Diploma Programme
Programme du diplôme
Programa del Diploma

Computer science
Higher level
Paper 2

29 October 2024

Zone A morning | **Zone B** morning | **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 – 13
Option D — Object-oriented programming	14 – 18

8824 – 6905

26 pages

Option A — Databases

1. After a crime has been committed, police officers carry out a number of interviews.

In each interview, a police officer will interview a witness (a person who saw the crime being committed).

- A witness may be interviewed many times.
- A police officer may carry out many interviews.

(a) Construct an entity-relationship diagram (ERD) that shows the relationships between the interview, the police officer and the witness. [2]

When creating a relational database, such as a police database, it is important to consider referential integrity.

(b) Identify **three** reasons why it is important to enforce referential integrity. [3]

(c) Explain how a query can provide a view of the database. [3]

2. (a) Define the term *secondary key*. [1]

In the police database, a WITNESS table, CRIME table and OFFICER table have been created, but are not shown. The primary keys for these tables are WitnessID, CrimeID and OfficerID respectively.

When a police officer interviews a witness, they ask for a statement about the crime.

Sometimes, a witness may need to be interviewed more than once by the same police officer.

A record of each interview is placed in the INTERVIEW table:

INTERVIEW(<u>WitnessID</u>, <u>CrimeID</u>, <u>OfficerID</u>, Statement, CrimeLocation, DateTimeStatement, PoliceStation, StationPhone)

The foreign keys WitnessID, CrimeID and OfficerID in the INTERVIEW table form a composite primary key for this table.

(b) (i) Outline why the use of WitnessID, CrimeID and OfficerID is not a suitable composite primary key. [2]

(ii) Outline why the INTERVIEW table is not in 3rd Normal Form (3NF). [4]

(Option A continues on the following page)

(Option A, question 2 continued)

Figure 1 includes data from two crimes.

Figure 1

Fields	Crime 1	Crime 2
CrimeID	56326	56543
CrimeType	Theft	Computer misuse
OfficerID and OfficerName	**OfficerID** **OfficerName** 0071 Delphine Allais 0056 Eric Demoncheaux	**OfficerID** **OfficerName** 0099 Christelle Chades 0082 Arthur Katalayi
CrimeAddress	18 Rue St Michel, 75011 Paris	127 Rue St Jacque, 75011 Paris
CrimeDateTime	2022-06-21 03:33:00	2022-06-21 02:27:00
EvidenceType	Statement Statement Photograph	Statement Statement Video
EvidenceDesc	Scott Ngatai statement Zixin Yang statement Photograph of broken window	Ciara Alinec statement Zixin Yang statement CCTV video 02:20 to 03:00
VictimID (VID) VictimName, and VictimPhone	**VID** **VictimName** **VictimPhone** HK20 Hana Kim 01029238484 ZY33 Zixin Yang 01029968565 SN21 Scott Ngatai 01029451869	**VID** **VictimName** **VictimPhone** CA19 Ciara Alinec 01021103431 ZY33 Zixin Yang 01029968565 HK20 Hana Kim 01029238484

The following database notation represents the fields in **Figure 1**.

CRIME(<u>CrimeID</u>, CrimeType, OfficerID, OfficerName, CrimeAddress, CrimeDateTime, EvidenceType, EvidenceDesc, VictimID, VictimName, VictimPhone)

The data type for VictimPhone is alphanumeric.

(c) Outline why an alphanumeric data type is used for the VictimPhone field. [2]

(d) Construct the database in 3rd Normal Form (3NF) for all of the data shown in **Figure 1**. You should use database notation as shown in the CRIME table. [6]

(Option A continues on the following page)

(Option A continued)

3. Two police officers attempt to update a witness statement at the same time.

(a) Describe how data locking can be used to deal with this situation. [2]

Many countries have a centralized crime database.

(b) Discuss the advantages and disadvantages of using a centralized crime database. [5]

Data mining and data matching algorithms are used to analyse crime data.

(c) (i) Explain why data mining might be used to extract information from crime data held in a police database? [4]

(ii) Explain, using an example, how data matching could be applied to the crime data. [4]

Legislation has been created to protect the rights of individuals whose data is stored electronically. Most countries have similar regulations to the General Data Protection Regulation (GDPR) found in Europe.

One requirement of legislation is that data must be secure.

(d) (i) Outline **two** reasons why a password-only approach would not be secure. [4]

(ii) Identify **three** ways that legislation, such as GDPR, protects individuals whose data is stored electronically. [3]

(Option A continues on the following page)

(Option A continued)

4. The police are considering changing from a relational database to either a network database or an object-oriented database.

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

(b) Discuss the advantages **and** disadvantages of an object-oriented database compared to a relational database. [5]

The European Union Agency for Law Enforcement Cooperation (Europol) may develop a data warehouse to hold data related to crimes.

(c) Outline **two** advantages of creating a data warehouse to store crime data. [4]

Predictive modelling is a technique that could be applied to the data warehouse.

(d) Outline **two** benefits of applying predictive modelling to crime data. [4]

Europol would like to apply link analysis to investigate fraud, uncover terrorist networks and investigate criminal activities.

