Department that provides music lessons to students in a number of high schools in the city.

The *Jackson City University* Music Department teachers visit the different schools in the city to teach students a range of musical instruments.

The following diagram shows an unnormalized table of student data.

Student_ID	First_Name	Family_Name	Date_of_birth	Student_Choice	School	School_Post_Code	School_phone_no
305	JOE	PARKAR	02/09/2000	VIOLIN, PERCUSSION	Tampines	512233	065544335, 065544336
306	MARY	ALCHIN	01/06/2000	GUITAR, KEYBOARD	Arab Street	198876	065543213
307	RAM	WONG	23/03/1999	GUITAR	Bedok	467501	065433221, 065433222
308	KATE	MOHAMA	17/08/1999	PERCUSSION	Simei	512233	065444412
309	RYAN	TAN	02/09/2000	GUITAR	Seng Kang	208563	065667898
310	PHILIP	BALMER	10/10/1999	VIOLIN	Orchard	309893	065756379
311	MIKE	MOSES	10/11/2000	GUITAR	Clementi	308978	065745677
...			

- (a) Explain **one** benefit of normalizing a database. [3]
- (b) Identify **three** ways that incorrect data could be prevented from being added into the *School_phone_no* field. [3]
- (c) Outline what would be necessary to make the above unnormalized table conform to 1st Normal Form (1NF). [2]
- (d) Construct the 3rd Normal Form (3NF) of the unnormalized relation shown above. [8]
- (e) Explain the difference between 2nd Normal Form (2NF) and 3rd Normal Form (3NF). [3]

(Option A continues on the following page)

(Option A continued)

4. *Basking Coats* is a business that sells textile products such as shirts, coats and trousers. The company was formed in 1970 and has numerous shops in Europe and South-East Asia. To ensure their marketing is targeted at appropriate customers, *Basking Coats* has asked *Singalytics*, a data analytics company, to assist them in improving their marketing strategy.

(a) Describe how deviation detection can be used to analyse this data. [2]

(b) Outline why data warehouses tend to use unnormalized data sets. [3]

Extract, Transform, Loading (ETL) processes can be used to clean up data for use in a database warehouse. When ETL is carried out, certain precautions should be taken.

(c) Identify **three** precautions to be taken before extraction is carried out on the database. [3]

Data in the *Singalytics* data warehouse is stored with a timestamp.

(d) Outline why data warehousing is time dependent? [2]

(e) Explain why *Basking Coats* could use association analysis to improve the marketing of its products. [4]

Basking Coats has decided to use an object-oriented database rather than a relational database to store its data.

(f) Explain why *Basking Coats* would use an object-oriented database rather than a relational database to store its data. [6]

End of Option A

Option B — Modelling and simulation

5. Ralph owns a furniture store that trades in second-hand furniture, lamps and musical instruments. He buys these items, repairs them where necessary and then resells them.

He calculates the value of each item using the information below:

- type: furniture, lamp, musical instrument
- brand recognition: 1 (low), 2 (medium), 3 (high)
- condition: “very good”, “good”, “needs repair”
- estimated volume of item: maximum volume accepted is 2000 dm^3 (note that $1 \text{ dm}^3 = 1 \text{ litre}$).

Ralph is going to use a spreadsheet to model this information.

- (a) Copy and complete the following table showing the variables, each variable’s data type and range of values that would represent the information shown above. [4]

Variable	Data type	Range of values

The spreadsheet model will inform Ralph as to whether he should buy an item. The decision to buy is based on the following rules:

- the default values for the items are; furniture \$100, lamp \$30, musical instrument \$80
- multipliers are applied to these default values depending on the volume of the item; for volumes that are 500 dm^3 or above it is 0.7, for volumes less than or equal to 30 dm^3 it is 1.2, for any other volume it is 1.

Ralph will not buy any item with a value calculated to be over \$90.

