International Baccalaureate®
Baccalauréat International
Bachillerato Internacional

**COMPUTER SCIENCE**
**HIGHER LEVEL**
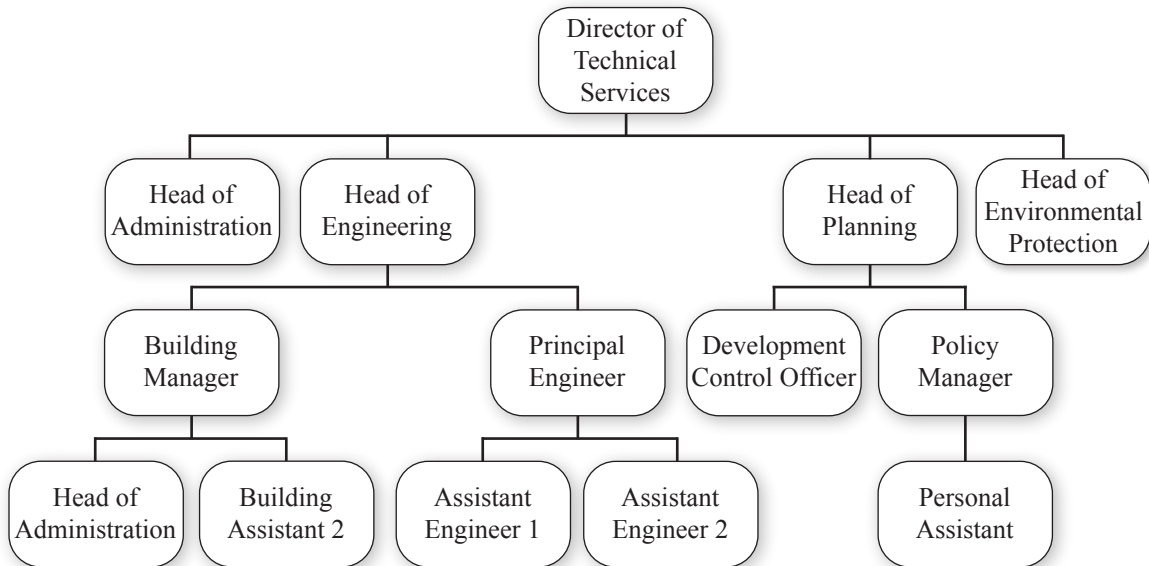**PAPER 2**

SPECIMEN PAPER

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.

**Option A — Databases**

**A1.** A company has a division called Technical Services. This division is organized into departments. The following chart shows how the various staff are organized.



The Human Resources Department maintains a flat file database to store information about all the staff members for salary, holiday and promotion purposes. A sample of the stored information is shown below.

| Staff Number | Surname | Forename | Job Title | Line Manager | Salary |
|---|---|---|---|---|---|
| 948 | Kent | McKenzie | Head of Planning | Adrienne Mathis | 90875 |
| 156 | Barker | Rae | Head of Administration | Adrienne Mathis | 30465 |
| 815 | Riddle | Kevyn | Development Control Officer | Kent McKenzie | 34768 |
| 580 | Figueroa | Rina | Policy Manager | Kent McKenzie | 45078 |
| 871 | Mathis | Adrienne | Director of Technical Services | Russell Z. Harrell | 45800 |
| 457 | Neal | Paul | Head of Environmental Protection | Adrienne Mathis | 80670 |
| 297 | Oliver | Ralph | Principal Engineer | Daria Gilmore | 50796 |
| 51 | Gilmore | Daria | Head of Engineering | Adrienne Mathis | 90655 |

*(Question A1 continued)*

The Human Resources Department is experiencing many errors when compiling reports of personnel and has been advised that it should change to a relational database.

(a)  (i)   Define the term *database*.   *[1]*

    (ii)   Describe the difference between a *flat file database* and a *relational database*.   *[2]*

(b)  (i)   Identify **two** problems caused by data redundancy.   *[2]*

    (ii)   Outline using examples how data redundancy in this database could lead to errors.   *[2]*

(c)   The departments usually have many ongoing projects that occupy their team members. The Human Resources Department needs to produce lists showing who is involved in current projects.