(e) Explain how link analysis can be applied to crime datasets. [5]

End of Option A

Option B — Modelling and simulation

5. The probability of two people having the same birthday is 1/365 or 0.003.

The probability of two people (one pair) from a random sample of N people sharing a birthday is calculated as follows:

- Number of pairs (NP): (N*(N–1))/2
- Probability of a pair with different birthdays (PP): 364/365 = 0.997
- Probability of different pairs (DP) for a specified number of pairs (NP): PP^NP
- Probability of a pair sharing a birthday (SB): 1–DP

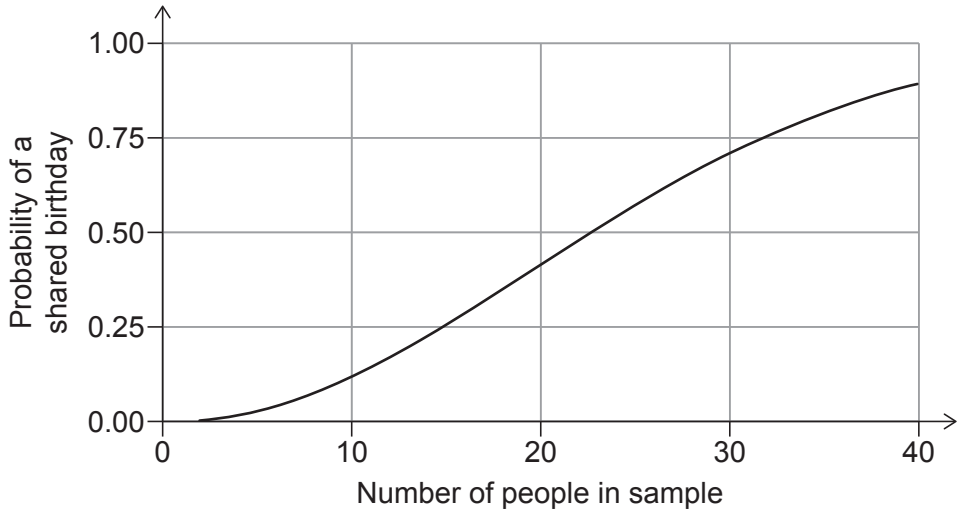
Table 1 shows how the probability of two people sharing a birthday is calculated in a sample of 23 people.

Table 1: Calculating probabilities of pairs sharing a birthday

	Formula	Result
Number of pairs (NP)	(N*(N–1))/2 = (23*22) / 2	253
Different pairs (DP)	PP^NP = 0.997^253	0.500
Pair sharing a birthday (SB)	1–DP	0.500

Figure 2 shows the probability of two people sharing the same birthday, where N is the number of people in the sample.

Figure 2: Probability of two people sharing the same birthday



(Option B continues on the following page)

(Option B, question 5 continued)

The graph in **Figure 2** was created from the spreadsheet shown below with the values shown in the first two rows:

N	NP	DP	SB
2	1	0.997	0.003
3	3	0.995	0.005

(a) Describe the steps to create a spreadsheet model that will produce the graph in **Figure 2**.

You will need to use the formulae given in the description. [4]

A spreadsheet model can be created to simulate whether two people share the same birthday by using random numbers.

The `=RANDBETWEEN(1,365)` function should be used to generate a random day between 1 and 365.

A cell should output `True` if any two birthdays are duplicated or `False` if there are no duplicate birthdays.

(b) Construct a spreadsheet model that generates random birthdays for 15 people and outputs whether there is at least one shared birthday. [4]

To test the hypothesis that the probability of a shared birthday for 23 people is 0.500, a simulation needs to be created.

This simulation will run 10 000 times and output a percentage of random people where there is at least one shared birthday.

A one-dimensional array named `DAY` has been created to store 23 random integers.

A function, `RANDINT(X,Y)`, is created to generate and return a random number between X and Y. This function will be used to populate the `DAY` array with random values between 1 and 365.

(c) Construct an algorithm in pseudocode that runs this simulation. [8]

(d) Outline why it was decided to run the simulation 10 000 times. [2]

(Option B continues on page 9)

Blank page

(Option B continued)

6. The global human population has grown from 1 billion in 1800 to 7.9 billion in 2020.

The French Institute for Demographic Studies website provides simulation software that allows users to simulate population changes.

(a) State **two** variables that might be needed for a simulation to calculate population change. [2]

(b) Explain why population change is a suitable topic to be modelled as a simulation. [5]

A population growth simulation is developed under the assumption that past events are effective predictors of future events. Testing is an important step in improving the accuracy of a simulation.

(c) Describe how you might test the accuracy of a population growth model. [5]

Unforeseen events can affect a model's accuracy. For example, the 1918–1920 H1N1 influenza outbreak is estimated to have caused 50 million deaths globally and reduced population growth.

