PurpleTrain.com

UNIVERSITY OF PORTSMOUTH

BACHELOR OF SCIENCE (HONS) IN COMPUTING

AUGUST 2004 EXAMINATION

17 AUGUST 2004

IT354 DATABASE DESIGN & MANAGEMENT

TIME: 3 HOURS

Students are permitted to refer the following books and course material:

1. "Fundamentals of Database systems ", Elmasri and Navathe, Addison Wesley.
2. "Database Systems – A practical Approach to Design, implementation and management", Thomas Connolly and Carolyn Begg, Addison Wesley
3. Course Contents.

NOTE: No handwritten notes and printed materials other than the above are permitted.

☞ **INSTRUCTIONS:**

- Choose any **THREE (3)** questions out of **FIVE (5)** questions.
- Each question is worth 33 marks. 1 extra mark will be added to all candidates scripts to construct a percentage.
- **DO NOT** answer more than **THREE (3)** questions in total.
- Please start every question on a new page.
- Answers will not be marked if they are illegible.
- Write the question numbers (in the order you have attempted them) in the boxes provided in the answer script.
- Write your **INDEX NUMBER** and **MODULE NUMBER** on the cover page of the answer booklet(s).

QUESTION 1

a) An Art gallery maintains a database of photographs that can be bought over the internet.

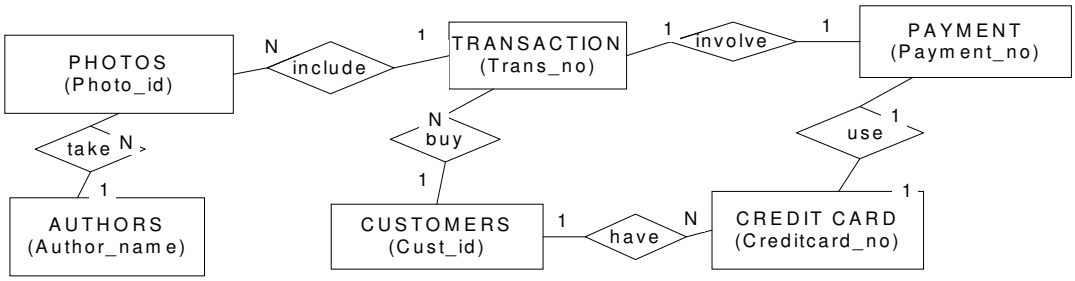
The database has three functions:

i) maintaining the catalogue with entities *photograph* and *photographer;*

ii) selling the photos with entities *transaction* and *customer;*

iii) maintaining the payment particulars with entities *payment* and *credit card details.*

The ER diagram of the existing system is shown in the diagram below. The primary keys are given in parenthesis.



You are required to enhance this ER model to an EER one by adding the following details:-

1) The database will store some technical information for every photograph, like the type (colour or black and white), type of film, shutter speed, f-stop, and price.

2) The photographs are to be grouped into three categories: Landscapes, Portraits and General (everything else). Landscape has information about the location of where the photograph was taken. Portrait has information about the model used The General category can be easily extended to describe other types of photographs, such as Underwater, Nature, Sports, War, etc.

A location is identified by a unique number, and a model by its name.

(15 Marks)

b) Consider the following two relations *R* and *S*:

A	B
a	b
c	b
d	e

R

B	C
b	c
e	a
b	d

S

Compute:

(i) $R \rhd\lhd S$

(ii) $S \div \{b, c\}$

(iii) $R \rhd\lhd_{B<C} S$ (where < means alphabetic order on letters)

(18 Marks)

[TOTAL FOR QUESTION 1: 33 MARKS]

QUESTION 2

a) A small database with only three relations, R1, R2 and R3 is distributed over three sites, S1, S2 and S3. You are to allocate each of the three relations R1, R2 and R3 to one of the three sites S1, S2 and S3 using the nonredundant best fit method, (which is one technique which helps DBAs to decide how to distribute data across sites).

There are only three transactions run on this database: T1, T2 and T3.
The originating sources for each transaction (that is the sites at which they are run) and the percentage frequency with which the transactions run are shown below:-

Transaction sources and frequencies

T1	sources S1 and S2	20%
T2	sources S1 and S3	30%
T3	sources S2 and S3	50%

Note that transactions are equally likely to originate from any of their sources.