    The following information has been provided:

    **Project**: Project_Ref, Project_Name, Budget
    **Department**: Dept_Ref, Dept_Name
    **Staff**: Staff_Number, Forename, Surname, Job_Title, Salary
    The projects are owned by a department.
    Each member of staff is a member of a department.

    (i)   Construct an Entity Relationship Diagram to show how a relational database could be designed to make this possible with the least chance of producing errors caused by data redundancy.   *[4]*

    (ii)   Identify the steps to create a query that could be constructed to produce a list of staff members whose salary is greater than $50 000 who are involved in a project called "Relocation_2012".   *[4]*

**Turn over**

**A2.** A bank wants to transfer money from one account to another. Both accounts are held in the bank's database management system (DBMS). The account DBMS carries out a series of separate operations in order to achieve this transaction. These include removing of money from one account and adding it to the other. An error may occur if this process is interrupted, for example due to a power failure. This could compromise the integrity of the account database.

    (a)    (i)    Identify **two** characteristics of a *transaction* in terms of a database. *[2]*

            (ii)    Distinguish between a database *schema* and a database *state*. *[2]*

            (iii)    Explain why the interruption described above could compromise the integrity of the account database. *[2]*

    (b)    Identify the steps that should be taken by the software in order to maintain the integrity of the account database. *[4]*

    (c)    The bank account data is shared between various applications, such as the bank's internal accounting systems and the online transaction system. Explain how problems of concurrency can arise and how they can be avoided. *[6]*

**A3.** (a)    (i)    Identify **two** functions of a *database management system* (DBMS). *[2]*

            (ii)    List **two** tools usually provided with a DBMS. *[2]*

    (b)    Database management systems allow data sharing and multiple views. Explain why these features are important to an organization that uses databases. *[4]*

    (c)    Explain why it is important for data to be independent of the application software that manipulates it. *[6]*

**A4.** Meubles de France is a large company that manufactures and sells furniture. It maintains a large factory and offices in Provence and over a hundred retail outlets. It employs hundreds of workers. It holds huge amounts of data on separate computer systems to handle:

- customer orders
- employees
- sales data
- production data
- finance
- budgeting.

All these systems have been acquired from different software suppliers. The management of Meubles de France has been advised that a data warehouse would be of great benefit to the business.

(a)  (i)   Identify **two** characteristics of a data warehouse.  *[2]*

    (ii)  Identify **four** features of Meubles de France's business that indicate the need for a data warehouse.  *[4]*

    (iii) Identify **four** transformations that may be necessary in order to produce usable data in a data warehouse.  *[4]*

Meubles de France uses data mining to uncover patterns of consumer spending so they can optimise their marketing for future products and services.

(b)  Contrast the use of association and cluster analysis for this purpose.  *[4]*

Many organizations such as Meubles de France use data mining to build a complex profile of its customers.

(c)  Explain why civil liberty groups may be concerned about Meubles de France having such detailed information.  *[6]*

**Turn over**

**Option B — Modelling and Simulation**

**B1.** Mathematical models are used to provide information to building designers.

(a) Identify **two** characteristics of a mathematical model. *[2]*

Any new building must be constructed to resist the force of a possible earthquake. The force (V) that the building must resist in order to prevent collapse varies according to the following variables:

| Variable | |
|----------|---|
| Z | The Earthquake zone (1–3) in which the building is constructed |
| I | Importance of the building according to its intended use (*e.g.* school, hospital, *etc.*) |
| C | Structure index based on the dimensions, the weight and the shape of the building |
| R | Flexibility of the building |

A simplified version of V can be calculated by the formula:

$$V = (Z * I * C) / R,$$

where Z and I are given and the values of C and R are obtained from published tables.

A building designer needs to be able to calculate the value of V for any one of a large number of buildings that is being designed. Z is fixed for all buildings; I, C and R can vary from building to building.

(b) Outline why the use of a spreadsheet is appropriate for this modelling task. *[2]*

(c) Construct a diagram to show the way in which data and calculations would be implemented in a spreadsheet. *[4]*

(d) Outline how data from the look up tables could be input into the spreadsheet. *[4]*

Within a city in an earthquake zone there are many different designs of buildings. It is decided to test all buildings for their resistance to a possible earthquake in that zone.