(d) Outline **one** alternative strategy to testing to ensure the model predicts population growth accurately. [2]

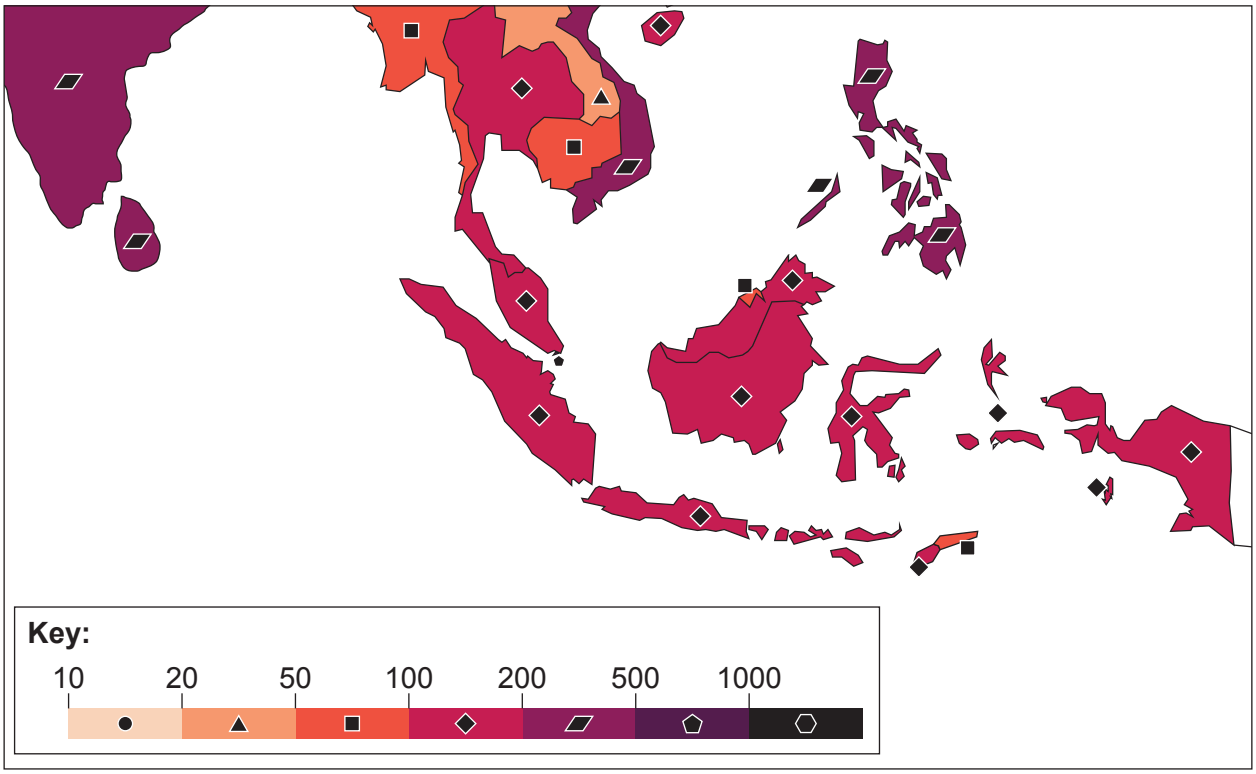
(e) Explain **two** social consequences of inaccurately simulating population growth. [6]

(Option B continues on the following page)

(Option B continued)

7. Population density is the number of people in a square kilometre. **Figure 3** shows a 2D visualization of population density in South East Asia.

Figure 3: 2D visualization of population density in South East Asia



(a) Define the term *visualization*. [1]

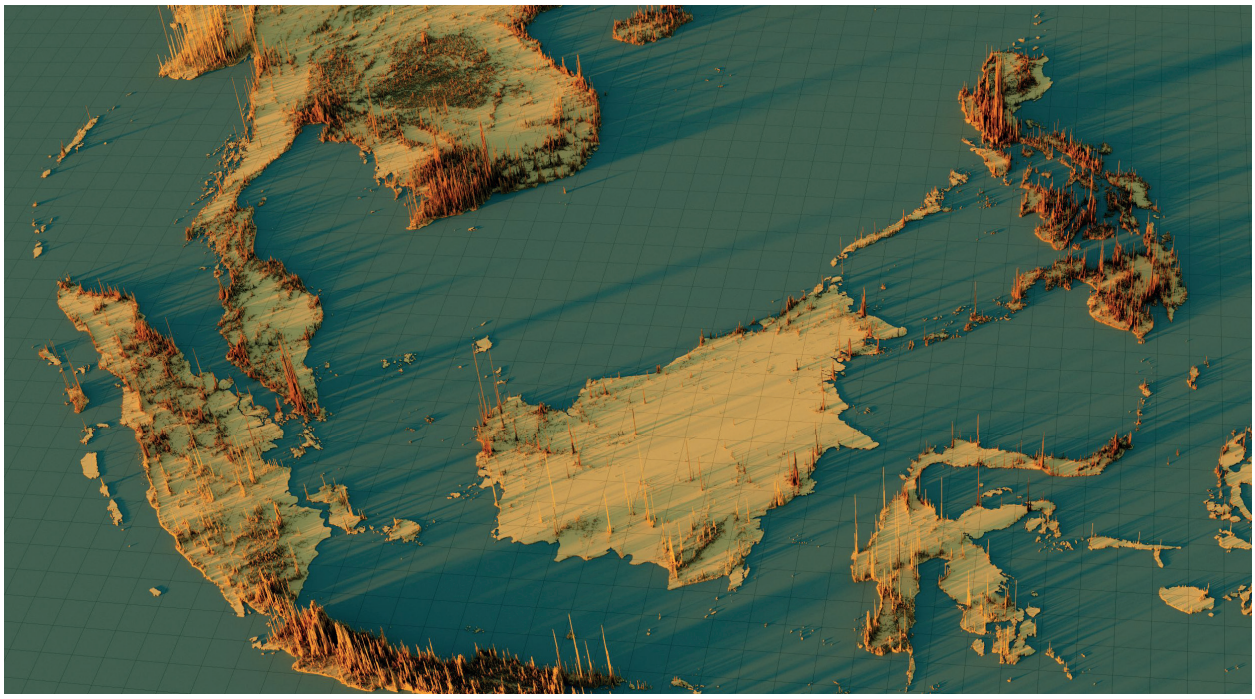
(b) Outline the memory needs of 2D visualization. [2]

(Option B continues on the following page)

(Option B, question 7 continued)

Figure 4 shows a 3D visualization of population density in South East Asia.

