UNIVERSITY OF LONDON

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 1996

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

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

PAPER I3.2 / M311

DATABASES

Tuesday, May 14th 1996, 3.00 - 5.00

Answer THREE questions

For admin. only: paper contains
5 questions
5 pages (excluding cover page)

1 Consider the following relation schema for the movies database:

Films(<u>title</u>, <u>director</u>, year, cost)

Takings(<u>title</u>, <u>director</u>, value)

Actors(<u>name</u>, <u>title</u>, <u>director</u>)

Prizes(<u>title</u>, <u>director</u>, <u>award</u>)

where the following functional dependencies hold.

title,director -> year,cost

title,director -> value

a Formulate the following queries in relational (tuple) calculus:

i) Which actors have appeared in more than two films, and what are the films each such actor played in?

ii) Which actors made their last film before 1990?

iii) A profitable film is one where the value of the takings is in excess of the cost. Give the name of the director and actors for each profitable oscar-winning film.

b Repeat part a) using SQL.

c i) Define a view **BoxOfficeHit** on the base relations **Films**, **Takings** and **Prizes** to store the attributes **title**, **director**, **value** and **award** using relational algebra, where the film has been awarded either an Oscar or a Bafta, and has takings value in excess of $50,000,000.

ii) Give the title and director of **BoxOfficeHit** films that have won both an Oscar and a Bafta award. Formulate the query in relational algebra on the **BoxOfficeHit** view.

iii) Repeat part ii) but formulate the query on the base relations.

iv) Is **BoxOfficeHit** in BCNF? Explain your answer.

2a i) Give Armstrong's Axioms for reasoning with functional dependencies.

ii) Derive the Additivity Axiom from Armstrong's Axioms.

b Consider the following relation scheme R and set FD of functional dependencies:

R(A, B, C, D, E, F, G)

FD:

AB→CDEFG
BC→E
E→F
AB→AG
G→D

i) Show that R is not in the third normal form (3NF).

ii) Normalise R into a set of 3NF relations with a decomposition that is lossless and dependency preserving. For each of the resulting relations indicate the candidate keys. At each stage of the decomposition show that it is lossless.

Show that the decomposition is dependency preserving.

c i) Define the fourth normal form (4NF).

ii) Show that any 4NF relation scheme is also in the second normal form.

iii) Give an example of a 4NF relation scheme which is not 3NF.

Turn over ...

3a Consider relation schema R(A, B, C), S(A, B, C), T(D, E, F). Assume R and S are union compatible.

i) Give an algebraic expression equivalent to $[R \cap (R-S)] \cup [S-(S-R)]$ which is as simple as possible, i.e. has as few algebraic operators as possible.

ii) For each of the following give an equivalent algebraic expression in terms of the set operators:

R join T
R join S

b A library keeps records of its books, sections and librarians in a database of relations with the following schema:

Books(BID, Title, Section, Year, Publisher)
abbreviated to **B(BID, T, S, Y, P)**

Library(Section, LibrarianName, Budget)
abbreviated to **L(S, LN, B)**

where
a tuple <bid, t, s, y, p> in **Books** indicates that the book identified by code bid has title t, is published by p in year y and is kept in section s of the library, and

a tuple <s, l, b> in **Library** indicates that section s of the library is the responsibility of librarian l and has a budget of b pounds.

Note that two different copies of the same book have different codes (BIDs), but the same title, publisher and year of publication.

Formulate the following queries in relational algebra:

i) Find all sections that have more than one copy of any book.

ii) Find the names of all librarians whose sections have no books that were published later than 1993.

iii) Find all the sections which have the highest budget.

iv) Find all the sections that only stock books from a single publisher

c On the basis of the relation schema of part b, formulate queries b(i) to b(iv) in relational calculus.

4 This question is mainly about entity-relationship modelling and partly about relational calculus.

The Department of Computing wishes to maintain the following information about its courses and staff.

Dept-info:

The students in the Department are divided into degree years (for example second year BENG, third year MENG, etc.). For each degree year we wish to record the degree, the year and the number of students registered in that degree year.

The Department runs a number of lecture courses. Each lecture course is identified by a code and has a title.

Each degree year includes a number of compulsory lecture courses (at least one) and may include some optional lecture courses. Each lecture course is offered to at least one degree year as a compulsory course, and may or may not be offered to any degree year as an optional course.

Each lecture course is taught by at least one member of the academic staff. Each member of the academic staff has a specialisation, teaches at least one lecture course and may co-ordinate the curriculum of at most one degree year. Each degree year has exactly one academic curriculum co-ordinator.

Members of the academic staff are staff members. Each staff member is identified by a name and has a rank (for example lecturer or professor).

Teaching assistants (TAs) are also staff members. Each TA provides tutorial help for at least one lecture course and co-ordinates the laboratory work of exactly one degree year. Each lecture course has one or more TAs as tutorial helper, and each degree year has exactly one TA as its laboratory co-ordinator. Some TAs are also registered as part-time PhD students.

a Represent the information in Dept-info in an entity-relationship diagram. Be careful to indicate all the existence constraints. Do not include the Department as a separate entity, but include all the entities that are indicated in the text as relevant to the Department.

b Translate the entity-relationship diagram into a set of relation schema. Indicate the primary keys. You do not need to indicate the foreign keys.

c Using the relation schema of part b formalise the following integrity constraints in relational tuple calculus:

i) No professor co-ordinates the curriculum of any degree year.

ii) TAs who are registered as part-time PhD students do not provide tutorial help for more than one lecture course.

d Repeat part c), formalising the integrity constraints in domain calculus.

Turn over ...

5a i) What is a transaction?

ii) Explain the SQL commands COMMIT and ROLLBACK.

b i) Explain when two schedules consisting of the same set of transactions are equivalent.
Explain when two schedules are conflict equivalent.

ii) Explain what is meant by a conflict serialisable schedule.
Explain what is meant by a serialisable schedule.

iii) Show, by an example, that not all serialisable schedules are conflict serialisable.

c i) Show informally, but generally, that all conflict serialisable schedules are serialisable.

ii) Show that under the two-phase locking protocol any schedule will be serialisable provided it does not deadlock.

Parts a, b and c carry, respectively, about 20%, 30%, 50% of the marks.

End of paper