

22147012

**COMPUTER SCIENCE  
HIGHER LEVEL  
PAPER 2**

Monday 19 May 2014 (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]*.

<b>Option</b>	<b>Questions</b>
Option A – Databases	1–4
Option B – Modelling and simulation	5–8
Option C – Web science	9–13
Option D – Object-oriented programming	14–17

Blank page

**Option A — Databases**

1. A pharmacist can access the database of the Healthcare Regulatory Agency online to verify the existence and specification of authorized products.
  - (a) Outline the difference between an *information system* and a *database*. [3]
  - (b) Identify the advantages of having the data stored in a central database, instead of locally in each pharmacy. [2]
  - (c) Identify the database function that is available to the pharmacist. [1]
  - (d) Explain concurrency in the given scenario. [3]
  - (e) Discuss how the end-user can interact with the database. [6]
  
2. A company is in charge of providing the computing infrastructure to a business, and has developed a blueprint for a relational database management system (RDBMS).
  - (a) Define the following terms.
    - (i) *RDBMS* [1]
    - (ii) *Schema* [1]
  - (b) Identify **two** features that will be handled at the physical level. [2]
  - (c) Discuss the importance of data modelling in the design of a relational database. [6]

*(Option A continues on the following page)*

(Option A continued)

3. OnlyScience publishes several scientific journals. There must be one editor for each journal, but an editor can be responsible for more than one journal.

(a) Construct the entity-relationship diagram (ERD) for this scenario. [3]

The editor works with a pool of reviewers to review papers submitted for publication. Editors may be reviewers for other editors' journals. Their operations and activities are described in a database system that includes the following three tables, which begin as shown.

person	
<u>name</u>	phone
Ada	123
Boris	456
Tanja	789
Hugo	101
Anupam	126
...	...

activity		
<u>pers-id</u>	journal-id	role
Ada	Astrophysics	reviewer
Boris	Astrophysics	editor
Tanja	Chemistry	reviewer
Hugo	Nuclear Physics	editor
Anupam	Astrophysics	reviewer
...	...	...

journal		
<u>title</u>	description	number
Astrophysics	...	25
Biology	...	3
Chemistry	...	Special
...	...	...

(b) Define the following terms.

(i) Primary key [1]

(ii) Foreign key [1]

(c) Identify the answer to the following query.

```
SELECT a.journal-id
FROM activity a
WHERE a.pers-id="Boris"
```

[1]

(Option A continues on the following page)

(Option A, question 3 continued)

- (d) Construct, in any form, a query to find the phone numbers of all the people involved in the production of a journal with a “Special” number. [5]

The publisher only needs to see the activities of the editors, not of the reviewers.

- (e) Outline how a query can provide a view of a database. [3]

- (f) Discuss whether a *view* is physically stored in the database. [3]

Each submitted paper that will be reviewed is assigned a unique identifier. The following table gives an overview of the current papers to review, for which the editors have already found a reviewer.

submission				
<i>paper-id</i>	<i>reviewer</i>	<i>journal-title</i>	<i>number</i>	<i>editor</i>
34Wz678	Ada	Astrophysics	25	Boris
38Ty263	Ada	Astrophysics	25	Boris
96Ug900	Ada	Astrophysics	25	Boris
678HuT5	Anupam	Astrophysics	25	Boris
67Ji780	Hugo	Astrophysics	25	Boris
Z678	Boris	Nuclear Physics	3	Hugo
798YY	Giancarlo	Nuclear Physics	3	Hugo
...	...	...	...	...

- (g) Transform the relation “submission” into 3rd Normal Form (3NF), limited to the information explicitly visible. [3]

(Option A continues on the following page)

*(Option A continued)*

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

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

One of the applications on data warehousing is data mining.

(c) (i) Define the term *data mining*. [1]

(ii) Describe how cluster analysis is used in data mining. [3]

(d) (i) Explain why data in a data warehouse needs to be time dependent. [3]

(ii) Suggest how this time dependence can be achieved. [2]

(e) Discuss the implications on privacy of using data mining techniques in e-commerce. [5]

**End of Option A**

**Option B — Modelling and simulation**

5. A simplified version of a country’s debt in a three-month period,  $i$ , can be measured by:

$$\text{current debt } (CD_i) = \text{government spending } (GS) + \text{previous debt } (CD_{i-1}) - \text{income } (I)$$

The debt is recalculated every three months according to the latest figures and expressed as a percentage of the value of the country’s production (GDP).

For example, in June 2011, the figures for France were:

$$\text{debt (in €): } 1\,717\,256 \quad \text{GDP (in €): } 1\,996\,583 \quad \text{\%debt: } 85.8 \quad \left( \text{\%debt} = \left( \frac{\text{debt}}{\text{GDP}} \right) \times 100 \right)$$

A model is to be used to hold the figures for each three-month period over several years.

