UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2000

BEng Honours Degree in Computing Part II
MEng Honours Degrees in Computing Part II
BEng Honours Degree in Mathematics and Computer Science Part II
MEng Honours Degree in Mathematics and Computer Science Part II
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 C230=MC230

DATABASES I

Friday 5 May 2000, 16:00
Duration: 90 minutes
(Reading time 5 minutes)

Answer THREE questions

Paper contains 4 questions

1a Describe the main features of a *data dictionary*. Discuss how such a facility can help in maintaining the integrity and security of a database system.

b A large trading company is divided into several departments. Each department, identified by a department number, is manned by several employees whose activities are supervised by a manager (who is also an employee). Employees are associated with only one department. The attributes recorded about employees are employee number, name and salary.

The company stocks a large number of items in its warehouse with each department having responsibility for selling some of these items. No more than one department would deal with a given item. The company purchases its items from a large number of suppliers, with the relationship between items and suppliers being many-many. The price charged to the company for a given item may vary from supplier to supplier; the company would, however, decide on a common price to be charged for the item when selling it to its customers. Suppliers are identified by their name and address.

The company has numerous customers who place orders for the items stocked. Each order is identified by a unique order number and the information recorded about an order would include the date of the order, the item ordered (identified by item number), the quantity of the item required and the customer (identified by name and address) placing the order.

i) Construct an entity-relationship diagram to include all entities, attributes and relationships implied by the above information. Do not include 'company' as a separate entity, but include all those entities (eg department, customer, supplier etc) of interest to the company. Identify any dominant/subordinate and strong/weak entities.

ii) Transform the entity-relationship diagram into a set of relations, indicating the primary key of each relation.

The two parts carry, respectively, 35% and 65% of the marks.

2a Explain what is meant by a *superkey.*

Distinguish between *candidate*, *primary*, *alternate* and *foreign* keys.

b The personnel section of a large organisation keeps a record of the courses taken by members of staff and the grades obtained at the conclusion of these courses. An example of the type of data recorded is given below.

Employee No	Name	Course Title	Course Code	Run No.	Start Date	Grade
1234	Green	DB - introd.	D456	73	01.2.99	A
		DB - advanced	D473	24	08.2.99	F
			D473	25	12.3.99	C
9876	Brown	Basic Comp.	B246	58	01.9.98	B
		DB - introd.	D456	73	01.2.99	C
9988	Black	Basic Comp.	B246	65	12.2.99	A

Over the years a course is repeated a large number of times. The Run Number gives the number of times that a course has been mounted - thus Run No. =24 means that this is the 24th time that this course is being presented. It may be assumed that an employee is not able to follow more than one course at a time. A, B, C are pass grades, F indicates failure in which case the course must be repeated.

Illustrate in a diagram the functional dependencies existing in this system stating clearly any assumptions you make regarding the semantics of the data.

State with reasons, in which normal form the data is currently held and, if necessary, decompose the data (in a non-loss manner) into Boyce-Codd normal form (BCNF). Explain clearly why the decomposition is non-loss.

c Using the BCNF relations, express the following query in *relational algebra:*

List the employees (by name) who have failed more than one course.

The three parts carry, respectively, 20%, 60% and 20% of the marks

3a State Armstrong's axioms of Reflexivity, Augmentation and Transitivity.

Using only these axioms, prove the following two additional rules:

i) If $A \rightarrow B$ and $A \rightarrow C$, then $A \rightarrow BC$

ii) If $A \rightarrow B$ and $BC \rightarrow D$, then $AC \rightarrow D$

b State the conditions necessary for two sets of functional dependencies to be equivalent.

A relation has been defined over the attributes A, C, D, E, H.

The following two sets of functional dependencies have been defined for the relation:

1. $A \rightarrow C$, $AC \rightarrow D$, $E \rightarrow AD$, $E \rightarrow H$

2. $A \rightarrow CD$, $E \rightarrow AH$

Determine whether these two sets of functional dependencies are equivalent, clearly explaining your answer.

c Define what is meant by a *view* in the context of relational databases.

Retrieval operations based on views can be implemented by using either the *substitution* approach or the *materialisation* approach. Explain what is meant by these and discuss the advantages and disadvantages of each such approach.

The three parts carry, respectively, 35%, 35% and 30% of the marks.

4a Define what is meant by a *complete history*.

Distinguish clearly between a *serial* history and a *serialisable* history.

Transactions T_1 and T_2 are to be run concurrently. The following complete history has been suggested for their execution:

$$H_1 = w_1[a]\ w_1[b]\ r_2[g]\ w_2[a]\ r_2[b]\ w_2[b]\ c_2\ w_1[d]\ c_1$$

Determine whether this history is serialisable. Comment on the actions that would be required to ensure the overall consistency of the database, and the feasibility of carrying out those actions, if the last step in H_1 were a_1 (T_1 aborting) rather than c_1.

b Explain what is meant by the *two phase locking protocol* (2PL) and distinguish between *Basic* 2PL and *Strict* 2PL.

Explain why serialisability of a history is guaranteed if the transactions participating in that history conform to the 2PL protocol. (Hint: Consider the serialisation graph of the history).

c Discuss what is meant by *livelock.*

Compare and contrast the *Wound-Wait* and *Wait-Die* approaches to deadlock prevention and show how livelock is prevented under these approaches.

The three parts carry, respectively, 40%, 30% and 30% of the marks.