The usage made of each relation by each transaction is given by:-

Transaction behaviour

T1	2 reads to R1 and 2 writes toR1 2 reads to R2
T2	3 reads to R1 2 reads to R3 and 1 write to R3

T3	2 reads to R2 and 1 write to R2 2 reads to R3

(15 Marks)

b) Consider the following set, *F*, of functional dependencies:

$AB \rightarrow C$	$D \rightarrow EG$
$C \rightarrow A$	$BE \rightarrow C$
$BC \rightarrow D$	$CG \rightarrow BD$
$ACD \rightarrow B$	$CE \rightarrow AG$

Two analysts, Mike and Dave have each produced a different minimal cover for *F* (given below).
Decide and give reasons for your decision whether each or either of these actually is a minimal cover for *F*.

Mike's "minimal cover":
AB → C
C → A
BC → D
CD → B
D → E
D →G
BE → C
CG → D
CE →G

Dave's "minimal cover":
AB → C
C → A
BC → D
D → E
D →G
BE → C
CG → B
CE →G

(18 Marks)

[TOTAL FOR QUESTION 2: 33 MARKS]

QUESTION 3

Tomes On Line, an internet only bookseller, requires a database to keep information on all of its offered books, together with their authors, publishers, and warehouse locations where the books are stocked.

For each publisher, the database must record the publisher code, name, and the publisher's address(es). A publisher can have several different addresses. No two publishers have the same name.

For each book, the database must record its ISBN code, title, the publisher code, the price, and published format (Possible formats are: hardback, paperback and/or audioCD). Note that for any title published in several formats, each separate format has a distinct ISBN code. In addition, the database must record the identity number and name of each of the authors of the book. If there is more than one author, they must be listed in the order in which they are listed on the book. This may or may not be ordered alphabetically. Furthermore the warehouse location and number in stock at that warehouse are also recorded for each book.

For each warehouse, the database must record the identity number, the name, the location, and the number of employees at that warehouse.

a) Identify the possible functional dependencies.

(14 Marks)

b) Hence give the BCNF relations (do ***not*** go through the intermediate normal forms 1NF, 2NF, 3NF). In each of your relations identify the primary key.

(6 Marks)

c) A previous attempt at constructing this database had produced the relation

WROTE (ISBN_CODE, AUTHOR_NUMBER, FORMAT)
with some sample data as shown in the table below:-

ISBN_CODE	AUTHOR_NUMBER	FORMAT
0-321-44567	3345	Hardback
0-321-44567	3345	Paperback
0-321-44567	2178	Hardback
0-321-44567	2178	Paperback
0-321-44567	4356	Hardback
0-321-44567	4356	Paperback
0-456-38921	3345	Hardback

Show that a multi-valued dependency exists within this relation and normalize the relation to 4NF.

(13 Marks)

[TOTAL FOR QUESTION 3: 33 MARKS]

QUESTION 4

a) Draw the conflict graph for the following schedule, S, of four transactions:
S: W1(x)W2(x)R3(x)R3(y)R1(y)R1(x)W4(y)
Hence determine whether or not the schedule is serializable.

(10 Marks)

b) Consider the following schedule of four transactions. Show how this schedule would be executed under strict 2PL(two phase locking).
R1(y)W4(x) R4(y)R3(w)R2(y)W1(y)W1(x)W2(x)W2(z)W3(x)R4(z)

(12 Marks)

c) i) What will happen when a 2PL is applied to the following schedule?
C = R1(x)R2(y)W2(y)R2(x)W2(x)R1(y)C1C2

(5 Marks)

ii) Consider the following protocol:
"If T_j is not blocked (i.e. not waiting for some locked item), then T_i is blocked and allowed to wait; otherwise abort T_i and restart T_i after a time delay" (the cautious waiting protocol).
What happens if we apply this protocol to the schedule C?

(6 Marks)

[TOTAL FOR QUESTION 4: 33 MARKS]

QUESTION 5

a) A distributed database system has been set up for the Klatchian Delivery Service and is located at two regional sites, Laotan and Elharib. The HQ site is at Smale, and no data is actually held there.

The relational schema and the details of records in the respective relations are given below.

.

