

Two hours

SECTION A AND SECTION B ARE COMPULSORY

**UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE**

Fundamentals of Databases

Date: Thursday 15th January 2015

Time: 09:45 - 11:45

The Paper is in THREE Sections

You must answer Section A, worth 10 marks

You must answer Section B, worth 10 marks

You must answer ONE of the two questions in Section C, worth 20 marks.

Use a SEPARATE answerbook for each QUESTION.

This is a CLOSED book examination

The use of electronic calculators is NOT permitted

[PTO]

Section A

This section contains one question.
You **MUST** answer it using a **SEPARATE** answerbook.

1. a) The questions in this part are about the relational approach to data modelling.
 - i) Briefly explain the difference between a superkey and a key in the relational model. (1 mark)
 - ii) What conceptual role do foreign keys play in the relational model? (1 mark)

- b) The questions in this part are about the relational languages. They are based on the following relational schema:

```
Emp( eId:int, eName:str, age:int, salary:flt )
WorksIn( eId:int, dId:int, percentageTime:flt )
Dept( dId:int, dName:str, budget:flt, managerId:int )
```

- i) Write a SQL query that, per department, returns the department name and the number of employees that work in that department 100% of the time. (2 marks)
- ii) Write a relational-algebraic expression that returns the name and salary of employees that work at least 50% of the time for the department with dId 101. (2 marks)
- iii) Write a SQL statement that reduces the `percentageTime` that every employee works (in any department) by 2%. (1 mark)
- iv) There is a foreign key constraint binding the `Dept` and `WorksIn` relations in the above schema. As a designer, what are the four options you have for specifying the enforcement of that constraint when an attempt is made to delete a tuple from `Dept`? (2 marks)
- v) In SQL, how would you capture the requirement that every department must have a manager? (1 mark)

Section B

This section contains one question.
You **MUST** answer it using a **SEPARATE** answerbook.

2. a) The questions in this part are about functional dependencies (FDs). Suppose that we have the following three tuples in a legal instance of a relation schema S with the three attributes ABC , listed in order: $(1, 2, 3)$, $(4, 2, 3)$, and $(5, 3, 3)$. Now, let

$$F = \{A \rightarrow B, BC \rightarrow A, B \rightarrow C\}$$

be a set of postulated FDs over S .

- i) Briefly explain whether or not one can infer which of the postulated FDs in F above do hold over S on the basis of the information provided. (2 marks)
- ii) Show which postulated FDs in F above, if any, do not hold over S on the basis of the information provided. (2 marks)

- b) The questions in this part are about normalization theory.

- i) Consider the attribute set $R = ABCDEGH$ and the FD set

$$F = \{AB \rightarrow C, AC \rightarrow B, AD \rightarrow E, B \rightarrow D, BC \rightarrow A, E \rightarrow G\}$$

Now, consider the attribute set $R' \subset R = ABC$. Name the strongest normal form that is not violated by R' and explain the reason why in terms of the FDs in F . (2 marks)

- ii) Suppose you are given a relation $R(A, B, C, D)$. Assume that

$$F = \{B \rightarrow C, D \rightarrow A\}$$

are the only dependencies that hold for R . Identify the candidate key(s) for R . (1 mark)

- iii) Briefly explain, in informal terms, why the result of mapping a good ER model into a relational schema often makes it unnecessary to engage in normalization to improve the outcome of the mapping. (3 marks)

Section C

Answer one of the two questions in this section.
Use a SEPARATE answerbook for the question you choose.

3. The questions in this part are about ER modelling and ER-to-relational mapping.
- a) A company database needs to store information about employees, their children, and the departments the employees work in. Each employee has a unique ID. Information about the salary and the phone number of each employee is also stored. The children of an employee each have a different name, but it is not unknown for children of different employees to share a name. A child's age is also stored. We are not interested in information about a child once the parent leaves the company. An employee must work in one, and only one, department and, of course, a department may have many employees working for it (including none). Each department has a unique ID, as well as a name and a budget. Each department must be managed by one and only one employee. Draw an ER diagram that captures this information. (10 marks)
- b) Apply the seven-step ER-to-relational mapping procedure that you were taught in the course and show the stepwise derivation of the relational schema corresponding to the ER model you gave as your answer to the previous question. Note that some steps may not be applicable, in which case you should note that fact. Recall that a later step sometimes modifies the outcome of a previous step, in which case you should note that fact too. (7 marks)
- c) The questions below are about mapping specializations and category types to relations.
- i) Assume that an entity type S is a category over entity types T and U . Assume further that T and U have different keys. Explain what kind of additional attribute must be added to S in order to map this conceptual model to a relational one. (1 mark)
 - ii) Assume that an entity type S is specialized into disjoint sub-entity types T and U . Would the approach of using a type (or discriminating) attribute to map the hierarchy into a single relation corresponding to the super-entity type work in this case? Assuming that the specific attribute sets of T and U are disjoint, what would be an inconvenient consequence of this approach? (1 mark)
 - iii) Can the approach of mapping both super- and sub-entity types to distinct relations be applied to overlapping sub-entity types? And what about disjoint sub-entity types? (1 mark)

4. a) The questions in this part are about SQL triggers and stored procedures. They are based on the following relational schema:

```
Emp( eId:int, eName:str, age:int, salary:flt )
WorksIn( eId:int, dId:int, percentageTime:flt )
Dept( dId:int, dName:str, budget:flt, managerId:int )
```

- i) Briefly explain what business rule (i.e., a domain-specific integrity constraint) the following (obscurely named) SQL trigger is enforcing:

```
CREATE TRIGGER Mysterious
AFTER UPDATE ON Emp
WHEN old.salary < new.salary
FOR EACH ROW
BEGIN
    UPDATE Emp M
    SET M.Salary = new.salary
    WHERE M.salary < new.salary
    AND M.eId IN (SELECT D.managerId
                   FROM Emp E, WorksIn W, Dept D
                   WHERE E.eId = new.eId
                   AND E.eId = W.eId
                   AND W.dId = D.dId);
END;
```

(3 marks)

- ii) The following trigger aims to ensure that departmental budgets are still sufficient in the wake of an increase to an employee's salary. However, as written, the trigger contains several errors. Describe four of them and show how to fix each one.

```
CREATE TRIGGER EnsureSufficientDepartmentalBudget
AFTER UPDATE ON Emp
WHEN old.salary < new.salary
FOR EACH STATEMENT
DECLARE
    increment REAL;
BEGIN
    increment := new - old;
    UPDATE Dept D
    SET D.budget = D.budget + increment
    WHERE D.dId IN (SELECT W.dId
                   FROM Emp E, WorksIn W, Dept D
                   WHERE E.eId = new.eId
                   AND E.eId = W.eId
                   AND D.dId = W.dId
                   AND D.budget < (SELECT Sum(E.salary)
                               FROM Emp E, WorksIn W
                               WHERE E.eId = W.eId
                               AND W.dId = D.dId));
END;
```

(8 marks)

b) The questions in this part are about transactions and concurrency.

i) Draw the state transition diagram for a transaction as taught in this course unit.
(3 marks)

ii) Assume there are three transactions to schedule, viz., T_1 and T_2 and T_3 . Assume that:

- T_3 reads the value x , which is written by T_2
- T_1 writes the value y after it has been read by T_3
- T_2 writes the value z after it has been written by T_1

Use the information above to draw the corresponding precedence graph and conclude from it whether the schedule $T_1; T_2; T_3$ is conflict serializable or not.

(3 marks)

iii) Name two general techniques for handling deadlock discussed in this course unit and describe a drawback associated with each, if any. (3 marks)