- (b) (i) Using the above rules, construct the pseudocode that will help Ralph in deciding whether to buy an item. [4]

- (ii) Describe **two** items that would have a calculated value of more than \$90. [2]

(Option B continues on the following page)

(Option B, question 5 continued)

The model is also used to calculate his final selling price for each item. This price is calculated using the following rules:

- initial selling value is 5 times the original default value
- a multiplier of 1, 2 or 3 is applied to reflect the brand recognition (a “top brand” receives a multiplier of 3)
- another multiplier is applied to reflect the volume occupied. This multiplier is 5 when the volume is greater or equal to 500 dm^3 , 3 for volumes between 30 and 500 dm^3 and 1 for other volumes
- a further multiplier is applied depending upon the item’s condition. This multiplier is 2 or 3 when conditions are good or very good, but is 0.8 if the item is damaged.

- (c) Calculate the selling price of a top brand guitar with a volume of 96 dm^3 that was damaged. You should show your working. [2]
- (d) With the help of a diagram, suggest an appropriate design for a spreadsheet used to calculate the final selling price, following this model. [4]
- (e) Identify **two** tests that should be included in the test plan for this model. [2]

6. In a small shop two people prepare and sell bread rolls. The space for preparing the bread rolls is small. This space includes a sink and a cash register. Each of the ingredients used to fill the bread rolls is kept in a separate bowl with its own fork or spoon. Sometimes the bowls need to be refilled.

Preparing a bread roll requires the following steps:

- slice the bread roll
- spread its base with some sauce
- add lettuce
- fill with at most two chosen fillings
- cover the roll with the other half
- wrap it in paper.

After that, the roll is ready for payment and collection.

Even when the shop is busy, each person only prepares one bread roll at a time.

- (a) Outline **two** problems with this method of preparation that could affect the time it takes the two people to prepare an order. [4]
- (b) Describe how the two people could improve the efficiency of their work, without compromising on the quality of service to the customers. [3]

(Option B continues on the following page)

Turn over

(Option B, question 6 continued)

A larger restaurant is already using simulation software for the preparation of their dishes.

The restaurant has a menu with eight possible dishes. These dishes may require a number of steps such as cutting, mixing, boiling and grilling to prepare them, as well as different cooking times. The restaurant has a cook and an assistant who prepare **one** dish at the time.

- (c) Explain the difference between a model and a simulation. [3]
- (d) Identify **three** elements that the simulation software might consider, in addition to the information already described above. [3]

The simulation program will group all orders received in an interval of 10 minutes. The program will then produce a sequence of all the cooking steps so that these orders can be completed as quickly as possible. Once the kitchen has completed the orders for one interval, it is ready to accept orders for the next interval.

Customers are impressed by the rapidity of service, but not by the quality of prepared food.

- (e) Suggest **two** elements that the software simulation may have not considered that may lead to complaints from the customers. [6]
7. The electronic control unit (ECU) of a vehicle is an on-board computer that is constantly monitoring the performance of several components of a vehicle. For example, the ECU:
- controls the functioning of the lights, the brakes, the airbag, and the fuel-level signal
 - permits the scale being changed in some digital displays, such as switching the speedometer from miles/hour to km/hour.

The software embedded in an ECU receives input data from a variety of sources when it runs auto-diagnostic tests.

- (a) With relation to the activities that the software of an ECU has to perform:
 - (i) identify **two** of the sources that provide input data to the ECU; [2]
 - (ii) suggest **one** reason why the auto-diagnostic program in the ECU depends upon the make and model of the vehicle. [2]

Vehicles are tested for their exhaust gas emissions using simulation software at specialist garages. During a period of 5 minutes, a vehicle with the engine switched on is monitored for emissions of carbon dioxide (CO₂) and fine particulates. The software uses 3D visualization techniques to display these parameters on a screen for the whole duration of the test.