- (a) Outline the way in which a mathematical model of the debt and %debt for one country could be formulated to hold the figures for each three-month period over a number of years. You should include the necessary inputs for each period. [4]

Government income includes taxes and the sale of bonds (investments).

- (b) With reference to a spreadsheet, or any other simulation software with which you are familiar, describe how the model could be adapted to include these taxes and bond sales and used to simulate the trend in debt over the next few years. [6]

The trend in rising debt is important for financial and political reasons.

- (c) Discuss the advantages and limitations, to governments, of using this simulation to control the level of debt. [6]

*(Option B continues on the following page)*

*(Option B continued)*

6. A simple learning game for young children consists of random arithmetic questions being presented for the player to answer. The questions consist of two randomly selected numbers and an arithmetic operator, randomly selected from addition, subtraction, multiplication and division. The game is scored and each time the correct answer is given the score is incremented.

- (a) Identify the variables needed to implement the game as a program and describe how they would change each time a question is generated. [4]

It is decided to make the game more interesting by using 3D visualization so that numbers and arithmetic operators are represented as characters such as animated monsters and wizards. Correct answers lead the player successfully through a castle maze and wrong answers lead the player into dungeons with scary animals.

- (b) Outline the need for rendering in the creation of the visualization of the characters. [2]

- (c) Discuss **two** technical and **two** social implications of implementing an arithmetic game in this way. [8]

*(Option B continues on the following page)*



*(Option B continued)*

7. Many cities face problems of traffic congestion at busy times of the day. These problems can be reduced by various methods, for example:

- build more roads
- introduce more one-way systems
- replace traffic lights with roundabouts (traffic circles)
- change the timing of traffic lights at different times of the day.

Before introducing any of these methods it is necessary to research the current traffic patterns and to consider the effect that these methods would have on the situation. One way of doing this is to create a computer simulation.

- (a) Suggest ways in which data could be collected that describes the current traffic situation. [4]
- (b) Describe how current traffic and infrastructure could be simulated. [4]

The simulation can be run with changes in rules to reflect the four methods described above.

- (c) Suggest **two** criteria that could be used to evaluate the results of the simulation, to help decide which of the four methods should be implemented. [4]

Other factors could be incorporated into the model, such as improved bicycle lanes and public transport.

- (d) Outline the difficulties in incorporating these other factors into the simulation. [3]

8. (a) By using **three** different contrasts between human and machine languages, discuss why natural language processing is difficult to achieve. [6]
- (b) Outline **three** potential benefits of successful natural language processing. [6]

Neural networks are used extensively in character recognition.

- (c) With the aid of a diagram, outline the structure of a neural network which has three input nodes and one hidden layer. [2]
- (d) With reference to your diagram, explain how a neural network could be trained to recognize handwritten characters. [6]

**End of Option B**

**Option C — Web science**

9. The process of downloading web pages onto a personal computer requires the combination of applications and protocols.

(a) Outline the interaction between client and server if the HTTP protocol is used. [2]

Information sent over the internet can be intercepted.

(b) Describe **two** ways in which HTTPS provides for extra security. [4]

Consider the following piece of code found inside the HTML of a downloaded web page:

```
<script language="JavaScript" src="anyFile.js">
</script>
```

(c) With reference to the above piece of code,

(i) identify the consequence of its execution; [1]

(ii) identify **two** ways in which the executing of this code could improve the functionality of the web page. [2]

Many websites allow users to purchase goods online. These websites may also make offers to recognized customers.

(d) Discuss the role that server-side scripting can play in these purchases. [6]

*(Option C continues on the following page)*

*(Option C continued)*

**10.** The principal aim of search engines is to return a list of the pages most relevant to a search query. Web crawlers are used to find and return web pages for later analysis.

(a) Outline the steps by which a web crawler moves through the web. [3]

There are many pages belonging to the *deep web* that search engines will not find.

(b) (i) Outline **one** reason why web crawlers are not able to find these pages. [1]

(ii) Identify **one** way in which these pages are generated. [1]

(c) Outline how hubs and authorities are used by some search engines in determining the relative importance of a web page. [4]

Search engine users expect to find the pages most relevant to their query at the **top** of the returned list.

(d) Discuss the moral responsibilities of search engine designers and website developers in ensuring that this happens. [6]

*(Option C continues on the following page)*

*(Option C continued)*

**11.** The internet is continually evolving in structure, content and the way that it is being used. Two such changes are the increasing use of both ubiquitous computing and mobile computing.

- (a) Define, with the use of an example, the term *ubiquitous computing*. [2]
- (b) Identify **one** limitation of mobile computing. [1]

A major organization that analyses data from space telescopes has traditionally used powerful supercomputers for large-scale processing. However, it is now considering changing over to grid computing.

