

UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2004

BEng Honours Degree in Computing Part I
MEng Honours Degrees in Computing Part I
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*

PAPER C130

DATABASES I

Friday 7 May 2004, 10:00
Duration: 90 minutes
(Reading time 5 minutes)

Answer THREE questions

Paper contains 4 questions
Calculators not required

- 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 Entity-relationship diagrams typically contain binary relationships. Using an appropriate example, discuss circumstances where binary relationships cannot adequately model the real world situation.
- c An individual (known as a *policy-holder*) may take out a number of insurance policies with a given insurance company. Each such policy, which relates to a single policy-holder, has a unique policy number and gives the date the policy started.

The insurance company's policies can provide cover for different types of risk eg. a motor policy may provide cover for third party, fire, theft, accident damage etc. Each such level of cover is known as a *product* and any one policy may include several such products (eg a policy-holder may choose a motor policy which includes three products - third party, fire and theft).

The insurance company's products (and so also, policies) are grouped by *business area* eg. life insurance, motor insurance, house contents insurance etc. Any given product belongs to only one business area.

Insurance Brokers (identified by a unique Broker number) act as agents for the insurance companies and sell policies to individuals. They receive a commission for each policy sold; the rate of commission is negotiated annually between the insurance company and the Broker. A Broker may be associated with only one insurance company, though any one company may have several Brokers associated with it. Individuals may only purchase policies through Brokers.

- i) Construct an entity-relationship diagram to include all entities, attributes and relationships implied by the above information. 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 three parts carry, respectively, 35%, 15% and 50% of the marks.

2a State Armstrong's axioms of Reflexivity, Augmentation and Transitivity.

Explain what is meant by an *irreducible cover* for a set F of functional dependencies.

Let D be the set of functional dependencies:

$$A \rightarrow B, ABCD \rightarrow E, EF \rightarrow G, EF \rightarrow H, ACDF \rightarrow EG$$

Determine the irreducible cover for the set D.

- b i) Explain what is meant by a *non-loss decomposition* and state the condition under which a given decomposition will be non-loss.
- ii) The personnel section of a large organization 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	01.12.03	A
		DB – advanced	D473	24	08.2.04	F
			D473	25	12.3.04	C
9876	Brown	Basic Comp.	B246	58	01.09.03	B
		DB – introd	D456	73	01.12.03	C
9988	Black	Basic Comp.	B246	65	12.2.04	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.

- iii) Using the BCNF relations obtained in (ii) express the following query in *relational algebra*:

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

The two parts carry, respectively, 40% and 60% of the marks.

- 3a Explain what is meant by a *schedule* in the context of concurrent running of transactions.

Distinguish between a *serial* and a *serialisable* schedule.

The following two schedules have been suggested for transactions T1 and T2:

	<u>Schedule 1</u>		<u>Schedule 2</u>	
Time	T1	T2	T1	T2
1	read (X)		read (X)	
2	X = X - A		X = X - A	
3		read (X)	write (X)	
4		X = X + B		read (X)
5	write (X)			X = X + B
6	read (Y)			write (X)
7		write (X)	read (Y)	
8	Y = Y + A		Y = Y + A	
9	write (Y)		write (Y)	

Considering each schedule in turn and by attempting to reorder the operations within the schedule determine whether the given schedule is serialisable. Confirm your results by drawing the appropriate serialisation graph.

- b Describe the Timestamp Ordering scheme for scheduling concurrent transactions.

Discuss whether the schedule below will lead to a serialisable execution of the transactions T3, T4 and T5 if concurrency is to be controlled using:

- Timestamp ordering
- Basic two-phase locking

Time	T3	T4	T5
1		read (Z)	
2		read (Y)	
3		write (Y)	
4			read (Y)
5			read (Z)
6	read (X)		
7	write (X)		
8			write (Y)
9			write (Z)
10		write (X)	
11	read (Y)		
12	write (Y)		

The two parts carry, respectively 50% and 50% of the marks.

- 4 The process of recovery sometimes involves *undoing* the effect of certain transactions and *redoing* the effect of other transactions.
- a Explain why the *undo* and *redo* operations are said to be *idempotent*.

Explain carefully the process whereby a database can be restored to a consistent state following a system failure as a result of which the contents of main memory and the input/output buffers have been lost whilst leaving the database itself intact. Identify any features of this process which may be redundant.

- b Discuss how recovery from system failure could be guaranteed
- i) without ever having to *undo* any transactions;
- ii) without ever having to either *undo* or *redo* any transactions.
- c A travel agent making a <booking> for a customer would typically be involved in three activities
- booking an airline seat;
 - making a hotel reservation;
 - arranging a car rental.

Discuss the advantages to be gained in the areas of concurrency and recovery if each activity within <booking> is considered to be a separate transaction as opposed to considering <booking> to be one large transaction.

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