Diploma Programme
Programme du diplôme
Programa del Diploma

Computer science
Higher level
Paper 2

Monday 5 November 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 – 8
Option C — Web science	9 – 12
Option D — Object-oriented programming	13 – 16

17 pages

International Baccalaureate
Baccalauréat International
Bachillerato Internacional

Blank page

Option A — Databases

1. (a) Outline the difference between an information system and a database. [3]

A bank maintains a database that stores details of clients and their accounts.

A client wants to transfer money between two accounts held at the same bank.

(b) Explain how the ACID (Atomicity, Consistency, Isolation, Durability) properties would apply in the context of this database transaction. [8]

A bank holds large volumes of financial and personal information about its clients in its database.

(c) Discuss whether this database should be open to interrogation by the police or the Government. [6]

2. (a) Define the term *database management system (DBMS)*. [1]

(b) Identify **three** characteristics of logical schema. [3]

(c) Outline the purpose of a data definition language (DDL). [2]

(d) Identify **two** tasks that a database administrator carries out to ensure the security of the database. [2]

(Option A continues on the following page)

(Option A continued)

3. A shop sells only fruits and vegetables. Data about the products sold is held in the relation Greengrocer as follows:

GREENGROCER

Prod_ID	Prod_Name	Prod_Price	Supp_Name	Supp_Contact	Supp_Phone
P211	Apples	5.00	Fruit and Veggie	Ines Bond	542 9800
P118	Celery	11.00	Veggy Co.	Mia Abiss	234 7654
P121	Onions	3.00	Fruit and Veggie	Ines Bond	542 9800
P122	Onions	2.45	Veggy Co.	Mike Oberst	235 7656
P219	Pears	8.50	Fruit and Veggie	Kim Ng	542 9801
P220	Pears	9.00	Veggy Co.	Mike Oberst	235 7656
P111	Pumpkin	6.00	Fruit and Veggie	Mia Abiss	542 9802
P265	Tomatoes	7.45	New Fruits	John Smith	677 7895

Prod_ID is the primary key.

The relation above would be represented using the following notation:

Greengrocer (Prod_ID, Prod_Name, Prod_Price, Supp_Name, Supp_Contact, Supp_Phone)

(a) (i) Identify an appropriate data type for Prod_Price. [1]

(ii) Construct the steps in a query to output the names of all products supplied by *Veggy Co.* with prices in the range from 4.00 to 10.00 inclusive. [3]

(b) Outline the redundant data in this relation. [2]

(c) Explain the problems that may arise for each of the following functions performed on the given relation. In your answer, you should give an appropriate example.

(i) Inserting a new tuple with an item that is supplied by existing supplier. [2]

(ii) Deleting a tuple from the relation. [2]

(iii) Modifying a specific attribute value in the tuple. [2]

(d) Construct the database in the 3rd normal form (3NF) using the notation specified above. You must show all your workings. [8]

(Option A continued)

4.   The *Ministry of Tourism* intends to include data from a number of sources in a data warehouse in an attempt to improve the services the tourism companies offer to their customers.

   (a)   Identify **four** sources of information that could be incorporated in the data warehouse.   [4]

   (b)   Outline **two** differences between a data warehouse and database.   [4]

   (c)   Explain how the Extract, Transform, Load (ETL) processes can be used to address the problems related to data migration.   [3]

   The *Ministry of Tourism* will use data mining to provide the Government with information that it can use to promote and support tourism in the country.

   (d)   Outline **one** ethical problem that may result from data mining.   [2]

   (e)   Explain how cluster analysis can be used to improve the advertising strategy of the tourism companies.   [4]

   (f)   Explain the importance of link analysis in exploring patterns in data mining.   [3]

End of Option A

Option B — Modelling and simulation

5.　A water authority supplies water to its customers. The amount of water supplied to each customer is measured by a meter. There are two types of meter; an old meter and a new meter.

　　The water authority also provides sewage services. This is when waste water is returned from the customers' houses to a central processing plant. Only the new meter is able to estimate the amount of waste water returned.

　　The authority reads the meters annually and the charges for these services are sent to the customers.