Figure 4: 3D visualization of population density in South East Asia



Creating similar 3D-rendered visualizations showing changes in population density over time will have considerable hardware implications.

(c) Describe the processing and memory requirements of 3D animation. [4]

(Option B continues on page 13)

Blank page

(Option B continued)

8. A team of scientists plan to travel to 54 capital cities in Africa to gather population data. They plan to use a genetic algorithm to determine the shortest route between the 54 cities.

A random set of solutions is placed into a mating pool.

(a) (i) State what data is stored in each solution. [1]

(ii) Outline why the set of solutions are random. [2]

The mating pool is a component of a genetic algorithm.

(b) Identify **two** other components of a genetic algorithm. [2]

The scientists come across people who do not know their age. They decide to use a supervised machine learning model to estimate the age of these people from photographs.

(c) Describe the steps to create a supervised machine learning model to estimate ages from photographs. [5]

The scientists plan to develop a natural language processing (NLP) application as a translation tool.

(d) Explain the differences between the cognitive and heuristic approaches to machine language learning. [4]

An unsupervised learning approach is being considered to develop an NLP application.

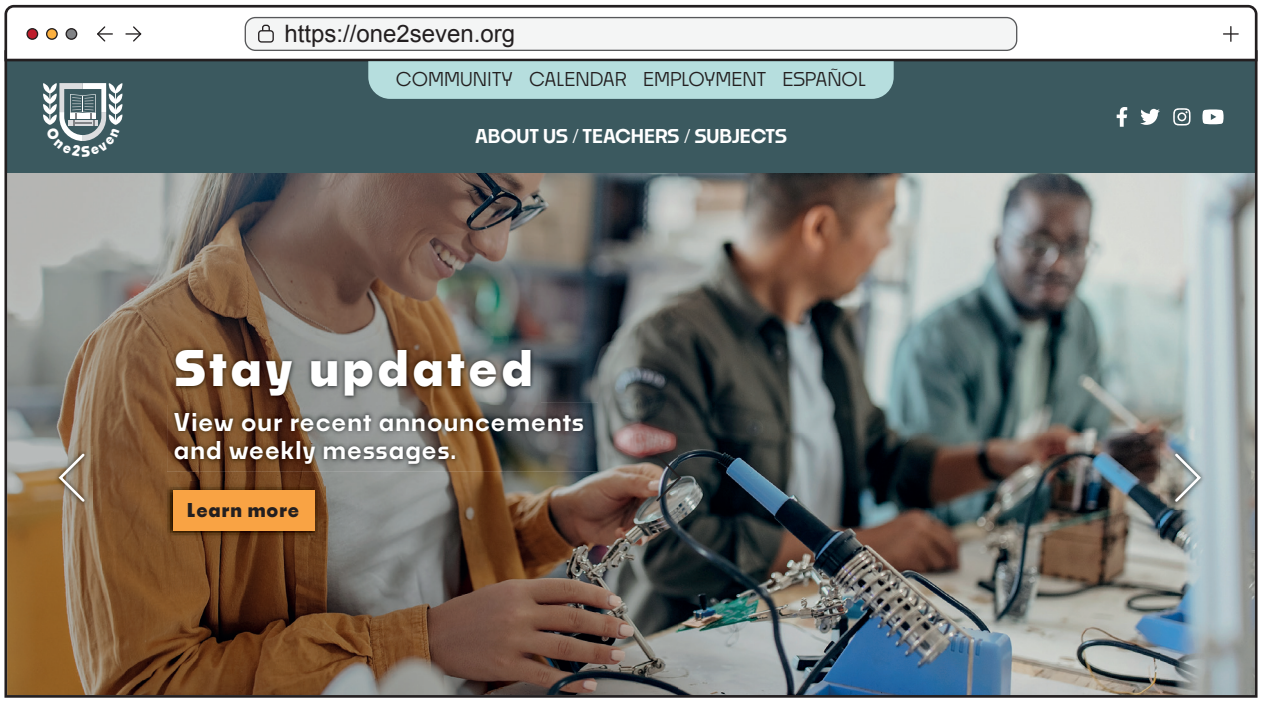
(e) Evaluate the suitability of unsupervised learning in developing an NLP application. [6]

End of Option B

Option C — Web science

9. Schools use websites to provide information to students, teachers and parents. **Figure 5** shows the home page of a school's website that was highly rated for its design.

Figure 5: Webpage from the *One2Seven* school website



(a) State the domain name in the uniform resource locator (URL) in **Figure 5**. [1]

The school's website consists of several components.

(b) Identify **two** components of the school's website. [2]

The school's website is viewed using a browser.

(c) Identify **two** functions of a web browser. [2]

The source code below shows the website uses a cascading style sheet (CSS).

```
<link rel="stylesheet" media="all" href="/assets/application-6d2c1.css"/>
```

(d) Explain **one** advantage of developing a website using cascading style sheets. [3]

(Option C continues on the following page)

(Option C, question 9 continued)

Figure 6 shows the script embedded in the homepage of the school website.

