UNIVERSITY OF EDINBURGH

FACULTY OF SCIENCE AND ENGINEERING

MASTER OF SCIENCE IN COMPUTER SCIENCE

DATABASE SYSTEMS

Monday 1 May 2000

2:00 to 3:30

Examiners:

I. Watson (External)

R.N. Ibbett (Chairman)

INSTRUCTIONS TO CANDIDATES

Answer any TWO questions.

All questions carry equal weight.

1. (a) With the aid of small examples, briefly describe the advantages of normalization in the design of a database for on-line transaction processing. [*5 marks*]

   (b) Why is it advantageous to *de-normalise* the database design for on-line analytical processing (OLAP) applications? [*5 marks*]

   (c) In some OLAP products, it is possible to partition a cube and to store part of the cube in a multidimensional data storage format and the rest of the cube in relational tables. If this stategy is adopted for a particular cube, how would you choose an appropriate partition in order to achieve high performance? [*5 marks*]

   (d) Explain why a consolidated dimensional hierarchy might be superior to a snowflaked hierarchy in the design of an OLAP application. [*5 marks*]

   (e) A simplified version of the syntax of a SELECT statement in the MDX is as follows.

```
SELECT <axis_specification> [, <axis_specification> ... ]
FROM   <cube_specification>
WHERE  <slicer_specification>
```

   Explain briefly the role of axis, cube and slicer specifications in formulating an MDX query. [*5 marks*]

2. (a) Briefly describe the following four isolation levels.

   i. *Uncommitted Read* [*2 marks*]

   ii. *Committed Read* [*2 marks*]

   iii. *Repeatable Read* [*2 marks*]

   iv. *Serializable.* [*2 marks*]

(b) Give an example of the *Phantom* problem, explaining the consequences of the problem for transaction isolation. [*5 marks*]

(c) When a database management system detects a deadlock cycle involving two or more transactions, it chooses one of the transactions as the 'victim' and rolls it back, undoing its work and releasing its locks, in order to let the other transactions in the deadlock cycle proceed.

A number of different victim selection strategies have been proposed, including:

- Choose the youngest transaction.
- Choose the transaction that has done the least amount of work.
- Chooose the transaction that completed the deadlock cycle.
- Choose a victim at random from among the deadlocked transactions.

With specific reference to transaction throughput and response time, describe the relative strengths and weaknesses of these victim selection techniques. You may assume that the server on which the database system executes is heavily loaded. [*8 marks*]

(d) In the context of transaction scheduling, define what it means for schedules to be *Cascadeless* and *Strict*. Give an example of a schedule that is Cascadeless but not Strict. [*4 marks*]

3. Let $R = A, B, C, D, E, F$ be a set of attributes and let

$$\Phi = \{AB \to C, C \to A, DE \to F, E \to D\}$$

be a set of functional dependencies for U.

(a) Find a minimal cover for Φ. [*5 marks*]

(b) What are the keys for R? [*5 marks*]

(c) Find a lossless join, dependency-preserving, 3NF database scheme for R and Φ. [*5 marks*]

(d) Obtain a lossless join, BCNF database scheme for R and Φ. [*5 marks*]

(e) Explain why it is not possible to obtain a lossless join, dependency- preserving, BCNF database scheme, given R and Φ. [*5 marks*]