Diploma Programme
Programme du diplôme
Programa del Diploma

Computer science
Standard level
Paper 2

Tuesday 21 May 2019 (morning)

1 hour

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 **[45 marks]**.

Option	Questions
Option A — Databases	1 – 3
Option B — Modelling and simulation	4 – 6
Option C — Web science	7 – 9
Option D — Object-oriented programming	10 – 12

2219 – 7015

Option A — Databases

1. *Marble Reading Book Stores* (*MRBS*) is a chain of bookstores based in London. The stores want to keep information about the books they sell, the authors of the books and the publishers they work with. The assumptions made when the database was created were:

- a publisher can publish books from one or more authors
- an author can write one or more books.

(a) Construct the entity-relationship diagram (ERD) for this scenario. [2]

Three of the tables in the *MRBS* database are shown below:

PUBLISHER

Publisher_Name	**City**	**Country**	**Telephone**
Orlando Crux	Melbourne	Australia	6187675423
Owen Troy	Taipei	Taiwan	8867843525
Philip Hall	Los Angeles	USA	1546838382

AUTHOR

Author_Num	**Author_Name1**	**Author_Name2**	**DOB**	**Publisher_Name**
OC80	David	Gully	05/06/1974	Orlando Crux
OC89	Clint	Donald	02/12/1957	Orlando Crux
OT66	Steve	Leking	11/07/1989	Owen Troy
PH54	Mary	West	23/12/1990	Philip Hall

BOOK

ISBN	**Book_Title**	**Genre**	**Author_Num**	**Year_Pub**
0-12763-777-1	Recipes of the East	Non-Fiction	PH54	2014
0-65432-187-1	Ken and his life	Fiction	OC89	2014
0-66655-916-2	All about the Grand Canyon	Geography	PH54	2012
0-76544-987-2	Tidings	Fiction	OT66	2015
0-87022-176-0	The fair price of life	Fiction	OC80	2014
0-98124-612-2	Seeking the truth	Non-fiction	OT66	2016

(b) Outline why data validation is difficult for the Book_Title attribute. [2]

(Option A continues on the following page)

(Option A, question 1 continued)

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

```
SELECT Book_Title
FROM BOOK
WHERE Genre = "Non-fiction"
AND ISBN = '0-98124-612-2'
```

(d) Construct a query to find the titles of the books published by “Orlando Crux”. [4]

The *MRBS* database undergoes many transactions.

(e) Outline why atomicity is important within a database. [2]

(f) Outline how data consistency can be maintained in transactions in this database system. [2]

Some data in the *MRBS* database is redundant.

(g) Outline **one** problem caused by redundant data. [2]

2. A school maintains a database of students’ details and teaching resources on a central server. This data can be accessed by all teachers in the school.

Teachers may need to edit resources when preparing their lessons.

(a) Explain how concurrent use of the school database is possible in this situation. [3]

When storing student details, data security is an important consideration.

(b) Describe **two** ways that data security in the school's database can be maintained. [4]

The school has appointed a database administrator (DBA).

A DBA is required to carry out tasks such as ensuring there is a strategy to recover the database if it becomes corrupted and that the data is shared ethically.

(c) (i) Describe **one** strategy that could be used to ensure the data can be recovered if the database becomes corrupted. [2]

(ii) Suggest how the privacy of student data can be ensured. [3]

(Option A continues on page 5)

Blank page

(Option A continued)

3. *Armour Hardware Company* has the following data about salespersons and the quantities of items sold.

Each salesperson can sell many different products.

SALES_PERSON

Product_ Number	Unit_ Price	Product_ Name	Date_And_ Time	Sales_ Person_ Number	Sales_ Person_ Name	Manager_ Number	Manager_ Name
19440	12.50	Saw	03/07/2018 12:23:34	102	Owen	16	Benson
32456	14.50	Hammer	03/07/2018 12:56:23				
35647	35.00	Drill	03/07/2018 12:35:02	199	Dahl	45	Rogers
67895	13.25	Wrench	03/07/2018 12:49:56				
98760	12.25	Pliers	03/07/2018 13:23:34				
67896	9.25	Wrench	03/07/2018 12:46:23	154	Fraser	16	Benson
32456	14.50	Hammer	03/07/2018 12:50:16				
35647	35.00	Drill	03/07/2018 12:55:09	234	Robert	16	Benson
32456	14.50	Hammer	03/07/2018 12:57:12				
67895	13.25	Wrench	03/07/2018 13:25:36				