　　The annual charges to customers are calculated using a model that takes into account the following variables:

METER_TYPE　　　The type of meter which can be either old or new.

WATER_IN　　　　The reading displayed on <u>both</u> meters (the reading is a 6-digit display in litres) which is used to measure the volume of water supplied.

WATER_OUT　　　The reading displayed on the <u>new</u> meter only (the reading gives a 6-digit integer display in litres) which is used in estimating the volume of waste water returned to the sewage system.

INFRA_CHARGE　　The cost of maintaining the infrastructure for supplying water and providing the sewage service. This is an annual charge of $204.80. This amount is reduced by 10 % if a new meter is used.

　　(a)　Identify the possible range of values **and** their data type for each of the above variables.　[3]

The digital displays on each meter function as follows:
- The digital display on each meter increments by one every time a litre of water passes.
- If the display reaches its maximum value it resets to 0 and then continues to function as before.

　　(b)　Outline how the yearly supply of water is calculated.　　　　　　　　　　　　[2]

The rules for calculating the annual bill for each household are as follows:
- The price for each litre of water supplied is $1.90.
- The total annual bill is the sum of the charge for the water supplied and the infrastructure charge.
- If the household has a new meter and at least 95 % of the water supplied is returned to the sewage system, then the infrastructure charge for that household is reduced by 10 %.

　　(c)　Construct the pseudocode that will calculate the annual bill for a household based on the information given above. You should introduce any new variables where necessary.　[7]

　　(d)　Identify **three** possible measures that customers could take to reduce their infrastructure charge.　　　　　　　　　　　　　　　　　　　　　　　　　[3]

(Option B continues on the following page)

(Option B continued)

6. The average weekly consumption for a family of four is 2700 litres of water. This includes the water needed for domestic cleaning, personal hygiene, cooking and drinking.

An online software tool is available to help families calculate how much water they are using for regular activities that require the supply of water. The software tool includes a model which estimates the water usage. A simulation can be run using this model.

(a) With reference to the scenario described above, explain the difference between a model and a simulation. [4]

The data given below estimates the amount of water (in litres) needed for different household activities:

Activity	Litres
Dishwasher, full load	15
Washing machine, full load	60
Bath	70
Shower	35
Toilet flush	9

(b) Explain how the choice of data in the table may affect the quality of the simulation. [4]

By using the data shown above, a family uses a spreadsheet to run some simulations of how to distribute the use of water.

(c) Outline how a simulation in a spreadsheet can be organized so that the family can ensure a consumption of no more than 2700 litres per week. [5]

The same spreadsheet is used by a group of four young people who share a house. The simulation, however, does not reflect their actual situation and the weekly target is often exceeded.

(d) Identify **two** reasons why the simulation may not reflect the actual situation in this case. [2]

(Option B continues on the following page)

(Option B continued)

7. Medical students require extensive training in preparation before they qualify as a surgeon. Part of this training involves the use of simulation software.

    (a)   Outline **one** advantage and **one** disadvantage of training medical students with simulation software.   [4]

    The simulation software projects the surgery scenario onto a touchscreen display. Students wear "smart" gloves to hold tools for performing the virtual operation on the touchscreen monitor. The touchscreen display shows realistic surgical tools.

    (b)   Identify **three** technical aspects that the gloves have to address so that they can be used in this way for the simulation of operations.   [3]

    (c)   Define the term *visualization*.   [2]

    The medical school is considering updating its simulation software, which currently displays the simulation in 2D, to one that displays it in 3D.

    (d)   Discuss whether this change to a 3D display would improve the quality of the training.   [6]

8. Genetic algorithms follow an iterative process in developing a solution to a problem.

    (a)   Outline the steps involved in this iterative process.   [6]

    A parasite is destroying olive trees in a region in southern Europe. This area is monitored by taking aerial digital images at midday each day. A neural network is then used in the analysis of these images.

    (b)   (i)   Outline why neural networks need to be trained before they can be used in this analysis.   [4]

          (ii) Explain why supervised learning may be preferred to unsupervised learning for training these networks to make predictions on the progress of this parasitic disease.   [6]

    Many companies now use an automated response system when answering basic telephone queries from customers. This system uses speech analysis together with recorded messages.