(e) Describe the data collection and data input needed to use the spreadsheet to test all buildings. *[6]*

**B2.** Despite many attempts, the simulation of an earthquake by computer has proved inaccurate and planned buildings are tested by making a physical model and applying the relevant forces to see if it can withstand an earthquake in the given zone.

(a)　Explain the relationship between a *simulation* and *mathematical modelling*.　　*[2]*

(b)　With reference to a simulation with which you are familiar, explain the importance of accurate rules and data in a simulation.　　*[4]*

Simulation has proved successful in predicting weather patterns.

(c)　Describe **one** advantage and **one** social consequence of using simulations for weather forecasting.　　*[4]*

**B3.** Cyclone Yasi hit Northern Australia in February 2011. There were many after effects such as flooding and landslides.
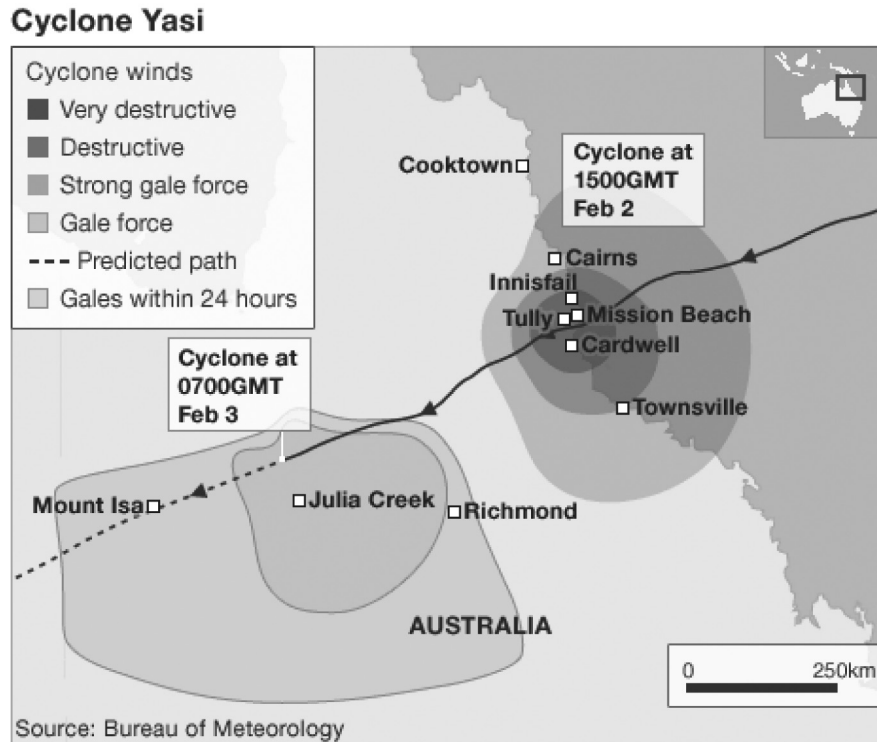


[Source: Australian Government Bureau of Meteorology (accessed 07 February 2011)]

To ensure that the emergency services could be located where they will be most needed a 2D visual model of likely danger areas, shown above, was created and distributed to surrounding areas in real time.

(a) Outline the ways in which visualization can be used to display the possible effects of the cyclone in real time. *[4]*

(b) Explain the technical difficulties that could arise in data collection and processing when attempting to predict in real time. *[4]*

(c) Explain the advantages of using visualization in this case. *[6]*

(d) Compare the use of 2D visualization with 3D visualization in this situation. *[3]*

**B4.** "Rescue robots capable of understanding the changing and unpredictable environment of disaster scenarios may one day be deployed to search for survivors in the aftermath of earthquakes.

This is the vision of inventors Erwin Prassler and Ivan Bratko, who have developed a software algorithm that takes data from a robot's sensors as it moves through an area to create models and predict how objects in the vicinity will change their position relative to its movements.

Using the same algorithm, Bratko said that the robot can learn physical concepts such as whether an object is moveable and where it can be moved to. He added that it also gives the robot the ability to learn 'abstract concepts' such as the structural stability of an object."