(a) Outline **two** reasons why databases are normalized. [4]

(b) Outline why the **SALES_PERSON** table is not in 1st Normal Form (1NF). [2]

(c) Construct the 3rd Normal Form (3NF) of the unnormalized relation shown above. [8]

(d) Outline why it is necessary to ensure that referential integrity is maintained in databases. [2]

(e) Outline why a primary key may consist of more than one attribute. [2]

End of Option A

Option B — Modelling and simulation

4. The rise in global temperatures has led to the melting of the sea ice in polar regions. Scientists have developed a number of computer models that can be used to make predictions about the rate of sea ice melting in these polar regions and its effect on coastal areas globally. In order to create a computer model a number of variables are identified.

A computer model of the effects of the melting of the sea ice in the Arctic Ocean may include the following variables:

- average ocean surface temperature (°C)
- albedo of the ocean (the proportion of light reflected from the ocean surface)
- precipitation (mm)
- salinity of the ocean (grammes of salt in one kilogramme of water)
- area of sea ice (km^2).

(a) Copy and complete the following table showing each variable's data type and a suitable range of values that would represent the information shown above. [3]

Variable	Data type	Range of values
Ocean surface temperature		
Albedo		
Area of sea ice		

In this model the following rules have been determined:

- for every 0.01 °C increase in ocean surface temperature, the area of sea ice decreases by 1 %
- for every 1 % decrease in the area of the sea ice, the sea level rises by 20 mm.

The initial values are:

- area of sea ice = 1 000 000 km^2
- average surface temperature of the ocean is 0.00 °C.

(b) Using the rules and initial values above, construct the pseudocode that would enable the area of the sea ice and the sea level rise to be calculated if there was an increase of 0.04 °C in the ocean surface temperature. [4]

(c) Using the information above state:

(i) the area of the sea ice. [1]

(ii) the change in sea level. [1]

(Option B continues on the following page)

(Option B, question 4 continued)

The scientists observed when running the model numerous times using historical data there were significant differences between observed and expected results.

A second model was developed that included new variables and rules.

The surface of the ocean reflects the heat from the sun. The ratio between the area covered by the sea ice and the area where there is no sea ice (open ocean) affects the value of the average albedo. The lower the albedo, the quicker the sea ice will melt.

The average albedo is calculated using this formula:

$$\text{Average albedo} = \frac{(\text{area of sea ice} \times \text{albedo of sea ice}) + (\text{area of open ocean} \times \text{albedo of open ocean})}{(\text{area of sea ice} + \text{area of open ocean})}$$

Note:
- area of sea ice = 1 000 000 km^2
- area of open ocean = 1 000 000 km^2
- albedo of sea ice = 0.6
- albedo of open ocean = 0.1.

The average albedo will change with every iteration of the model. Each iteration is 2 years after the previous.

The rules for the model are:
- the initial albedo is 0.35
- the rate of decrease in sea ice every 2 years is $\frac{0.3}{\text{average albedo}^2}$
- the starting year is 2019
- the sample rate is every 2 years.

(d) Using the formula, rules and initial data given above, construct the pseudocode that would calculate the year that the area of sea ice will be less than 10 000 km^2. [7]

(e) Identify **two** ways that this model could be implemented. [2]

(f) Explain why the accuracy of the simulation in predicting the area of the sea ice is critical. [3]

(Option B continues on the following page)

(Option B continued)

5. Many health agencies are using simulations in an attempt to understand how their resources could be used in the future. With many countries experiencing aging populations, health agencies have worked with computer scientists to develop simulations that will enable them to manage their resources more effectively.

One of the key features of these simulations is the development of "what-if" models.

(a) Describe the main features of a "what-if" model. [4]

The following variables can be considered as part of a model to be used to simulate the management of an aging population:

- quality of health education
- lifestyle choices such as smoking
- residential region.

(b) Identify **three** other variables that could be included in this model. [3]

(c) Explain the ethical issues that may arise from the collection of information for this model. [5]

(d) Explain why the model would be converted to a simulation. [3]

6. An old set of 2D animated cartoons from the 1940s has been discovered and it is decided to modify them to turn them into 3D animation.

