UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2002

MSc in Computing Science
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
MSci 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

Friday 10 May 2002, 10:00
Duration: 120 minutes

Answer THREE questions

Paper contains 4 questions
Calculators not required

1a Let R(A,B,C,D,E,F,G) be a relation scheme with the following set S of functional dependencies:

$AB \rightarrow CD$	$C \rightarrow E$
$E \rightarrow A$	$ED \rightarrow F$
$CA \rightarrow D$	$ECD \rightarrow G$
$EB \rightarrow D$	

i) Explain why EBG is not a candidate key for R.

ii) Give a canonical cover for S.

iii) Give three candidate keys for R.

iv) Is R in BCNF (Boyce-Codd normal form)? Give a brief explanation for your answer.

v) Which attribute in R can never have a NULL value? Justify your answer in detail.

b Normalisation of a relation involves a decomposition of it. A property that a decomposition might have is dependency preservation. In general what advantages does a decomposition that is dependency preserving have over one that is not?

c A relation R is in 4NF if and only if whenever it has a multivalued dependency $M \rightarrow\rightarrow N$ then either $M \rightarrow R$ or $M \rightarrow\rightarrow N$ is a trivial multivalued dependency. Prove in detail that a relation that is in 4NF according to this definition is also in BCNF.

Parts a, b, c carry, respectively, 70%, 10%, 20% of the marks.

2a For each of the following state whether or not it is true. In either case give an example to illustrate your answer.

i) $(A*B) \div B = A$

ii) $(A \infty B) \div B = A$

b Consider the following relation schemes dealing with the administration of a chain of gyms called GYMFIT. The primary keys are underlined.

R(<u>GymID</u>, City, AreaCode, ManagerID, AverageAnnualIncome)
S(<u>GymID</u>, <u>Facility</u>)
T(<u>ManagerID</u>, ManagerName, Salary)

Tuples in R give the city, area code, manager ID and the average annual income of the gyms run by GYMFIT. Tuples in S give the facilities available in the gyms, and tuples in T give the name and salary of the gym managers.

Consider the following queries.

Q1) List all the cities that have more than one gym run by GYMFIT located in them.

Q2) List the names and salaries of all gym managers whose gyms have the lowest average annual income.

Q3) List the names of all gym managers whose gyms have both a swimming pool and a sauna.

Q4) List the Ids of all gym managers each of whom runs all the GYMFIT gyms in any one city.

i) Formalise Q1, Q3, Q4 in relational algebra.

ii) Formalise Q2, Q4 in relational calculus.

Parts a, b(i), b(ii) carry, respectively, 25%, 45%, 30% of the marks.

3a i) What is transaction atomicity in databases?

ii) What information has to be kept on the database log to help ensure transaction atomicity?

iii) What is meant by the *idempotence* of redo and undo operations? What is its significance in transaction management? How can it be ensured?

b What is the *check option* that can be used in SQL view definitions? What is its purpose?

c Consider the following relation schemes and views.

R1(A,B,C,D)   R2(A,B,F)

CREATE VIEW V1 AS SELECT A,C FROM R1;

CREATE VIEW V2 AS SELECT A,B,D FROM R1 WHERE D<500;

Let AB be a foreign key in R2 referencing R1.

Let T1, T2, T3 be the transactions below.

T1={delete from R1 the tuple <a1,b1,c1,400>, add to R1 the tuple <a1,b2,c1,600>}
T2={add to V1 the tuple <a2,c2>, delete from V1 the tuple <a3,c3>}
T3={add to R2 the tuple <a4,b4,f4>, delete from V2 the tuple <a5, b5, 300>}

i) For each of T1,T2,T3 list all the problems that the transaction might have in terms of its ambiguity and the violation of the integrity of the relations and the views.

ii) Explain how an active rule might be employed to avoid one of the problems you have listed in c(i).

Parts a, b, c carry, respectively, 45%, 15%, 40% of the marks.

4 Cats-Heaven is an organization that looks after pet cats whilst their owners are away. It would like to keep data about the cats it takes in. The data has to reflect the following:

Cats-Heaven Data:
Each cat is identified by a number, has a type (ginger, tabby, etc.), has an owner and is housed in a room. The owners are identified by their name and give an address. An owner may place several cats, and at least one, in Cats-Heaven. Each room is identified by a number and may contain any number of cats or none at all. The cats in any one room are all of the same type. Each room is looked after by at least one member of staff. Each member of staff is identified by a name and may look after one or more rooms.

a Consider a relation
CATS(cat_no, type, owner_name, owner_address, room_no, staff_name)
as an attempt to design a relation to store the data required by Cats-Heaven.

i) Identify and represent all the non-trivial functional dependencies in the relation CATS.

ii) Show that CATS is not in 2NF (second normal form).

iii) Suppose CATS is decomposed into two relations
The_Cats(cat_no, type, owner_name, owner_address, staff_name)
The_Staff(staff_name, room_no).
Is this decomposition lossless? If yes, prove it, and if no, show an example of instances of the three relations that demonstrates that the decomposition is not lossless.

b i) Represent the Cats-Heaven data in an Entity-relationship (ER) diagram. Use entities CAT, OWNER, ROOM, STAFF. Be careful to represent all the existence constraints and the relationship cardinalities.

ii) Translate the ER model into the relational model, giving the relation schemes and the primary keys. You do not need to give any foreign keys.

iii) Bearing in mind the data requirements of Cats-Heaven, identify three advantages that your relations resulting from b(ii) have over the relations The_Cats and The_Staff, as far as updates are concerned.

Parts a, b carry, respectively, 55%, 45% of the marks.

© University of London 2002