[Source: http://www.theengineer.co.uk, 1 September 2010]

In this question you can assume that the robot is equipped with distance and angle sensors.

Robots involved in rescue are placed in an unknown environment and have to model the environment. Part of this process involves genetic algorithms.

(a) Outline, using examples, the difference between *supervised learning* and *unsupervised learning*. *[4]*

(b) Explain the way in which *genetic algorithms* help in the learning process. *[6]*

(c) Describe the way in which the robot could model the situation in which it finds itself. *[4]*

Once the robot has found a victim, it needs to communicate this information back to emergency services.

(d) Suggest ways in which the robot and human rescue workers could communicate to make a successful rescue of a person. *[6]*

**Turn over**

**Option C — Web Science**

**C1.** A publishing company, ABC Publications, based in London has a large IT department. This department is responsible for:

- providing IT services to the company
- maintaining the company's web site
- creating and maintaining web based learning resources that are sold to schools and colleges.

The company is finding it difficult to recruit and retain sufficient high quality IT staff to keep these functions operating at an optimal level. It is investigating transferring at least part of its IT operations to a cloud computing solution. At the moment it has not been decided how much of this should be implemented by a private cloud and how much by a public cloud.

(a)    Define the term *private cloud*.    *[1]*

(b)    Distinguish between a cloud computing model and a conventional client server model in providing computing services.    *[4]*

(c)    Explain why ABC Publications might benefit from changing part of its IT provision to a cloud based model.    *[4]*

(d)    Comment on the privacy **and** security issues relating to ABC's use of cloud computing.    *[4]*

**C2.** A theatre box office maintains a web site that can display what productions are coming up and which seats are available for a particular production. It can then take orders online. The interface consists of dynamic web pages such as the one below, in which the underlying HTML interacts with client-side and server-side scripts.



Part of the source code for this page is:

```
<script type="text/javascript" src="http://assets.ophse.org/_inc/popdt/init_live.js"></script>
<script type="text/javascript" src="http://assets.ophse.org/_inc/popdt/src/flash.js"></script>
```

(a)     Identify **one** characteristic of Hypertext Markup Language (HTML). *[1]*

(b)     Identify the steps that the server would carry out so that the information in the events' calendar can be displayed on the client's computer. *[4]*

(c)     (i)     Identify **two** ways that a client-side script may be made available to a web browser. *[2]*

(ii)    Describe **one** reason why a client-side script may be used in preference to a server-side script. *[2]*

*(This question continues on the following page)*

**Turn over**

*(Question C2 continued)*

(d) The organizers of the theatre want to ensure their web pages appear higher up the ranking of search engines.

   Suggest whether the use of meta-tags can help achieve this aim. *[4]*

(e) It is common for dynamic web pages to make use of a mixture of client-side and server-side scripting. Explain how the interaction of HTML, client-side and server-side scripting have allowed the production of a web page such as the one shown here. *[6]*

**C3.** As the web has developed, data formats, communication protocols and standards such as XML or SQL have proven crucial to progress. Two fundamental concerns that have been central to this development are the issues of interoperability and that of open standards.

(a) (i) Identify **one** characteristic of XML. *[1]*

   (ii) Define the term *protocol*. *[1]*

(b) Describe, with the use of examples, how the use of open standards allows interoperability to occur. *[3]*

Music is distributed across the web in a variety of different ways such as peer-2-peer (P2P) networks.

(c) Discuss **two** factors that would affect the decision to use either lossless or lossy compression when transferring files across the Internet. *[6]*

(d) Explain **one** advantage of the use of a peer-2-peer (P2P) network for obtaining and downloading music and movie files. *[2]*

**C4.** The World Wide Web (web) can be regarded as a directed graph. This allows search engines to make use of algorithms based on graph theory.