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

(b) Outline the need for rendering in the creation of the animated 3D characters. [2]

(c) Explain **two** technical implications of implementing a 3D animation in this way. [6]

End of Option B

Option C — Web science

7. Sestra.com is a website maintained by a company who sell items made by local craftspeople.

The website is compatible with different screen sizes and formats ranging from desktop computers to mobile smartphones. All the site's pages contain the following code fragment:

```
<link rel = "stylesheet" href = "../css/default.css">
```

(a) Identify **two** ways that a cascading style sheet (CSS) can be used to ensure web pages are compatible with different screen sizes and formats. [2]

Visitors to the site can search categories of products (for example "Toys", "Bags", "Dresses" *etc*) selected from a drop-down menu. The menu is populated from the records stored in the `CATEGORY` table of the site's database.

Parts of the code of the file `search.php` is shown below:

```
// Other code present here
<?php
  $categoryquerytext = 'SELECT 'category_id', 'category_name' FROM
  'CATEGORY' ORDER BY 'category_name'';
  $categoryqueryresult = mysqli_query($con, $categoryquerytext);
?>
// Other code present here
<form action = "showresults.php" method = "post">
  <select name = "category">
  <?php
    while($row = mysqli_fetch_array($categoryqueryresult))
    {
      echo '<option value = "'. $row['category_id']. '">'.
      $row['category_name']. '</option>';
    }
  ?>
  </select>
  <button type = "submit">Search</button>
</form>

// Other code present here
```

(b) Explain the processing this code enables on the server before `search.php` is sent to the client. [3]

The owners of the company have noticed that Sestra.com does not appear very prominently in search engine results.

(c) Describe **two** ways in which the site developers could use white hat optimization to improve the site's search engine ranking. [4]

(Option C continues on the following page)

(Option C, question 7 continued)

The Sestra.com site includes:
- images of each product
- pdf documents giving background information about the craftspeople who produced the products.

(d) Distinguish between lossy and lossless compression. [2]

(e) Explain why the developers at Sestra.com would use lossless compression for the pdf documents. [3]

8. The Large Hadron Collider at CERN in Switzerland produces an average of 15 petabytes (15 million gigabytes) of experimental data every year. This data must be accessed and analysed by scientists around the world.

(a) With reference to the URL https://home.cern/topics/large-hadron-collider

(i) State the protocol used. [1]

(ii) Identify the steps taken by the domain name server when the scientist enters a URL such as https://home.cern into their web browser. [3]

CERN has established the *Worldwide LHC Computing Grid*.

(b) Explain **two** reasons why CERN would use grid computing to support its research. [6]

Instead of copyrighting its experimental results, CERN has decided to publish its experimental results using Creative Commons licensing.

(c) Explain **two** reasons why CERN would publish its experimental results using Creative Commons licensing. [6]

(Option C continues on the following page)

(Option C continued)

9. *Brownsville Council* run several public libraries in different areas of the city. The libraries use an Integrated Library Management System (ILMS) to manage all items (for example books, DVDs, *etc*) held by the libraries. Details of the items are stored in a database on a central server.

Below is part of the extensible markup language (XML) code used to describe an item.

```
<item id = "97812">
  <category>Book</category>
  <author>Stark, Elizabeth</author>
  <title>Handheld Device Usability</title>
  <genre>Computer Science</genre>
  <publisher>Taylor & Orams Inc.</publisher>
</item>
```

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

(b) Outline **one** advantage of XML for sharing data on the web. [2]

XML is based on open standards.

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

Library users interact with the ILMS through a web page that includes a form to search for items stored on the database.

(d) Describe the role of the common gateway interface (CGI) in processing search requests made via the web form. [2]

The library managers have decided to extend their web pages to include a blog and a forum, maintained by the head librarian, in order to increase engagement with library users.

(e) Distinguish between a blog and a forum. [2]

(f) To what extent has the use of social media, blogs and forums enabled the head librarian to be a more effective decision maker? [6]

End of Option C

Option D — Object-oriented programming

An international school organizes a regional swimming competition for students from 10 different schools. Each school will send a team of 5 to 15 swimmers.

Each swimmer can enter up to 5 events (such as the “50 m freestyle” or “100 m butterfly”).

Each event consists of one or more races. A race can be a qualifying heat, or a final. The final has the best 8 swimmers from all the qualifying heats in the event.

Each race has a maximum of 8 swimmers.

The UML diagrams for the classes `Swimmer` and `Race` are provided below

Swimmer
- String name - String school - String[5] eventID - **double**[5] time
+ constructor + accessor and mutator methods + addTimes()

Race
- Swimmer[8] swimmer - **double**[8] time
+ constructor + accessor and mutator methods + addSwimmers() + addTimes()

10. (a) Define the term *mutator method*. [1]

(b) State **one** additional instance variable of type `boolean` which could be added to the class `Race` as indicated above. [1]

(c) With reference to both class UMLs provided above, distinguish between a class and an instantiation. [3]

In this scenario, `Swimmer` objects are aggregated in a `Race` object.