    (c)   Explain why the content of the messages given by these systems must be carefully chosen.   [4]

End of Option B

Option C — Web science

9. The online edition of a newspaper includes a collection of audio podcasts. These podcasts are interviews with artists and scientists and last for up to 10 minutes. They may have been provided by other broadcasters and agencies.

(a) Identify **two** actions relating to the source of these podcasts that should be taken by the newspaper before publishing them. [2]

At any time, no more than eight podcasts are available on the online edition of the newspaper.

(b) Outline **one** technical reason for restricting the number of podcasts to eight. [2]

The podcasts can either be downloaded or streamed.

(c) State **one** format for an audio file. [1]

The podcast can be stored online using either lossy or lossless compression.

(d) Compare these **two** approaches for a person who downloads or streams these podcasts. [5]

A person is streaming one of these podcasts at home. However, he is experiencing buffering problems, where the podcast keeps stopping and starting.

(e) Outline **one** reason for this problem. [2]

These podcasts are normally displayed on the website in the chronological order in which they were added, with the most recent on top. However, users who regularly access the site for science podcasts, may see them presented in a different order, with all the scientific ones on top.

(f) With reference to the technology involved, explain how this is made possible. [4]

(Option C continues on the following page)

Turn over

(Option C continued)

10. *BesTea* is a company that sells tea and tea-related accessories online. Users browsing the website encounter the following URL:

    https://www.BesTea.com/customer/pages/delivery

    (a) Describe how this URL provides security in communication over the Internet. [2]

    A person enters the URL for *BesTea* into their browser.

    (b) Identify the steps that the Domain Name System (DNS) server will take in order to locate the correct IP address for this request from the browser. [4]

    The logos of the credit and debit cards accepted by *BesTea* are displayed as images in the footer of all of its web pages.

    (c) Outline how the footer of a web page can be made identical across all pages of a website. [2]

    The *BesTea* website includes a shopping basket facility that enables users to make purchases.

    The following fragment of PHP code is present in the script that allows users to place a completed order:

```php
<?php
   …
$_SESSION['sessionUserID'] = new_random_number();
$query = $db -> query ("SELECT * FROM users_db WHERE id = "
.$_SESSION['sessionUserID']);
$userRecord = $query -> fetch_record();
…
>
```

    (d) Suggest the function of this fragment of code. [4]

    Users may perform several actions on their shopping baskets. These actions require interaction with a central database of products. Each record in this database includes the identifier, the name and the price of one of these products.

    Consider the following fragment of code:

```php
<
$basket = new Basket
   …

If $_REQUEST['action'] == 'removeBasketItem' && !empty($_REQUEST['id']){
remove($basket, $_REQUEST['id']);
header("view_basket.php");
}
>
```

    (e) From the code, identify the information that the user has provided. [2]

(Option C continues on the following page)

(Option C continued)

11. A charitable organisation, *Harvest*, which operates in a less economically developed country (LEDC), uses old computers that have been donated to them. The organization relies on WiFi and satellite-based transmission for access to the Internet.

    (a)   Outline **one** technical challenge that *Harvest* may encounter in setting up its own website with the technologies that it has available. [2]

    (b)   Explain how the use of cloud-based services may help *Harvest* to overcome the limitations of the available technologies. [4]

Harvest is considering using a cloud-based service provided by a large IT company based in one of the more economically developed countries (MEDC).

    (c)   Describe **two** problems this may cause for an organizsation such as *Harvest*. [4]

Harvest wishes to improve its ranking in the major search engines.

    (d)   Discuss whether *Harvest* should use black hat search engine optimization techniques in order to improve its ranking. [5]

12. *Tapetum lucidum* refers to the layer of tissue in the eyes of many animals, such as cats and owls, which helps to improve their night vision by reflecting light. The eyes of these animals glow in the dark.

A student wishes to identify a particular animal which belongs to the above group.

    (a)   With the use of an example, explain how a multimedia search in the semantic web would be preferable to a text based search for this student. [4]

The web and social media are increasingly populated by pictures of pets and other animals. Most of these pictures are annotated with tags, such as those used on social networks.