- (c) Examine **two** features of grid computing that would be considered before making the decision to change. [4]
- (d) With reference to data stored on the web, distinguish between copyright and intellectual property. [2]

Many organizations are concerned that action from Internet Service Providers (ISPs), major companies or certain countries might undermine the democratic nature of the web.

- (e) With specific reference to online retailers, discuss the concept of the democratic web and the consequences should this democracy be lost. [6]

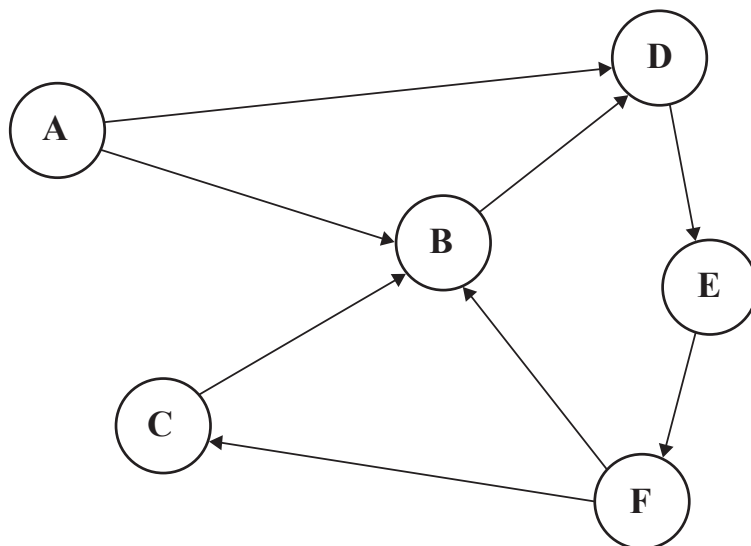
*(Option C continues on the following page)*

(Option C continued)

12. The following table contains a description of the web graph for the fictional university Stansfield.

Name	Type	Nodes	Edges	Description
web-Stans	Directed	524970	6345270	Web graph of Stansfield

- (a) Explain the information given by the following data items in this table.
  - (i) 524970 [1]
  - (ii) 6345270 [1]
- (b) Suggest an example of a sub-graph from this particular web graph described above. [1]
- (c) Consider the following web graph:



- (i) Describe the feature that makes this a **directed** web graph. [2]
- (ii) Outline why this web graph is **not** strongly connected. [2]
- (d) Searching the web for specific images is a challenging process for search engines. Contrast the use of both text based and image based queries in searching for a particular image. [4]

(Option C continues on the following page)

*(Option C continued)*

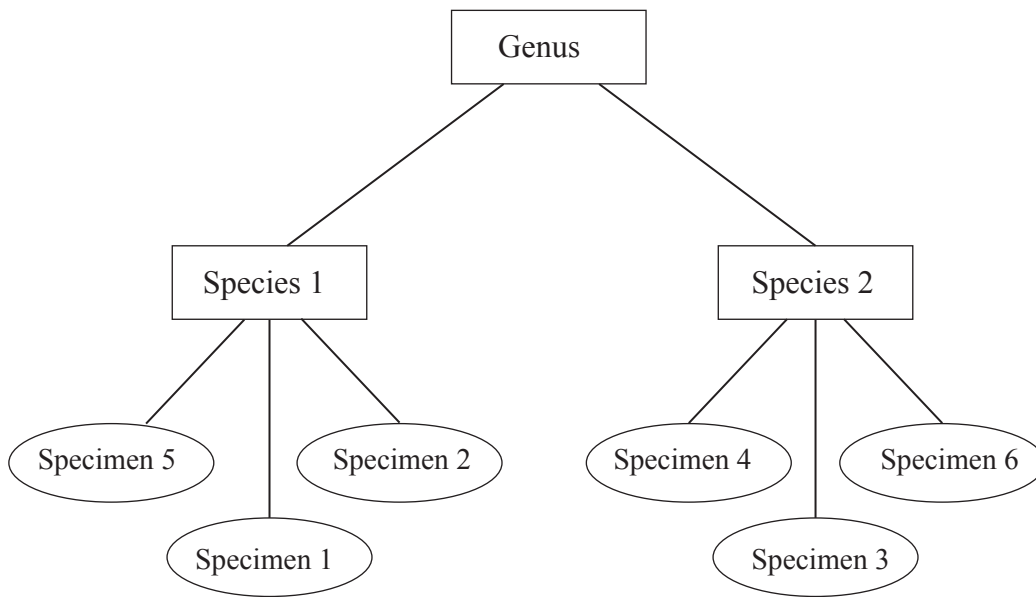
- 13.** (a) Outline the role of folksonomies in achieving a successful implementation of the semantic web. [3]
- (b) A mother and daughter live in a house equipped with ambient intelligence. The embedded computer systems react differently and automatically when a specific individual enters. By using examples of different reactions, discuss the technology needed for this to happen. [6]

**End of Option C**

**Option D — Object-oriented programming**

14. A large zoo has a collection of many individual animals of many different species. A computer program is being developed to keep track of all of the animals in the collection.