Figure 6: The embedded script on the *One2Seven* homepage

```
<script language="javascript">
if(!window.location.search.substring(1) == "check=false") {
   if(navigator.userAgent.includes('iPhone') ||
      navigator.userAgent.includes('iPad') ||
      navigator.userAgent.includes('Android') ||
      navigator.userAgent.includes('Blackberry') ||
      navigator.userAgent.includes('WebOs')) {
         window.location.replace("https://one2seven.org/mobile.html");
      }
}
</script>
```

(e) (i) Identify the type of scripting used in the JavaScript example in **Figure 6**. [1]

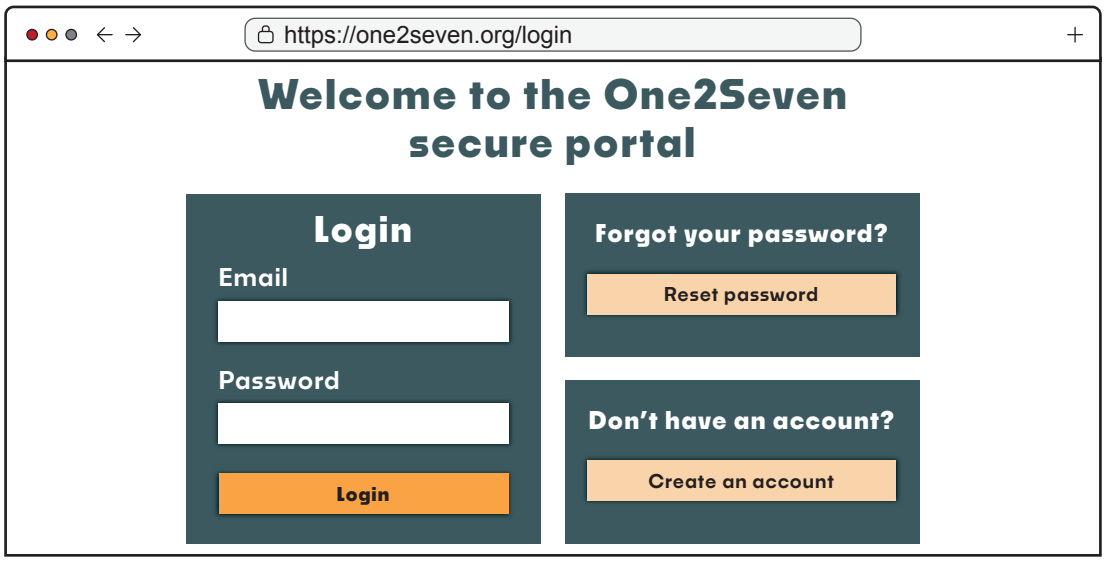
(ii) Describe the processing that occurs when this webpage is opened in a browser. [3]

(iii) Identify **one** reason why the user agent is used on this webpage instead of on the server. [1]

Some school websites contain sensitive information that is not shared with people outside of the school community.

Figure 7 shows the login page for the secure section of the school's website.

Figure 7: The login page for the secure section of the *One2Seven* school website



The school's website uses both client-side scripting and server-side scripting.

(f) Explain why the school would use client-side scripting **and** server-side scripting. [3]

(Option C continues on the following page)

(Option C continued)

10. Companies such as Amazon and Netflix use distributed cloud-based systems.

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

(b) Explain **one** reason companies such as Amazon and Netflix use distributed cloud-based networks rather than client–server networks. [3]

Amazon and Netflix allow users to download or stream videos.

(c) Outline why lossy compression may lead to a reduction in the quality of the videos that are downloaded or streamed. [2]

(d) Explain how the growth of web-based streaming services such as Netflix may lead to intellectual property issues. [3]

Videos downloaded from companies such as Amazon and Netflix can be viewed from any location in the world.

(e) Explain how the evolution of the web can lead to the removal of international boundaries. [3]

11. *XYZ Pharma* is a start-up company that is looking for ways to increase its rank in search engines such as Google, Yahoo or Bing.

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

(b) Identify **two** metrics used by a search engine. [2]

Google uses the PageRank algorithm to order the web pages it searches.

(c) (i) Identify **two** features of the PageRank algorithm. [2]

(ii) Outline why the effectiveness of the PageRank algorithm may be linked to the assumptions made during its development. [2]

(d) Explain the importance of HTTP(S) when a browser fetches a webpage from the server. [3]

XYZ Pharma is considering using a combination of white hat and black hat search engine optimization techniques to ensure its pages are ranked as highly as possible.