- (b) Explain how emissions of CO₂ and fine particulates could be represented in 3D by the software. [4]

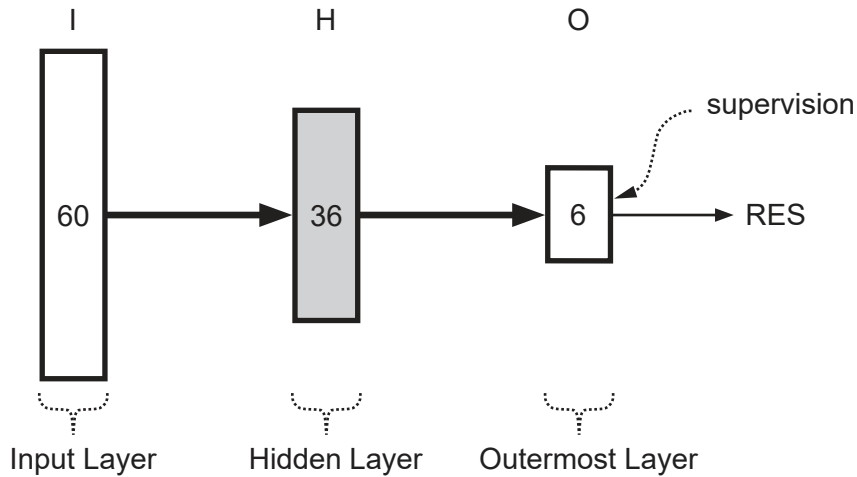
(Option B continues on the following page)

(Option B continued)

8. Neural networks are often used to model the complex behaviour of systems. These networks must first be trained in order to establish a reliable model. One such method of training makes use of supervised learning algorithms.

(a) Outline the basic steps involved in supervised learning algorithms. [4]

The following block diagram illustrates a neural network used by a supervised learning algorithm to calculate a resulting value RES, for a given set of inputs.



The neural network in the block diagram contains 60 nodes in the input layer (block I), 36 nodes in the hidden layer (block H) and 6 nodes in the outermost layer (block O). The solid black arrows indicate that *all* nodes in one layer are feeding *all* nodes in the successive layer. In particular, the value RES is calculated by using all nodes in O.

(b) Calculate the number of weights that the above neural network uses in producing a value for RES. Show the details of your calculation. [2]

(Option B continues on the following page)

(Option B continued)

9. An image processing system analyses colour images of architectural landscapes. The system compares any new images with existing ones stored in a knowledge base. The purpose of the system is to extend the knowledge base with images of the same scenes taken either from different perspectives or at different times of the year or with different weather conditions.

(a) Explain how cluster analysis can be used to achieve the aims of the system that is described above. [4]

(b) Discuss the appropriateness of using genetic algorithms to compare the processed images with those stored in the knowledge base. [6]

Machine translators are regularly used to translate text from one language to another. In spite of the advances that have been made in this field, the output may still need to be proofread by a human.

10. Describe **two** problems that these translators may encounter when translating from one language to another. [4]

End of Option B

Option C — Web science

11. Home banking allows individuals to perform operations over the Internet on their own bank accounts.

- (a) Outline the relationship between the Internet and the world wide web (WWW). [2]
- (b) Explain why the choice of browser should not affect a customer’s ability to access their bank account details. [4]
- (c) State **two** features that make HTTPS more suitable than HTTP in the context of home banking. [2]

Access to a bank’s home banking services requires, as a first step, identification and authentication of the user. Individuals log on the bank web site, and enter their own personal space by providing their full account number and a personal code that the bank gave them. The processing of this information takes place on the server side.

- (d) Explain why server-side processing is used in this case. [3]

The 20 most recent account transactions can be displayed on screen in a webpage that uses XML. A print-out of **all** transactions of the past three months may be obtained by clicking an onscreen button on the webpage. The print-out is landscape oriented and shows many more columns than are displayed on the screen.