    (b)   Identify **three** differences between an ontology and a folksonomy. [3]

    (c)   Discuss whether it is possible to integrate the expressivity of folksonomies with the authoritativeness of ontologies. [6]

Search engines use indexing for the storing of keywords that might be used in a search engine query.

    (d)   Explain the role that web crawlers perform in developing this index. [4]

Open source software, such as Linux, is often developed by using the collective intelligence of the developer community.

    (e)   Explain how collective intelligence can lead to the successful development of open source software. [3]

End of Option C

Option D — Object-oriented programming

An airport uses an object-oriented program to keep track of arrivals and departures of planes. There are many objects in this system and some are listed below.

Flight	Each `Flight` has many attributes including a unique ID code.
Arrival	In addition to a variable of the type `Flight`, each `Arrival` has a Scheduled Time of Arrival (STA) in 24-hour format, a runway number, a gate code, a delay time (in minutes) and an indicator whether the airplane has landed.
FlightManagement	The main class that manages the arrival and departure of all flights.

The code below outlines the `Arrival` class used in this program.

```
public class Flight
{ private String id;
  public String getId() {return this.id;}
  // ... more variables, accessor and mutator methods
}
public class Arrival
{ private Flight myFlight;
  private String sta;       // Scheduled Time of Arrival ("hh:mm")
  private int runway;
  private String gate;
  private int delay;
  private boolean landed;

  public Arrival(Flight myFlight, String sta)
  { this.myFlight = myFlight;
    this.sta = sta;
    this.runway = 0;
    this.gate = null;
    this.delay = 0;
    this.landed = false;
  }

  public void addDelay(int newDelay)
  { this.delay = newDelay;
  }

  public String getETA()
  { // calculates the Estimated Time of Arrival (ETA) of the flight
    // by adding the delay to the sta and returning the result as a
    // String ("hh:mm")
  }

  public int compareWith(String flightID)
  { if (myFlight.getID().equals(flightID)) { return 0; }
    else { return 1; }
  }

  public int compareWith(Arrival anotherArrival)
  { // missing code
  }

  // ... plus accessor and mutator methods
}
```

(Option D continues on the following page)

(Option D continued)

13. (a) (i)   Outline the general nature of an object.   [2]

   (ii) Describe **two** disadvantages of using Object Oriented Programming (OOP).   [4]

   (iii) Outline **one** advantage of using modularity in program development.   [2]

   (b) State the relationship between the `Flight` object and the `Arrival` object. [1]

   (c) Construct a UML diagram to represent the `Arrival` object.   [4]

The code below outlines part of the `FlightManagement` class used in this program.

For the purposes of this exam only arriving flights are being considered.

```
public class FlightManagement
{
  private Arrival[] inbound;        // array of inbound airplanes
  private int last = -1;            // index of last used entry

  public FlightManagement()
  { inbound = new Arrival[200];
  }

  public void add(Arrival newArrival)
  { // missing code that adds the newArrival to the array inbound
    // sorted by ETA, and updates last
  }

  private int search (String flightID)
  { // missing code that searches the array inbound and
    // returns the index of the Arrival object with flightID
  }

  public Arrival remove(String flightID)
  { Arrival result;
    int index = search(flightID);
    result = inbound[index];
    while (index < last)
    { inbound[index] = inbound[index + 1];
      index++;
    }
    last--;
    return result;
  }

  // ... many more methods
}
```

(Option D continues on the following page)

(Option D, question 13 continued)

The method `search` in the `FlightManagement` class searches the array `inbound` and returns the index of the `Arrival` object with the given `flightID`.

(d)   Construct the method `search()` implementing a linear search. Use the first `compareWith()` method in the `Arrival` object.

You may assume that an `Arrival` object with the given `flightID` exists in the array `inbound`.   [4]

(e)   (i)   Outline **one** advantage of using a binary search.   [2]

(ii) Outline **one** disadvantage of using a binary search.   [2]

14. (a)   Outline the use of data-hiding as a security feature in OOP.   [2]

(b)   (i)   The `Arrival` object has two methods named `compareWith`.