(e) Explain why *XYZ Pharma* may use a combination of white hat and black hat search engine optimization techniques. [3]

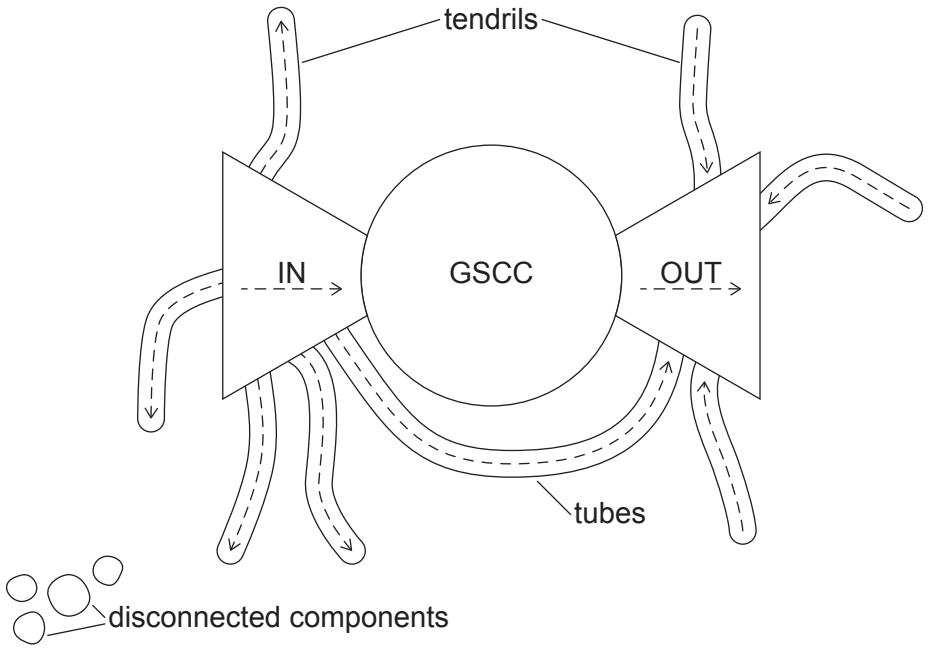
(f) Explain how the growth of search engines such as Google can create an unregulated monopoly. [3]

(Option C continues on the following page)

(Option C continued)

12. The World Wide Web (web) can be represented by models such as the bowtie structure in **Figure 8**.

Figure 8: The bowtie structure



(a) Outline **one** reason why models such as the bowtie structure are used to represent the web. [2]

(b) Explain why power laws are appropriate to predict the future development of the web. [3]

(Option C continues on the following page)

(Option C continued)

13. The web is growing exponentially and there are currently over two billion websites. As a result, the way users interact with the web has changed and they now need to use precise searching techniques to locate relevant information.

(a) Identify **two** ways a search could be made more precise. [2]

Flickr was created in 2004 when there were only 50 million websites. In the same year, the term *folksonomy* was used for the first time by Thomas Vander Wal.

Other websites, such as Raindrop, are examples of folksonomies.

(b) Identify **two** characteristics of a folksonomy. [2]

(c) Describe **one** way that sites such as Flickr and Raindrop are changing the web. [2]

Google Lens is an image recognition technology which allows the users to search for objects captured live from a camera phone.

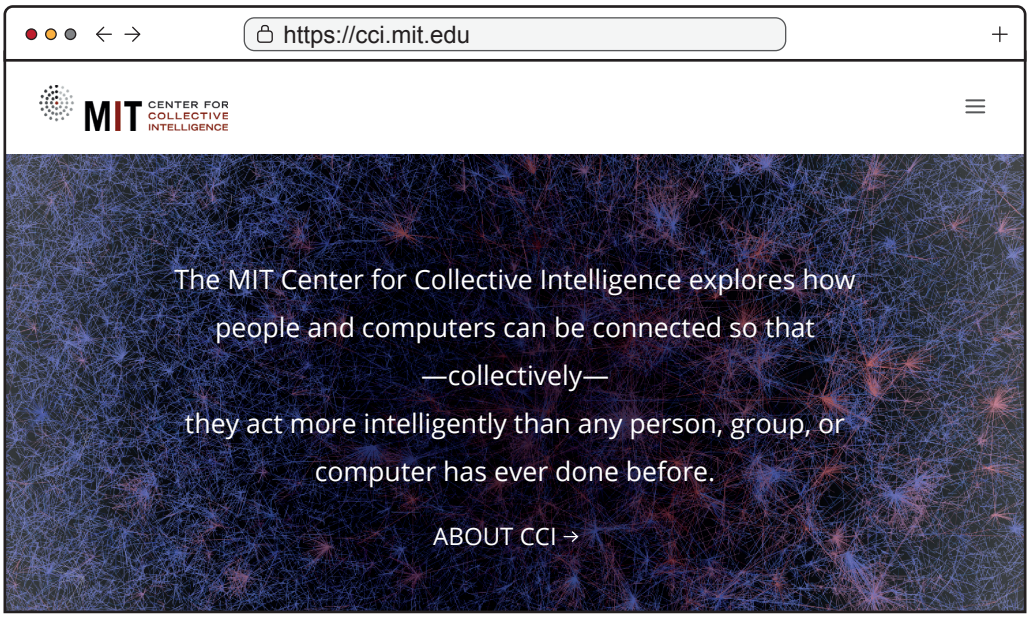
(d) Explain how Google Lens might perform this search. [3]

Folksonomies can also be considered as a form of collective intelligence.

Many organizations are using collective intelligence to solve complex problems.

One example is the Massachusetts Institute of Technology (MIT) who set up a Center for Collective Intelligence (CCI) (see **Figure 9**).

Figure 9: The MIT CCI website



(Option C continues on the following page)

(Option C, question 13 continued)

Professor Thomas W. Malone from MIT stated, “Migrating our Climate CoLab community into Wazoku’s global crowd of problem solvers can accelerate the development of solutions that will make for a greener, more sustainable future.”