- (e) Describe how this processing takes place with reference to the use of XML and XSLT. [3]

(Option C continues on the following page)

Turn over

(Option C continued)

12. *RunAndBeyond* is a sports knowledge website, offering expert knowledge and technical tips on sport disciplines. Official teams and associations are invited to include some of their own multimedia digital resources to *RunAndBeyond*.

A team of triathletes, finding their sport discipline under-represented in the website, wants to contribute content for *RunAndBeyond*. The team intends to use a wiki in collaborating online with other similar interested groups in triathlon when creating and updating their own resource.

- (a) Describe how a wiki can support the ongoing collaboration in producing the triathletes' resource. [3]

The triathlete's online resource provides a variety of references through external links, and one of them points to:

`ftp://files.tri-events.cc/site/index.php/en/general-info`

- (b) (i) State the reason why the above link is a URL. [1]
(ii) Outline the processing that takes place when the line of code above is executed. [3]

There are also external links on the website to videos and some of these are transferred using lossy compression. When these videos are transferred, they can either be streamed or downloaded for future viewing.

- (c) Suggest why lossy compression should be the compression technique used. [4]

The *RunAndBeyond* website continually automatically updates its content regarding a live sporting event.

- (d) Suggest how a dynamic web page would function in providing this service to the user. [4]

(Option C continues on the following page)

(Option C continued)

13. A web application (app) runs on mobile devices such as smartphones and tablets. It allows users to locate their position in real time on a map, as they walk around a city, as well as the surrounding attractions. The app uses icons to represent tourist attractions such as art galleries and museums. When the user clicks on the icon, further details are shown, such as opening times. The app includes some use of client-side scripting.

- (a) (i) Outline the functioning of this app. Include specific references to the technology and software involved. [3]
- (ii) With reference to the use on mobile devices, outline a feature of this application that may rely on client-side scripting. [2]

Many art galleries have websites that can be found by search engines. White hat techniques and practices allow website developers to optimize the search process. It is good practice to maintain the source code of websites up-to-date with actual information.

- (b) (i) State **two** metrics used by search engines. [2]
- (ii) Explain why maintaining a clean HTML source code of a website by removing old information optimizes the search process. [5]

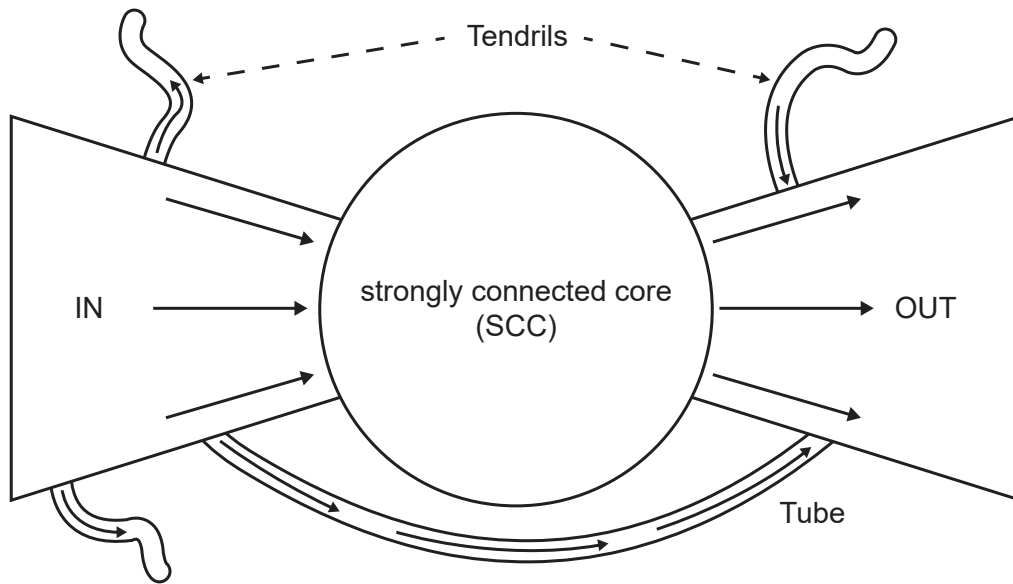
The evolution of the web, architectures, protocols and their uses has led to increasingly sophisticated services that run on peer-2-peer (P2P) architectures.