(a)  Identify how the web may be represented as a directed graph.    *[1]*

The web can be represented as having a bowtie structure as indicated in the diagram below.
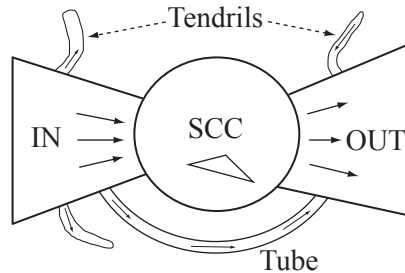


[Source: http://nlp.stanford.edu/IR-book/html/htmledition/the-web-graph-1.html (accessed 13 April 2011)]

(b)  (i)  Define the term *Strongly Connected Core* (SCC).    *[1]*

    (ii)  Outline the characteristics of web sites that are located in the IN portion of the bowtie.    *[2]*

With the growth of the web, web site developers have realized that there are possible concerns about the ability of being able to link to all web pages as well as ensuring a page is highly ranked by search engines.

(c)  Describe how power laws suggest it will be possible to link from one web page to any other web page despite the fact the web is growing so rapidly.    *[2]*

(d)  Explain how the relative importance of a web page can be determined in search engines.    *[4]*

The development of the web has changed the way that users interact with the web and with each other.

(e)  With reference to specific examples, distinguish between an ontology and a folksonomy.    *[2]*

The growth of the web has enabled the development of new ways to solve problems. Collective intelligence is one such approach.

(f)  Identify **two** characteristics of collective intelligence.    *[2]*

(g)  With the increasing world population and the effects of globalization, the world's population is facing new and complex problems such as the recent banking crisis. To what extent is it likely that collective intelligence could help to address these problems?    *[6]*

**Turn over**

**Option D — Object-oriented programming**

A bus company operates in a local city. The bus company operates along fixed routes where there are marked stops and sometimes bus shelters for people to wait in out of the weather. People (passengers) pay the driver a specified fare for travel when they enter a bus.

There are many objects in this company, here are some of them:

| Object | Description |
|---|---|
| *Bus* | A physical vehicle that carries *passengers* on a specific *route* and has a *driver*. |
| *Passenger* | A person that travels on a *bus*. |
| *Route* | A series of roads/streets the *bus* travels over from its start to its destination. |
| *Bus Stop* | A **named** place on a *route* where people wait for a *bus*. May be a simple marker or may have a shelter and seats. |
| *Driver* | A person qualified to drive a *bus* and trained to drive it over a given *route*. |

These two objects have already been defined for the bus company:

```
BusRoute
────────────────────────
Integer: route
String: start
────────────────────────
setRoute(Integer: route)
setStart(String: start)
Integer getRoute()
String getStart()
String toString()
```

```
Bus
────────────────────────
Integer: id
String: driver
BusRoute: busRoute
────────────────────────
setId(Integer: id)
setDriver(String: driver)
setBusRoute(BusRoute: route)
Integer getId()
String getDriver()
BusRoute getBusRoute()
String toString()
```

The `toString()` method returns a String implementation of an object.

These are implemented in code as follows:

```java
public class BusRoute
{
  private int route;
  private String start;
  public BusRoute(int r, String s)
  {
    setRoute(r);
    setStart(s);
  }
  public void setRoute(int r){ route = r; }
  public void setStart(String s){ start = s; }
  public int getRoute(){ return route; }
  public String getStart(){ return start; }
  public String toString()
  {
    return "Route: " + route + " start: " + start;
  }
}


public class Bus
{
  private int id;
  private String driver;
  private BusRoute busRoute;
  public Bus(int i, String d, BusRoute r)
  {
    setId(i);
    setDriver(d);
    setBusRoute(r);
  }
  public void setId(int i){ id = i; }
  public void setDriver(String d){ driver = d; }
  public void setBusRoute(BusRoute r){ busRoute = r; }
  public int getId(){ return id; }
  public String getDriver(){ return driver; }
  public BusRoute getBusRoute(){ return busRoute; }
  public String toString()
  {
    return "Bus id:" + id + " - " + driver + ": " + busRoute.toString();
  }
}
```

**Turn over**

**D1.** (a)  Explain the term *parameter variable*, using an example from the code.          *[2]*

   (b)  Describe **one** additional field that might have been included in the `BusRoute` object/class. Include data types and sample data.          *[2]*

   (c)  Identify the output produced by the following code fragment.

```
Bus bus = new Bus(1001, "N Prakesh", new BusRoute(431, "Klang"));
System.out.println(bus.toString());
```
          *[2]*