Because there are so many different kinds of species in the collection, and each species has some unique characteristics and some characteristics in common with other species, it was decided that the computer program would contain objects that correspond to different levels of the taxonomy used by biologists to classify all life forms. A genus is composed of a group of species that have similar common characteristics, as shown in the diagram.



A separate object, `Specimen`, is used to represent each individual animal in the zoo.

*(Option D continues on the following page)*

*(Option D, question 14 continued)*

The following code implements the Species and Specimen objects:

```

public class Species extends Genus
{
    private String speciesName;
    public Species( String s, String g )
    {
        super(g);
        setSpeciesName(s);
    }

    public void setSpeciesName(String s){ speciesName = s; }
    public String getSpeciesName(){ return speciesName; }
    public String toString()
    {
        return "Species: " + getGenusName() + " " + speciesName;
    }

    public boolean equals(Species s)
    {
        return speciesName.equals(s.getSpeciesName());
    }
}

public class Specimen
{
    private String name;
    private int cageNumber;
    private Species toa; // "Type Of Animal"
    public Specimen( String a, int c, Species s)
    {
        setName(a);
        setCage(c);
        setTOA(s);
    }

    public void setName(String a){ name = a; }
    public void setCage(int c){ cageNumber = c; }
    public void setTOA(Species s){ toa = s; }
    public String getName(){ return name; }
    public int getCage(){ return cageNumber; }
    public Species getTOA(){ return toa; }
    public String toString()
    {
        return name + " is a " + toa + " in cage " + cageNumber;
    }
}

```

*(Option D continues on the following page)*



*(Option D, question 14 continued)*

- (a) State the relationship between the `Genus` and `Species` objects. [1]
- (b) State the relationship between the `Species` and `Specimen` objects. [1]
- (c) Construct the unified modelling language (UML) diagram for the `Species` object. [4]
- (d) Outline **two** ways in which the programming team can benefit from the way the relationships between the three objects, `Specimen`, `Species` and `Genus`, have been represented in the code. [4]
- (e) The `Genus` class implements a `toString()` method that produces an output string that is different from the one produced by the `toString()` method in the `Species` class.

Consider the following code fragment:

```
Species human = new Species ( "homo", "sapiens" );  
System.out.println( human.toString() );
```

- (i) Outline why calling the `toString()` method in this code does not cause an error. [2]
- (ii) Identify the term for this property. [1]

*(Option D continues on the following page)*

*(Option D continued)*

15. (a) Define the term *encapsulation*. [1]
- (b) Outline **two** benefits provided by encapsulation. [4]
- (c) Identify an accessor method in the `Specimen` class. [1]
- (d) Identify an instance variable in the `Specimen` class. [1]
- (e) Construct code for the `Genus` object including a constructor, accessor methods and a `toString()` method. [3]

The `Specimen` object could have been designed as a sub-class of the `Species` object.

- (f) Outline **one** advantage and **one** disadvantage of having the `Specimen` object as a sub-class of the `Species` object. [4]
16. (a) Outline the changes that would be needed in order to add a description of each animal's individual markings to the program. [4]

An array is used to store the `Specimen` objects corresponding to the animals in the zoo.

- (b) Construct a method `countSpecimens( Specimen[] animals, Species s )` that will output the number of specimens of the given species in the zoo. [8]
- (c) Construct an algorithm in pseudocode for `listSpecies( Specimen[] animals )`, which will generate a list of the different species in the zoo. [6]

*(Option D continues on the following page)*

(Option D continued)

17. For some operations, it will be more convenient to use a linked list to hold the `Specimen` objects corresponding to the animals in the zoo. In answering this question, you should use the JETS `LinkedList` class and may use the following methods.

```
addHead()  
getHead()  
addTail()  
getTail()  
getNext()  
insert()  
delete()  
isEmpty()  
isFull()
```

- (a) Identify the features of an abstract data type (ADT). [2]

- (b) Construct a method, `makeList()`, that builds a linked list containing the `Specimen` objects from an input array.

```
LinkedList makeList( Specimen[] animals )  
{  
    // insert your code here  
}
```

[4]

- (c) Construct a method, `makeSpeciesList( LinkedList animals )`, that will return a linked list of `Species` objects, one for each animal specimen present in the zoo. Note that the list returned will contain duplicate `Species` objects if the zoo has more than one specimen of a given species. [6]

- (d) Construct a method, `makeSpeciesListUnique( LinkedList allSpecies )`, that will take a linked list of `Species` objects as described in (c) and produce a new linked list that contains only one `Species` object for each species in the zoo. [8]

**End of Option D**

---