(e) Evaluate the decision of MIT to use collective intelligence to solve complex problems such as climate change. [6]

End of Option C

Option D — Object-oriented programming

14. A school requires all its 250 students to go on a four-day field trip. The school offers 16 different trips. A maximum of 25 places are available on each trip.

Each student submits a request form with their top 10 trip preferences.

The trip coordinator assigns each student a place on a trip based on their preferences and the availability of places on that trip.

A trip will not run if fewer than 10 places are filled after all the student requests have been processed. For example, if 9 or fewer students have requested the second trip, it will not run.

The `Trip` class keeps the details of each trip. Part of the code for this class is shown below:

```
public class Trip {

    private int tripNumber;    // id of the trip
    private String tripTitle;  // trip Name
    private String tripDesc;   // trip description
    private int placesFilled;  // number of places filled
    private boolean isRunning; // status of the trip
    private ArrayList <String> tList = new ArrayList<String>();

    public Trip(int tripNumber, String tripTitle, String tripDesc) {
        this.tripNumber = tripNumber;
        this.tripTitle = tripTitle;
        this.tripDesc = tripDesc;
        this.placesFilled = 0;
        this.isRunning = false;
    }

    public String getTripTitle() {return tripTitle;}

    public ArrayList<String> getList() {return tList;}

    public int getPlacesFilled() {return placesFilled;}

    public boolean isRunning() {return isRunning;}

    public void setTripTitle(String tripTitle) {
        this.tripTitle = tripTitle;}

    public void updatePlacesFilled() {
        this.placesFilled++;
        // code missing to change the trip running status
    }

    public void addStudentToTripList(Student s){
        tList.add(s.getName());}
        // all accessor and mutator methods are present but not shown

}// end of Trip class
```

(Option D continues on the following page)

(Option D, question 14 continued)

The `Student` class keeps the details of a student. Part of the code for this class is shown below:

```
public class Student {

    private String sId; // unique id of the student
    private String sName; // name of the student
    private int sGrade; // grade level of the student
    private String tripAssigned; // trip Title assigned to the student
    public int[] sPreferenceList = new int[10]; // preferred trip numbers

    public Student(String sID, String sName, int sGrade,
    int[]sPreferenceList) {
        this.sId = sId;
        this.sName = sName;
        this.sGrade = sGrade;
        this.sPreferenceList = sPreferenceList;
        this.tripAssigned = "zzz";
    }

    public void setTripAssigned(String tripAssigned) {
        this.tripAssigned = tripAssigned;}

    public String getName() {return sName;}

    public String getTripAssigned() {return tripAssigned;}

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

}// end of Student class
```

(Option D continues on the following page)

(Option D, question 14 continued)

The `TripAllocator` class has the main method and other methods to assign trips and generate the information required.

```
public class TripAllocator {

    private Trip[] allTrips = new Trip[16];
    // trips are stored in the ascending order of the trip number

    private Student[] allStudents = new Student[250];

    public void assignTrip(Student s) // assigns a trip to a student
    {// code missing}

    public void displayPlacesLeft() {// code missing}

}
```

(a) State the reason why the size of the `allTrips` array is 16. [1]

(b) State the **two** parts of a method signature that are required in all methods. [2]

(c) Outline the concept of encapsulation with reference to the given code. [2]

(d) Describe how the `updatePlacesFilled()` method would change the trip running status to `true`. [2]

(e) Construct the `assignTrip(Student s)` in the `TripAllocator` class. [6]

Assume the student preference list has the most preferred trip number at position 0.

This method:
- should assign the most preferred trip to a student based on their preference list if a place is available on that trip
- should make required updates to the correct trip object.

(f) Outline why another method, `reAssignTrip(Trip t)`, will be required in the `TripAllocator` class. [2]

(Option D continues on the following page)

(Option D continued)

15. Some students asked the trip coordinator if they could submit a preference list jointly with one friend. This would ensure that they have at least one friend on their trip.

(a) Identify **one** modification that this request would require in the `Student` object. [1]

(b) Describe, without writing code, **two** modifications that this request would require in the method `assignTrip(Student s)` for the program to work correctly. [4]

Assume that the `assignTrip()` method has been executed for all students.

(c) Construct the method `displayPlacesLeft()` in the `TripAllocator` class. This method should output the trip number, trip title and the number of places left for each trip that is running but is not yet full.

Your method must present the output in the format below.

Example output:
Trip 1: Mountain trails Places left: 6
Trip 5: Local craft appreciation Places left: 3 [5]

It is possible that some students could not be assigned any trip after the first round of program execution.

(d) Construct a method `noTrip()` to return the number of students who could not be assigned any trip at the end of the first round of program execution. You should include the method signature as part of your answer. [5]

(Option D continues on the following page)

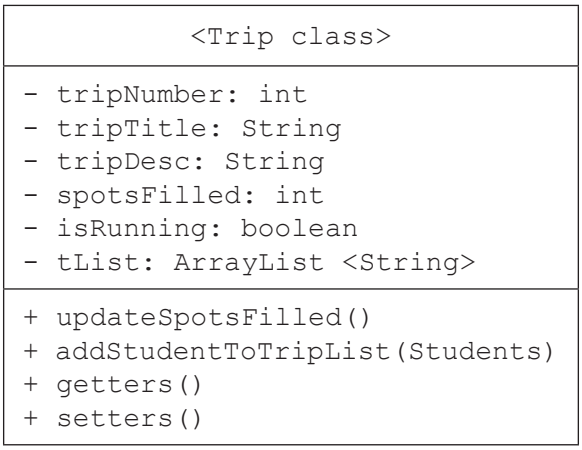
(Option D continued)

16. The `Trip` class is shown in **Figure 10**.

A trip may be classified as an adventure trip, a service trip or a culture trip based on certain characteristics. Hence, three new classes `AdventureTrip`, `ServiceTrip` and `CultureTrip` must be created.