Consider the code fragment below.

```
private static final int MAX_BUSES = 12;
private Bus[] buses = new Bus[MAX_BUSES];
buses[0] = new Bus(1001, "N Prakesh", new BusRoute(431, "Klang"));
buses[1] = new Bus(1010, "J Carey", new BusRoute(342, "Tanglin"));
buses[2] = new Bus(1014, "H Lee", new BusRoute(411, "Queenstown"));
buses[3] = new Bus(1015, "K Peters", new BusRoute(319, "Jamaica Street"));

showBusDrivers(buses, 1010);
```

   (d)  Construct the method `showBusDrivers(Bus[] b, int n)` which lists the drivers for all buses with a route number less than or equal to the parameter variable (`n`).          *[6]*

The company wishes to keep track of its drivers in more detail, including first and last name and employee number – this is a 4-digit whole number.

   (e)  Construct a suitable diagram for this `Driver` object.          *[3]*

**D2.** In relation to the Bus example:

(a) Outline how encapsulation is used. *[2]*

(b) Outline a disadvantage of using Object Oriented Design. *[2]*

(c) Explain how a programming team could benefit from an Object Oriented Design approach. *[4]*

Recall that a *Bus Stop* is one of many **named** places on a *route* where *buses* stop to pick up or drop off *passengers*. It may or may not have a **shelter** to protect *passengers* from the weather. The **distance** in km from the start point of the *Bus Route* is important information for planning.

(d) Design the Bus Stop Object using a simple object diagram. *[3]*

(e) Suggest how Bus Stop information for a given Bus Route instance could be stored, giving both sample data and sample code fragments to show how it could be implemented. *[4]*

**D3.** The company grows, offers more routes of different types and decides to use three different types of bus:

- A bus that operates on busy city routes – the Urban Bus – has only a driver.
- A smaller bus that operates on longer country routes – it carries an additional person to collect the fares.
- A van that can be used for heavy equipment deliveries rather than passengers – the Delivery Van – it carries a co-driver and a helper.

These vehicles will have some things in common, such as a *driver*, and other elements that are different, for example both Urban and Rural buses will operate on a fixed *route* whereas the Delivery Van will take equipment to specified *destinations* (such as factories or other businesses).

(a) Construct diagrams to show how you would re-design the `Bus` class to implement inheritance. *[8]*

(b) Explain the advantage of inheritance for this situation. *[4]*

A method is required in the subclasses that returns the number of employees per vehicle.

(c) Outline how polymorphism might apply in this design. *[3]*

**Turn over**

**D4.** The bus company decides to run a simulation over a particular route to see what happens when several buses are started on the route a set time apart. A queue will be used to hold the individual `Bus` instances.

(a)  Identify **three** features of a queue that make it suitable for this purpose. *[3]*

(b)  Construct a diagram of the queue after the following code has been executed.

```
public class BusSim
{
  private LinkedList<Bus> busQueue;

  public static void main(String[] args){ new BusSim(); }
  public BusSim()
  {
    // Create new LinkedList for Bus instances
    busQueue = new LinkedList<Bus>();
    BusRoute route = new BusRoute(903, "Nerang Creek Road");
    Bus bus1 = new Bus(2011, "C Humbley", route);
    Bus bus2 = new Bus(3943, "M Hillier", route);
    Bus bus3 = new Bus(4923, "J Inglis", route);
    busQueue.addFirst(bus1);
    busQueue.addFirst(bus2);
    busQueue.addFirst(bus3);
  }
}
```
*[3]*

Recall that the method of the `LinkedList` class `remove(int index)` removes the element at the specified position in the list while the method `size()` returns the number of elements in the list.

(c)  Construct the method `removeBus(int pos)` which attempts to remove a bus from the queue at the specified position, and returns true if successful and false if it fails. *[4]*

A large company might have several hundred buses running. Each one has a unique id stored with the `Bus` instance.

(d)  Explain how a binary tree could be used to store these ids such that they can be quickly retrieved (if they exist) by a search. *[3]*

The tree stores the ids 2045, 3474, 5877, 1099, 9644.

(e)  Draw a diagram of an ordered binary tree containing these keys assuming they were inserted in the order given. *[5]*

A binary tree node may be inserted iteratively or recursively.

(f)  Identify **two** disadvantages of the recursive algorithm. *[2]*