

BCS THE CHARTERED INSTITUTE FOR IT
BCS HIGHER EDUCATION QUALIFICATIONS
BCS Level 6 Postgraduate Diploma in IT

ADVANCED DATABASE MANAGEMENT SYSTEMS

Wednesday 2nd April 2014 - Afternoon

Answer **any** THREE questions out of FIVE. All questions carry equal marks
Time: THREE hours

Answer any Section A questions you attempt in Answer Book A
Answer any Section B questions you attempt in Answer Book B

The marks given in brackets are **indicative** of the weight given to each part of the question.

Calculators are NOT allowed in this examination
--

Section A

Answer Section A questions in Answer Book A

A1

- (a) Using a simple example of your own choosing, explain and discuss the following data warehouse design issues (and how they relate/differ to the ERD modelling process for an OLTP database):

- Star schemas
- Snowflake schemas
- Dimensional Modelling

You should particularly address the roles of primary and foreign keys, normalized & de-normalized data. Good diagrams are essential.

(10 Marks)

- (b) Discuss and explain the following data warehouse features and techniques:

- Summary Management
- Analytical Functions
- Indexing & Optimization Techniques

(10 Marks)

- (c) Describe the different *categories* (not specific products) of *data extraction* tools that an end-user may use to interact with a data warehouse, briefly highlighting the primary purpose of each.

(5 Marks)

A2

(a) Using your own simple examples and/or diagrams, describe the potential problems that may occur in a *multi-user* database if *concurrency control* techniques are not fully implemented. You should comment specifically on:

- Transaction scheduling techniques
- How data may be damaged, lost or misread
- The use of time-line diagrams to model two or more transactions accessing the same database.

(10 Marks)

(b) For each of the following transaction control terms, write a *single sentence* (no need for extended responses, examples or diagrams) explaining the key concept.

- Schedule
- Cascaded rollback
- Optimistic locking
- Pessimistic locking
- Checkpoint

Each is worth one mark.

(5 Marks)

(c) In your own words, describe what is meant by the following *transaction-processing* terms:

- *Two-phase locking* (and the function of each stage)
- *Serializability* (and the role of serial and interleaved schedules)

You should supply any suitable examples and/or diagrams that you deem appropriate to support your answer.

(10 Marks)

A3

Consider the following database that contains information about directors and the films that they have directed:

Film (filmNbr, title, year)

Director (directID, name)

Directs (directID*, filmNbr*)

(a) Consider the following query:

```
SELECT Film.title
FROM Film, Director, Directs
WHERE Film.filmNbr = Directs.filmNbr
AND Director.directID = Directs.directID
AND Director.name = 'Lucas';
```

Suppose this query is run by executing the following sequence of steps:

1. R1 = Join of Director and Directs
2. R2 = Join of Film and R1
3. R3 = Selection (name = 'Lucas') from R2
4. R4 = Projection (title) from R3

(i) What problem arises if the query is executed based on the sequence above.

(4 Marks)

(ii) Suggest a new sequence that will make the query more efficient. Hint: You may need to introduce extra steps and not just re-arrange the existing steps.

(12 Marks)

(b) Suppose there is an index on the column "title" of the "Film" table above. Explain how this index could be used when executing each of the following queries:

(i)

```
SELECT *
FROM Film
WHERE title = 'The God Father';
```

(3 Marks)

(ii)

```
SELECT *
FROM Film
ORDER BY title;
```

(3 Marks)

(iii)

```
SELECT COUNT(title)
FROM Film;
```

(3 Marks)

Section B

Answer Section B questions in Answer Book B

B4

A GIS is a database application that processes geo-spatial data, in other words information about objects that exhibit geographical features on a map. There are three basic types of geo-spatial objects the simplest is based on a single point or location having a geographical reference such as latitude and longitude. Consider the following scenario

A University has adopted a personal identity card (PID) system to improve security and to restrict access to different groups of people to certain locations such as buildings and rooms within buildings. Permission to enter a building does not necessarily allow access to every room in that building. To enter a building or room a person swipes their PID card through a card reader outside the door of the building or room. Staff and students can only enter designated buildings or rooms once their access rights are confirmed. When a card is swiped the PID reader records who has entered a building or room. Using the PID system it is possible to calculate the routes taken, when an individual enters or leaves a particular zone building or room.