- Adventure trips require students to be physically fit and bring trekking shoes.
- Service trips require students to have good communication skills.
- Culture trips require students to bring a camera.

Figure 10: UML diagram of the `Trip` class



(a) (i) Construct a UML class diagram that shows the relationships between the `Trip` class and the three new classes, including additional attributes and methods. [5]

(ii) State how the class relationship of part (a)(i) is written in code. [1]

(iii) Outline **two** advantages of inheritance in object-oriented programming (OOP). [4]

(iv) State why there is a need to reduce dependencies between objects. [1]

(b) Outline **two** advantages of modularity in program development. [4]

(Option D continues on the following page)

(Option D continued)

17. The popularity of a trip depends on the number of requests received for that trip.

Information on the number of student requests for each trip over the past five years is stored in a two-dimensional array, `popularity` (see **Figure 11**).

Figure 11: A representation of the two-dimensional array, `popularity`

Year / Trip number	2023	2022	2021	2020	2019
1	16	20	15	17	14
2	23	26	17	27	29
3	11	19	18	20	16
4	10	12	20	21	23
5	8	12	15	11	10
…	…	…	…	…	…

The rows represent the number of student requests for a trip, with the first row being Trip 1. The columns represent the year.

(a) Construct the method `avgPopularity(int[][]popularity, int[]tripPop)` to store the average popularity of each trip in the `tripPop` array. [4]

(b) Construct the code needed to output the trip titles that have a lower popularity value in 2023 compared to their average popularity value. You must make use of any previously created arrays. [3]

Two more attributes are specified for each trip: departure date and departure time.

To ensure correct departures, trip objects are inserted in a linked list in order of departure time with the earliest first.

On day one, Trip 3 is scheduled to depart earliest, followed by Trip 11, then Trip 2, and Trip 5 which is scheduled to depart last.

Due to unforeseen circumstances, Trip 11 has been delayed and will now leave after Trip 2.

(c) Outline the sequence of steps required when changing the pointers in the linked list to create the new departure order. You may use an annotated diagram. [3]

(Option D continues on the following page)

(Option D continued)

18. The trip coordinator gave students one week to submit their request form. Each student's unique ID is added to a stack in the order the request forms were submitted.

A queue may be used as an alternative data structure to hold this data.

(a) (i) Outline **one** advantage of using a queue in this scenario. [2]

(ii) Outline **one** disadvantage of using a queue in this scenario. [2]

(b) Consider the following recursive algorithm:

```
ArrayList <String> aList = new ArrayList <String>();
public void aMethod(int n)
{
    if (n<0)
        return;
    else
    {
    if(allTrips[n].isRunning()==false)
        for (int i=0; i<allTrips[n].getList().size(); i++)
            aList.add((String)allTrips[n].getList().get(i));
                aMethod(n-1);
    }
}
```

(i) Outline the purpose of this algorithm when n = 15. [2]

(ii) Construct a non-recursive method that will achieve the same purpose as the given `aMethod(int n)`. [2]

(iii) Outline **one** benefit of the styling convention used in the code to the programming team. [2]

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:

Figure 3 Hannah Ritchie, Lucas Rodés-Guirao, Edouard Mathieu, Marcel Gerber, Esteban Ortiz-Ospina, Joe Hasell and Max Roser (2023) - "Population Growth" Published online at OurWorldinData.org. Retrieved from: 'https://ourworldindata.org/population-growth' [Online Resource]. Data source: United Nations, Department of Economic and Social Affairs, Population Division (2024). World Population Prospects 2024, Online Edition - Copyright © 2024 by United Nations, made available under a Creative Commons license CC BY 3.0 IGO: http://creativecommons.org/licenses/by/3.0/igo/. Source adapted.

Figure 4 With permission from Alasdair Rae.

Figure 5 [logo] Latypova, G., n.d. *Education Book Logo. Vector design – stock illustration*. [image online] Available at: https://www.gettyimages.co.uk/detail/illustration/education-book-logo-vector-design-royalty-free-illustration/1221128440?phrase=school+logo&adppopup=true [Accessed 22 August 2023]. Source adapted.

[photo] supersizer, n.d. *We can make it work - stock photo*. [image online] Available at: https://www.gettyimages.co.uk/detail/photo/we-can-make-it-work-royalty-free-image/1403446891?adppopup=true [Accessed 16 August 2023]. Source adapted.

Figure 8 Fujita, Y., Kichikawa, Y., Fujiwara, Y., Souma, W. and Iyetomi, H., 2019. Local bow-tie structure of the web. Applied Network Science. https://www.researchgate.net/publication/332512547_Local_bow-tie_structure_of_the_web. Open access article under Creative Commons Attribution 4.0 International License http://creativecommons.org/licenses/by/4.0/. Source adapted.

Figure 9 With permission from the MIT Center for Collective Intelligence.

13. [Quotation] With permission from the MIT Center for Collective Intelligence.