Outline why calling the `compareWith` method does not cause a conflict.   [2]

(ii) State the name of this OOP property.   [1]

The method `compareWith(Arrival anotherArrival)` compares the Estimated Time of Arrival (ETA) of the current `Arrival` object with the ETA of another `Arrival` object. The method should return a negative value if the current `Arrival` object's ETA is less than the ETA of `anotherArrival`.

Remember that the `String` class contains a method `compareTo()` as follows:
given two `String` variables A and B,
`A.compareTo(B)` < 0 when A is lexicographically smaller than B
`A.compareTo(B)` = 0 when A is equal to B
`A.compareTo(B)` > 0 when A is lexicographically bigger than B.

(c)   Construct the code for this method `compareWith()`.   [3]

(Option D continues on the following page)

(Option D continued)

15. (a) Define the term *method signature*. [2]

The array `inbound` in the `FlightManagement` class is sorted by Estimated Time of Arrival (ETA).

(b) Construct a method `showDelayed()` that outputs the IDs of all delayed flights in the array `inbound` that have not yet landed and that have an ETA before a given time `t`. The time `t` is passed as a `String` parameter. [4]

(c) Without using a sorting algorithm, construct the method `add(Arrival newArrival)` that inserts a `newArrival` in the sorted array `inbound`. You may assume that `newArrival` has been instantiated and that the array `inbound` is not full. [6]

When a flight is delayed, a method in `FlightManagement` is used to find and update the flight with the delay and to reorganize the array so that it remains sorted for the ETA.

(d) Describe how this method would be implemented using the methods provided. [4]

(Option D continues on the following page)

(Option D continued)

16. A different airport has two runways for arriving flights. Two objects of the `LinkedList` class, named `runway1` and `runway2`, are associated with these two runways.
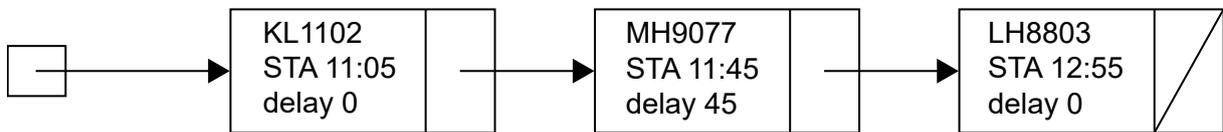
    (a)   Outline an advantage of using a library.   [2]

Arriving flights are added to a list depending on the runway they are assigned to.

Both linked lists contain `Arrival` objects, corresponding to arriving flights, and are sorted by ETA.

    (b)   Evaluate the use of dynamic linked lists to static arrays when managing arriving airplanes.   [4]

The following diagram represents `runway1`.

    (c)   Sketch the resulting linked list when an airplane with ID RO225, STA 12:05 and delay 0 is added to this list.   [2]

    (d)   (i)   Define the term *recursion*.   [1]

Consider the following method.

```
void process(LinkedList<Arrival> myList, int i)
{ if (i < myList.size())
  { process(myList, i + 1);
    System.out.println(myList.get(i).getETA());
  }
}
```

        (ii)   Trace the call `process(runway1,0)` given the diagram of `runway1` as drawn above. Copy and complete the following table.

i	id	output
0	KL1102	
…	…	…

    [4]

(Option D continues on the following page)

(Option D, question 16 continued)

In case of bad weather it is possible that one of the runways has to be closed and the two linked lists (`runway1` and `runway2`) need to be merged using a method `mergeLists()` into one new linked list `runway`.

(e)   Construct the method `public LinkedList<Arrival> mergeLists()` that merges the two sorted linked lists in to one sorted linked list `result` which needs to be returned. You may assume that `runway1` and `runway2` are accessible to the method and do not need to be passed. The original two lists are allowed to become `null` as a result of the merging.   [7]

In answering this question, you may use the following methods from the JETS `LinkedList` class, in addition to any method provided or developed in this exam paper.

```
addFirst(<object>)  // adds the object at the head of the list
addLast(<object>)   // adds the object at the tail of the list
getFirst()          // returns the head object in the list
getLast()           // returns the tail object in the list
removeFirst()       // removes and returns the head object
removeLast()        // removes and returns the tail object
size()              // returns the number of objects in the list
isEmpty()           // returns true if the list is empty
```

End of Option D