With specific reference to the scenario above:-

- a) Give an example of **two further** basic types of geo-spatial data made up of points that would be used to model the PID system.

(5 Marks)

- b) **EITHER**

Drawing upon examples from the scenario above, discuss the limitations of the relational model in modelling the relationship between geo-spatial objects

OR

Describe the characteristics of modern geo-spatial database applications that have allowed them to become very popular and easily accessible over the WWW and describe how these could apply to the PID system.

(20 marks)

B5

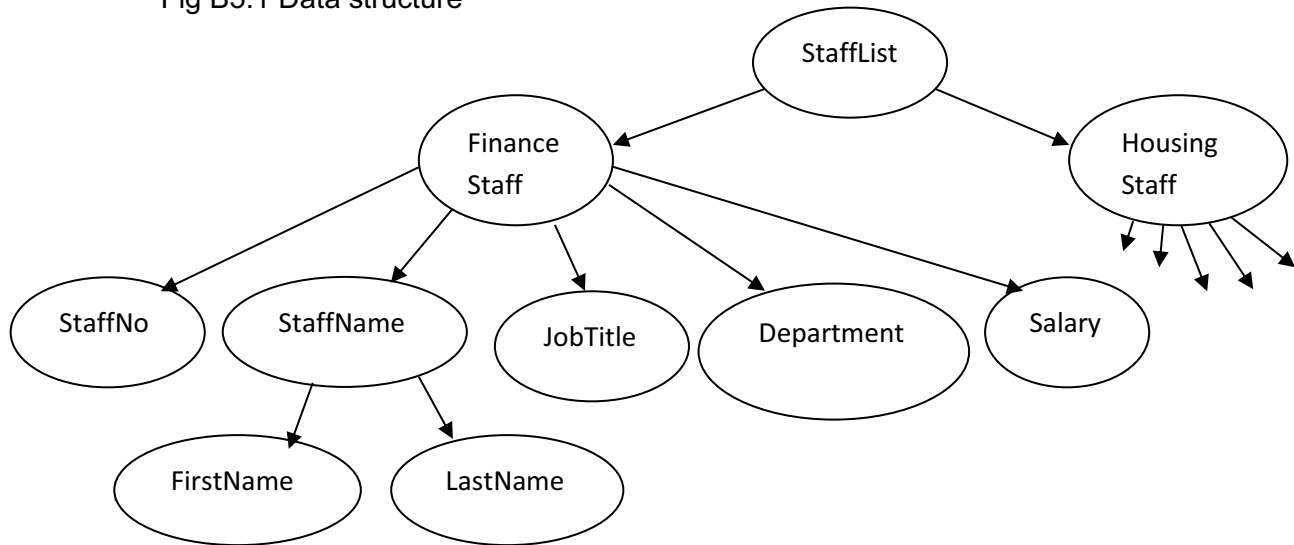
- a) Represent the contents of the relation StaffList as an XML document using the data structure given in Fig B5.1 and a namespace called *StaffList* with a url = <http://www.Idealhomestaff.co.uk/branch>.

(8 Marks)

StaffList

StaffNo	FirstName	LastName	Department	Salary	JobTitle
123	Robert	Rogers	Housing	34300	Manager
126	Duncan	Boult	Finance	22000	Accountant

Fig B5.1 Data structure



- b) XQuery is a language used to query XML data.

Describe with the aid of examples using the XML document you produced in fig B5.1 how each of the following types of XQuery expressions are formulated

- path expressions,
- conditional expressions,

(10 Marks)

- c) Explain the type of expression and the function of the following XQuery expression

```

{
  LET $doc := document("stafflist.xml")
  FOR $t IN distinct($doc/staff/salary)
  LET $p := $doc/staff[jobtitle = $t]/salary
  RETURN
  <avgsalary jobtitle={ $t/text() }>
  {
    avge($p)
  }
  </avgsalary>
}

```

(7 Marks)