- (c) Explain how a P2P network can provide more reliability than a client-server model. [4]

(Option C continues on the following page)

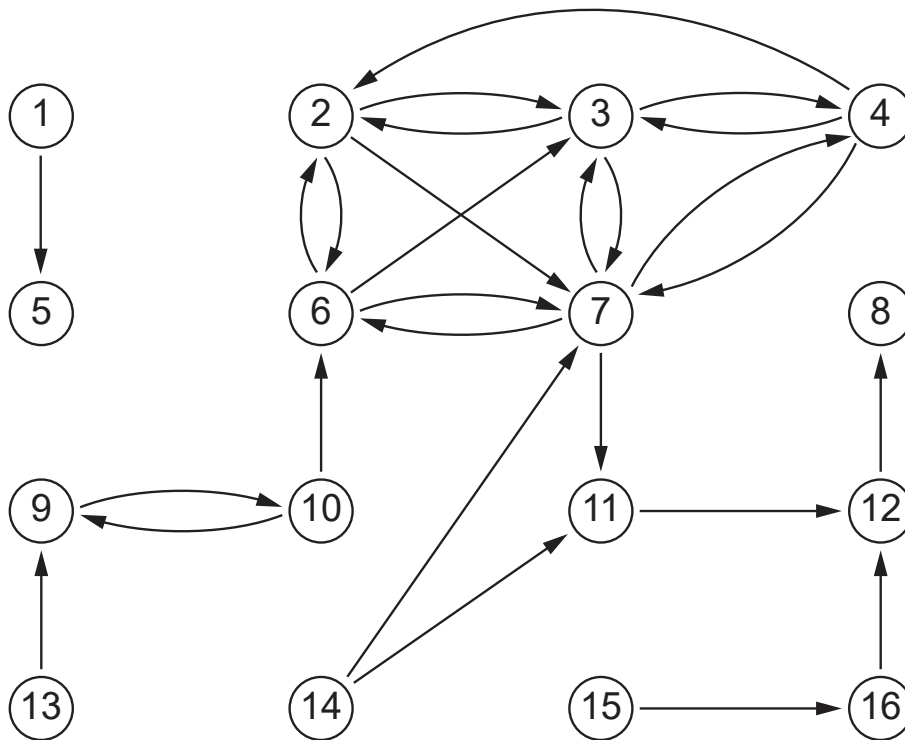
(Option C continued)

14. In the bow-tie model of the web, illustrated below, we can quickly identify some components of a very large network, represented as a graph.



[Source: Reprinted from Computer Networks, 33/1-6, Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, Janet Wiener, *Graph structure in the Web*, 309-320, Copyright (2000), with permission from Elsevier. <https://www.sciencedirect.com/science/article/pii/S1389128600000839>]

Consider the following diagram:



(Option C continues on the following page)

(Option C, question 14 continued)

- (a) Outline why this diagram can be considered a sub-graph of the World Wide Web. [2]
- (b) With reference to the diagram provided and the terminology of the bow-tie model:
 - (i) identify the set of IN nodes; [2]
 - (ii) identify the OUT nodes; [1]
 - (iii) outline why node 7 is part of the strongly connected core; [1]
 - (iv) outline why node 15 is not an IN node. [1]

A study has been commissioned to research the role of soft drinks in childhood obesity. As part of this investigation, web crawlers are used to search the web.

- (c) Explain how a web crawler can access information on the web. [3]

Some of the audio-visual documentaries on the topic, that could provide valuable documentation for this study, have not been retrieved by these web crawlers.

