UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2000

MSc in Computing Science
BEng Honours Degree in Information Systems Engineering Part III
MEng Honours Degree in Information Systems Engineering Part III
BEng Honours Degree in Mathematics and Computer Science Part III
MEng Honours Degree in Mathematics and Computer Science Part III
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the Associateship of the City and Guilds of London Institute
This paper is also taken for the relevant examinations for the Associateship of the Royal College of Science

PAPER M311=I3.2

DATABASES

Tuesday 16 May 2000, 14:30
Duration: 120 minutes

Answer THREE questions

Paper contains 4 questions

1a Consider relation R(A, B, C, D, E).

i) Explain precisely what the functional dependency AB→CD means.

ii) Assume that AB→CD is the only non-trivial functional dependency that holds in R. Discuss briefly any design shortcomings that R may have, in terms of redundancies and update difficulties.
What is the highest normal form that R is in?

b Consider a relation scheme S(A, B, C, D, E, F, G) with the following set of functional dependencies.

AB→CD	ABE→G
D→E	DF→G
CE→F	G→AB

i) Using Armstrong's axioms, and set theoretic concepts only, show that the functional dependency CD→G holds in S.

ii) Give three candidate keys for S.

iii) Is S in BCNF? Justify your answer.

iv) Consider a decomposition of S into two relation schema S1(D,E) and S2(A,B,C,D,F,G). Is the decomposition lossless? Is it dependency preserving? Justify your answers.

The two parts carry 25%, 75% of the marks, respectively.

2 A small business maintains the following relation schema. The primary key of each scheme is underlined.

EMP (ID, Name, Job, Salary)
BONUS (ID, Amount, Month)
ABSENT (ID, Day, Month)

Relation EMP stores information about the employees, each of which is given a unique identifier, ID. (Examples of employee's jobs are checkout, stock control, etc.) Relation BONUS records any bonuses paid to employees and the month of the payment. Relation ABSENT records the dates any employee is absent.

Consider the following queries to the relations above.

(Q1) Find the names of all employees who received bonus payments both in January and February.

(Q2) Find the names and jobs of all employees who have never been absent.

(Q3) Find the names of all checkout employees, each of which received a bonus payment in a month during which he/she was absent on at least two days.

a) Formulate all the queries above in relational algebra.
b) Formulate queries (Q2) and (Q3) in tuple relational calculus.
c) i) Formulate query (Q2) in SQL.
ii) Formulate the following query in SQL:
Find the names of all the checkout employees who had at least one bonus payment exceeding 200 pounds in January.

Parts a, b, c carry 45%, 35%, 20% of the marks, respectively.

3 This question is concerned with entity-relationship (ER) modeling and the relational model.

The Department of Computing would like to keep data about examination papers and laboratory tests that are to be set for various groups of students. The data has to reflect the following:

The students are grouped into degree years. For each degree year we have to store the degree (MSc, ISE, etc.), the year (1^{st}, 2^{nd}, etc) and the number of students in that degree year. (The MSc degree has only one year, which, we assume will be denoted as 1^{st}.)

The department offers a number of lecture courses, each of which is identified by a code, and is given a name. Each course is offered to at least one degree year, and each degree year is offered at least one course. Some courses are optional for some degree years, and others are compulsory.

The assessment of each course for each degree year is either by a laboratory test or by a written examination. For each assessment we have to record an identifying code, the number of hours it takes, and the maximum possible mark. Each course is assessed by at least one assessment. Each assessment involves exactly one course, and is taken by at least one degree year. Each degree year takes at least one assessment.

For each laboratory test we have to record what programming language it involves and what operating system it uses. For each written examination we have to record its title. The laboratory tests are set by teaching assistants. A teaching assistant may set any number of laboratory tests (possibly none). Each laboratory test is set by exactly one teaching assistant.

The written papers are set by the academics. Each paper is set by at least one academic, and each academic may set any number of papers (possibly none). Teaching assistants and academics are identified by their ID.

a Represent the data above in an ER diagram. Be careful to represent all the existence constraints and the relationship cardinalities.

b Translate the ER model into the relational model, giving the relation schema and the primary and foreign keys. You do not need to give any foreign key rules, and you do not need to indicate which relations the foreign keys refer to. But for each foreign key state whether or not it can accept NULL values.

Parts a, b, each, carry 50% of the marks.

4 a i) What are active rules? Give their general form, and very briefly describe their main components.

ii) What are foreign keys?

iii) What is referential integrity?

b The directors of an opera production maintain information about their opera in the following relation schema. The primary key of each scheme is underlined.

Singers (Name, Voice, Salary)
Cast (Name, Part)
Characters (Part, Voice)

Singers stores information about the singers, their voices (soprano, etc) and their salary. *Cast* stores information about which singer plays which part in the opera. *Characters* stores information about all the parts in the opera and the voices (soprano, etc) that they require.

i) Give two foreign keys that may be imposed on these relation schema to help data integrity. For each foreign key give the relation in which it occurs and the relation it should reference.

ii) Using informal notation, suggest two active rules aimed at helping to maintain the referential integrity based on the foreign keys you suggested in b(i).

c Apart from helping to maintain database integrity, give two other possible applications for active rules in general.

The three parts carry 40%, 40%, 20% of the marks, respectively.