(d) (i) Outline **one** advantage of using aggregation in this context. [2]

(ii) Outline **one** disadvantage of using aggregation in this context. [2]

(e) Construct code for the constructor of the class `Swimmer` that instantiates an object with parameters `name` and `school`. The event IDs should be set to “empty” and the times to 0.0 [4]

Many swimmers in the event have names that cannot be represented using basic character sets such as ASCII.

(f) Describe **one** feature of modern programming languages that allows the wide range of students’ names to be represented correctly. [3]

(Option D continues on the following page)

(Option D continued)

11. A generic `Event` class is defined as follows:

```
class Event
{
  private String eventID;
  private int numberOfRaces;
  private Race[] races;
  private Race finals;

  public Event(String ID, int numberOfRaces)
  {
    eventID = ID;
    races = new Race[numberOfRaces];
    for(int i = 0; i < numberOfRaces; i++)
    {
      races[i] = new Race();
    }
    finals = new Race();
  }

  public void addSwimmers()
  {
    // fills the qualifying heats with swimmers
  }

  public void fillFinals()
  {
    // fills the finals race with the best 8 from the qualifying heats
  }

  // more methods()
}
```

(a) The same method identifier `addSwimmers` is used in both classes `Race` and `Event`.

Explain why this does not cause a conflict. [3]

The `Event` class above assumes that the event has more than 8 swimmers and requires qualifying heats. However, an event with less than 9 swimmers has no qualifying heats, so the original `Event` class was inherited by a new class `FinalsOnlyEvent`.

(b) Outline **two** advantages of the OOP feature “inheritance”. [4]

(c) Outline how method overriding can help to create the new class `FinalsOnlyEvent`. [2]

(Option D continues on the following page)

(Option D continued)

12. An `Event` has been instantiated with 2 qualifying heats for a total of 11 swimmers.

```
Event free100 = new Event("100 m free style",2);
```

The swimmers were added to the two `Race` arrays and after the races, their times were recorded as shown in the table.

(For the purpose of this question, the name represents the full swimmer object.)

`races[0]`

`swimmer`	Andy	Bella	Chris	Duc	Eric	null	null	null
`time`	34.2	33.8	40.9	36.3	34.6	0	0	0
	[0]	*[1]*	*[2]*	*[3]*	*[4]*	*[5]*	*[6]*	*[7]*

`races[1]`

`swimmer`	Fiona	George	Hetty	Idan	Jo	Karl	null	null
`time`	41.2	36.6	37.6	35.2	48.8	37.2	0	0
	[0]	*[1]*	*[2]*	*[3]*	*[4]*	*[5]*	*[6]*	*[7]*

The method `fillFinals()` will select the 8 fastest swimmers, in ascending order of time, from both `swimmer` arrays and copy them to the `swimmer` array in the `finals` race.

(a) Sketch the resulting `swimmer` array in `finals`. [3]

To help with this selection, all entries from `races[0]` and `races[1]` will be copied into two new parallel arrays of size 16, one array for swimmers and one array for their times.

(b) Construct the code fragment for the given situation that will copy swimmers and times into two parallel arrays named `tempSwimmer` and `tempTime`. [6]

(Option D continues on the following page)

(Option D, question 12 continued)

The two temporary arrays will be sorted using the following code.

```
int i,j;
Swimmer swapSwimmer;
double swapTime;
for(i = 0; i < 15; i++)
{
  for(j = 0; j < 15; j++)
  {
    if(tempTime[j] > tempTime[j + 1])        // if wrong order then…
    {
      swapSwimmer = tempSwimmer[j];          // swap the swimmer and…
      tempSwimmer[j] = tempSwimmer[j + 1];
      tempSwimmer[j + 1] = swapSwimmer;
      swapTime = tempTime[j];                // swap the time
      tempTime[j] = tempTime[j + 1];
      tempTime[j + 1] = swapTime;
    }
  }
}
```

(c) (i) State the name of this sorting algorithm. [1]

(ii) Outline **two** improvements to this code that would make the algorithm more efficient. [4]

(d) Construct the code fragment that will copy the names of the 8 fastest swimmers in ascending order of time from the array `tempSwimmer` to the array `swimmers` in the race `finals`. [6]

End of Option D