- (d) Explain **two** reasons why web crawlers could not retrieve these documentaries. [4]

The development of the semantic web would offer opportunities to incorporate collective intelligence in returning more useful search results for the understanding of the relationship between soft drinks and childhood obesity.

- (e) Discuss whether searches based on the use of the semantic web that are able to incorporate collective intelligence might provide more relevant results than purely text based searches. [6]

End of Option C

Option D — Object-oriented programming

A hotel chain has a loyalty scheme in which customers are awarded 1000 points for each day they stay in one of their hotels. With these points, customers can achieve one of three status levels: Gold, Silver or Bronze. The level will determine the extra services to which they are entitled.

The total number of points collected during the **current** year will determine which of the three status levels they are assigned for the **following** year: For example **only** the points collected in 2018 will determine the status level for 2019.

Occasionally, new customers receive additional **bonus** points as part of a promotion.

The `Points` class keeps details of the points and status levels of each customer.

```
public class Points
{
    private String memberId; // id of the hotel customer
    private int totalPoints; // this year's points
    private int bonusPoints; // any bonus points given to this year's new member
    private String statusNow; // current(this year's)status
    private String statusNextYear; // following year's status
    private Visits[] allVisits = new Visits[366]; // details of each visit
                                                // during this year
    int y; // number of visits this year

    public Points(String id) // constructor for new member
    {
        memberId = id;
        bonusPoints = 0;
        y = 0;
        statusNow = "Bronze";
    }

    // constructor for new member given bonus points (valid for current year only)
    public Points(String id, int bp)
    {
        memberId = id;
        bonusPoints = bp; // multiples of 1000 - maximum number is 5000
        y = 0;
        statusNow = "Bronze";
    }

    // all the accessor and mutator methods are present but not shown

    public Visits getAllVisits(int v)
    {
        return allVisits[v];
    }

    public void addVisit(Visits v) // adds a new Visit object to the array
    {
        allVisits[y] = v;
        y = y + 1;
    }

    isGold() {code missing}
    calculateTotalPoints() {code missing}
    daysMissing() {code missing}
}
```

(Option D continues on the following page)

(Option D continued)

15. The instance variables in the `Points` class are preceded by the modifier `private`. The choice of modifier affects the way in which these variables are accessed or used.
- (a) With the use of **two** examples other than `private`, outline how the choice of this modifier affects the way in which these variables are accessed or used. [4]
 - (b) With reference to the two methods with the same name in the `Points` class, explain the OOP feature that makes it possible to successfully implement either of these methods. [4]

The customers will be assigned one of three levels for the following year (Gold, Silver or Bronze) depending upon the current year's total points as follows.

- Bronze = less than 10 000 points
- Silver = 10 000 or more but less than 50 000
- Gold = 50 000 or more.

In 2018, Tim became a member for the first time and was awarded a bonus of 1000 points. So far, in 2018, Tim has stayed three times at one of these hotels. The first visit lasted 2 days, the second visit lasted 1 day and the third visit lasted 6 days.

- (c) State the status level that Tim has been assigned, for 2019, following these visits. [1]

The different `Points` objects are stored in an array which is declared globally in the main (driver) class as follows: `Points[] allPoints = new Points[10000];`

- (d) State how an individual object can be identified using this array. [1]

The attribute `statusNow` is assigned its correct value at the beginning of every year for existing members. It cannot be changed during the year.

- (e) Construct the method `isGold()` in the `Points` class, which will return whether the current status is "Gold". [3]

(Option D continues on the following page)

(Option D continued)

16. The details of hotel stays during the current year are stored in the variable `allVisits` which is an array of the `Visits` class. `allVisits` is used in determining the total points awarded in the current year.

