

BCS THE CHARTERED INSTITUTE FOR IT

BCS HIGHER EDUCATION QUALIFICATIONS
BCS Level 6 Professional Graduate Diploma in IT

ADVANCED DATABASE MANAGEMENT SYSTEMS

Friday 27th March 2015 - 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

In recent years there has been a trend to represent traditional relational data in a semi-structured data format such as XML.

- a) List and explain three of the main motivations for this trend over recent years. **(6 marks)**
- b) State an application and explain why the stated application warrants the need to store and process semi-structured data. **(6 marks)**
- c) Explain the support required for storing semi-structured data in a relational database. Hint: use a simple example such as a file that contains product details. **(7 marks)**
- d) Discuss the performance implications for retrieving a particular XML element from a XML file. **(6 marks)**

[Turn Over]

A2

Because of the increasing storage capacity of memory, a typical database application can now cache most of the application's data requirements in internal memory.

- a) Explain the concept of data persistence and explain the impact on data persistence given the above statement. **(8 marks)**
- b) Consider the following scenario that describes the processing of examination results at a college on a database that holds information on student assessment.

Students are assessed on a particular course by taking 4 exams. Each exam is the only assessment for a module on which they have enrolled. The students from different courses share the same module. Exam marks for a particular student are entered in sequence. A report is generated showing the end-of year assessment results with the following column headers :-

Student_ID, Student_name, Date_Assessed, Average_mark, Grade

- i) Using this information derive a simple CLASS (Object Oriented) model using a defined notation (for example UML) **(5 marks)**
- ii) A database trigger could be used to implement the following business rule.

Business Rule:-

If the mark entered is less than 30% then an overall grade of FAIL is recorded. When all 4 marks are entered then the average mark is calculated and a grade of PASS or FAIL recorded. For a PASS the average mark must be 40% or more with no single mark less than 30% otherwise a FAIL is recorded.

Explain with the aid of sample data and pseudo-code how this could be achieved and discuss the advantages and disadvantages of using triggers in this way.

(12 marks)

A3

- (a) Consider the following three linked tables that contain information about employees and the projects they work on:

```
employees (empID, name, salary)
project (projNbr, title, budget)
workload (empID*, projNbr*, duration)
```

Consider the following query:

```
SELECT P.title, E.name
FROM employees E, project P, workload W
WHERE E.empID = W.empID
AND P.projNbr = W.projNbr
AND E.salary > 15000
AND W.duration < 20;
```

- (i) Draw an initial relational algebra tree for the above query. **(4 marks)**
- (ii) Apply a series of transformations to the tree obtained in part (i) to make the query more efficient. Discuss each step and state the heuristic used. **(12 marks)**

(b) Suppose you have the following table:

Employees (empID, lastName, firstName, salary)

Suppose that the most frequently executed queries on this table are as follows (the “?” indicates any given value):

- `SELECT * FROM Employees WHERE firstName = ? AND lastName = ?`
- `SELECT * FROM Employees WHERE lastName = ?`
- `SELECT empID FROM Employees ORDER BY salary`

Suppose there are four indexes as follows:

- Index1 (empID)
 - Index2 (firstName)
 - Index3 (lastName)
 - Index4 (salary)
- (i) There is a need to improve the performance of all three of the above queries. Select **two** indexes from the above list that if built and available will result in an improved performance being obtained for all the above queries. Explain your answer. **(6 marks)**
- (ii) Describe three disadvantages of using indexes. **(3 marks)**

[Turn Over]

Section B

Answer Section B questions in Answer Book B

B4

- (a) *Using your own suitable examples and diagrams*, explain in your own words what the following database concepts mean:
- (i) Database Transaction **(3 Marks)**
 - (ii) ACID Properties **(3 Marks)**
 - (iii) Isolation Level **(3 Marks)**
- (b) *Using your own examples and suitable diagrams*, discuss the following transaction-processing concepts:
- (i) COMMIT **(3 Marks)**
 - (ii) TWO-PHASE COMMIT **(3 Marks)**
 - (iii) ROLLBACK **(3 Marks)**
 - (iv) SAVEPOINT **(3 Marks)**
- (c) For each of the following transaction control terms, write a *single sentence* (no need for extended responses, examples or diagrams) explaining the key concept.
- (i) Cascaded rollback **(1 Mark)**
 - (ii) Optimistic locking **(1 Mark)**
 - (iii) Pessimistic locking **(1 Mark)**
 - (iv) Checkpoint **(1 Mark)**

B5

- (a) Describe the key characteristics of a *data warehouse* and how it differs in content, structure and function from an *on-line transaction processing (OLTP)* database. You should support your discussion with suitable diagrams and examples. **(10 Marks)**
- (b) For each of the following items, explain what the term means, the underlying concepts involved, any associated benefits or limitations, typical applications and features along with any additional technical or implementation points you think appropriate to mention. You should support your discussion with suitable diagrams and/or examples.
- (i) OLAP (*hint: think different implementations of OLAP, SQL and OLAP, aggregation etc.*) **(5 Marks)**
 - (ii) Multi-Dimensional Data (*hint: think what each dimension could represent, roll-up, pivoting*) **(5 Marks)**
 - (iii) Data Mining (*hint: patterns and prediction, techniques to identify these, data preparation, tools*) **(5 Marks)**