Laotan: i) DRIVER (Driver_id, Driver_name, Licence_no, date_of_birth)
Number of Records: 100
Each attribute consists of 15 bytes

ii) DRIVER_VEHICLE(Driver_id, Vehicle_id)
Number of Records: 2000
Each attribute consists of 15 bytes

Elharib: i) VEHICLE(Vehicle_id, Vehicle_make)
Number of records = 20
Each attribute consists of 15 bytes

ii) TRIPS(Trip_no, date, delivery_address, Vehicle_id)
Number of records = 1000
Each attribute consist of 15 bytes

You can assume that the cost of data transfer from one site to another is $1.00 per kb. and that all other costs are negligible compared with the data transfer cost. Furthermore you can assume that one Vehicle will make only one trip in a day and that during the period 01-01-2004 to 30-04-2004, a vehicle will make a maximum of 90 trips

At HQ in Smale, the SQL Query given below is used to get information about the number of trips made by the vehicle with vehicle_id= 'SH4004' during the period 01-01-2004 to 31-03-2004.

```
SELECT   Driver. Driver_id,  Vehicle.Vehicle_id, Count(Trips.Trip_no)
FROM     Driver ,  Driver_vehicle , Vehicle , Trips
WHERE    Driver. Driver_id = Driver_vehicle. Driver_id
```

AND Driver_Vehicle.Vehicle_id = Vehicle.Vehicle_id
AND Vehicle. Vehicle_id= Trips. Vehicle_id
AND Driver Driver. Driver_id
AND Vehicle_id = 'SH4004'
AND date >01-01-2004
AND date <31-03-2004

i) Identify two different query strategies to execute the query at Smale.
(4 Marks)

ii) Calculate the cost involved in executing the two strategies and hence identify the optimum price-related strategy.
(7 Marks)

b) Consider the relation Student_Unit with attributes and data given by:-

Student_id	Student_name	Date of birth	Course_id
010191012222	Tan Kok	12-03-1984	IT354
010191012222	Tan Kok	12-03-1984	IT356
010191012222	Tan Kok	12-03-1984	IT355
010191021234	Chai Wei Mng	26-10-1980	IT356
010191021234	Chai Wei Mng	26-10-1980	IT355
010191021234	Chai Wei Mng	26-10-1980	IS352
010191021234	Chai Wei Mng	26-10-1980	IS358
010191021234	Anton	21-09-1981	IS352
030091015390	Joseph Robert	10-03-1978	IT355
010191021234	Anton	21-09-1981	IT354
010191021234	Anton	21-09-1981	IT356
400091013456	Ramesh	04-04-1978	IS352
400091013456	Ramesh	04-04-1978	IT354
400091013456	Ramesh	04-04-1978	IT356
020091016578	Jannice	12-12-1979	IS352
020091016578	Jannice	12-12-1979	IT354
020091016578	Jannice	12-12-1979	IT356
030091015390	Joseph Robert	10-03-1978	IT352

A horizontal fragmentation of this relation into relations F1, F2, F3, and F4 has been made with the fragments given as:-

F1

Student_id	Student_name	Date of birth	Course_id
010191012222	Tan Kok	12-03-1984	IT354
010191021234	Anton	21-09-1981	IT354
030091015390	Joseph Robert	10-03-1978	IT355
400091013456	Ramesh	04-04-1978	IT354
020091016578	Jannice	12-12-1979	IT354

F2

Student_id	Student_name	Date of birth	Course_id
010191012222	Tan Kok	12-03-1984	IT355
010191021234	Chai wei Mng	26-10-1980	IT355
030091015390	Joseph Robert	10-03-1978	IT355

F3

Student_id	Student_name	Date of birth	Course_id
010191012222	Tan Kok	12-03-1984	IT356
010191021234	Chai Wei Mng	26-10-1980	IT356
010191021234	Anton	21-09-1981	IT356
400091013456	Ramesh	04-04-1978	IT356
020091016578	Jannice	12-12-1979	IT356

F4

Student_id	Student_name	Course_id
010191021234	Chai Wei Mng	IS352
010191021234	Anton	IS352
400091013456	Ramesh	IS352
020091016578	Jannice	IS352
030091015390	Joseph Robert	IT352

i) Identify the predicate used to create the fragments.

(1 Mark)

ii) Write relational algebra expressions to create the horizontal fragments based on the predicate identified.

(1 Mark)

iii) Show whether the given fragments satisfy complete, disjoint and reconstruction conditions.

(7 Marks)

c) Consider the following relational schema:

DOCTORS(Dr id, Name, Specialty, Telephone, Years of experience)

PATIENTS(Patient id, Name, Sex, Address, Age)

RECORDS(Dr id, Patient id, Date, Diagnosis)

Write relational algebra expressions for the following queries:

i) List the names of doctors and patients in pairs, so that, for each pair, the patient is over 60 years old and saw the doctor in 1999.

(5 Marks)

ii) Find the names of all male patients who have never seen a doctor.

(4 Marks)

iii) Find the patient name and diagnosis for all patients treated by doctors with the specialty "Neuro surgery"

(4 Marks)

[TOTAL FOR QUESTION 5: 33 MARKS]

- End of Paper -