The `Visits` class is outlined below:

```
public class Visits
{
    private String hotelCode; // id of the hotel
    private int days; // number of days of the visit

    public Visits(String h, int d)
    {
        hotelCode = h;
        days = d;
    }

    public int getDays()
    {
        return days;
    }
}
```

- (a) Construct a UML diagram for the `Visits` class.

[3]

The main (driver) class manages the `Points` and `Visits` classes. It contains the following code:

```
Points[] allPoints = new Points[10000]; // declared globally

allPoints[0] = new Points("m100");
allPoints[1] = new Points("m101", 5000);
allPoints[2] = new Points("m102", 2000);

Visits v1 = new Visits("h003", 3);
Visits v2 = new Visits("h013", 1);
Visits v3 = new Visits("h013", 2);
Visits v4 = new Visits("h005", 6);

allPoints[0].addVisit(v1);
allPoints[0].addVisit(v2);
allPoints[0].addVisit(v3);
allPoints[0].addVisit(v4);
allPoints[1].addVisit(v1);
allPoints[1].addVisit(new Visits("h004", 6));
```

(Option D continues on the following page)

(Option D, question 16 continued)

- (b) State the output given by the following statements:
 - (i) `System.out.println(allPoints[2].getMemberId());` [1]
 - (ii) `System.out.println(allPoints[0].getBonusPoints());` [1]
 - (iii) `System.out.println(allPoints[1].getAllVisits(1).getDays());` [1]
- (c) Construct the method `calculateTotalPoints()`, in the `Points` class, which will calculate and return the total number of points awarded so far in the current year. [5]
- (d) Construct the method `daysMissing()`, in the `Points` class, that will return the number of extra days that a customer needs to stay in order to keep the same status the following year as they have in the current year. [7]

The hotel chain maintains the details of the extra benefits that each status provides in appropriate classes.

- (e) Suggest how the hotel chain might make use of the inheritance feature of OOP when designing the classes from pages 18 and 20. [4]

At midnight on 31st December each year the systems are temporarily made unavailable as the hotel chain makes the changes needed to prepare the system for the new year.

- (f) Identify **three** changes that would have to be made to the classes previously described in order for the system to function correctly in the new year. [3]

17. Full details of each hotel are stored as objects of the `Hotel` class.

For research purposes, the managers want to identify the name of the hotel in which a particular customer has stayed the most days during the current year.

Without writing code, outline the steps that would have to be taken to accomplish this. You should include reference to any classes, methods or attributes that will be created or will be needed. [7]

(Option D continues on the following page)

(Option D continued)

18. Full details of the customers are stored as objects of the Customer class. This class is partially shown below:

```
public class Customer
{
    private String memberId;
    private String email; //email address (assume only 1 per customer)

    public Customer(String a, String b)
    {
        memberId = a;
        email = b;
    }

    public String getMemberId()
    {
        return memberId;
    }

    public String getEmail()
    {
        return email;
    }
}
```

The objects can be accessed through the linked list `allCustomers` which is declared in the main (driver) class as follows:

```
LinkedList<Customer> allCustomers = new LinkedList<Customer>();
```

- (a) Outline why a linked list structure has been chosen for `allCustomers`. [2]
- (b) Construct the method `goldMails()` that will return an `ArrayList` containing the email addresses of all current “Gold” members. You should make use of any previously defined methods. [6]

(Option D continues on the following page)

(Option D continued)

- 19. (a) Identify **two** essential features of a recursive algorithm. [2]
- (b) Consider the following recursive method.

```
public void recursionEx(int x)
{
    if(x != 0)
    {
        System.out.println(x);
        recursionEx(x - 1);
        System.out.println(x);
    }
}
```

Copy and complete the following table to show the output when the method is called by: `recursionEx(3);`

parameter passed to method	output
3	

[4]

- (c) Explain what would happen if the method was called by `recursionEx(-3)`. [3]
- (d) Identify **three** features which should be included in either program code or UML diagrams to help programmers understand and modify programs in the future. [3]

End of Option D