

22/11/18

BITS PILANI, DUBAI CAMPUS
DUBAI INTERNATIONAL ACADEMIC CITY, DUBAI
SECOND SEMESTER 2017 - 2018
COURSE: DATABASE SYSTEMS (CS F212)
COMPONENT: Test 1 (CLOSED BOOK) DURATION: 50 Minutes
WEIGHTAGE: 20% (40 Marks) DATE: 26.02.2018 TIME: 8.30-9.20 am.

This question paper has 2 pages.

1. What is the difference between LOGICAL DESIGN and PHYSICAL DESIGN w.r.t. database design? [2M]
2. How do **each** of the following professionals use a Database - **Naive User, Application Programmer, Analyst, Database Administrator**? [4M]
3. What are the responsibilities of **Storage Manager** in a DBMS? [2M]
4. Consider the following collection of relation schemes:
 professor(profname, deptname)
 department(deptname, building)
 committee(profname, commname)

Write the **Relational Algebra** queries for each of the following:

- a. Find all the professors who are in any one of the committees that Professor Smith is in.
- b. Find all the professors who are in at least all those committees that Professor Smith is in.
- c. Find all the professors who are in exactly (i.e., no more and no less) all those committees that Professor Smith is in.
- d. Find all the professors who have offices in at least all those buildings that Professor Smith has offices in. [2+3+4+3 =12M]

5. Micro loans are small loans, which is beginning to gain popularity especially among borrowers in developing countries. The idea is to bring venture lenders together using information technology. Typically, the loans will be used to finance startup or development of the borrower's company, so that there is a realistic chance for repayment. The money in a loan can, unlike traditional loans, come from many lenders. The following information form the basis for creating the model:

- Each borrower and lender must be registered with information about name and address.
- A loan starts with a loan request, which contains information about when the loan should at latest be granted, The total amount being discussed (US-dollars), and how long the payback period is. Also, a description is included of how the money will be used. The rent on the payment is calculated in the loan amount, which is to say, the full amount is not paid.
- Lenders can commit to an optional portion of the total amount of a loan request.
- When the commitments for the loan request covers the requested amount, the request is converted to a loan. If not enough commitments can be reached, the loan request is cancelled. A borrower can have more than one request, and more than one loan at time, but can at most make one request per day.
- The loan is paid through an "intermediary", typically a local department of a charity, who has a name and an address.
- The borrower chooses when he or she will make a payment. Every payment must be registered in the database with an amount and a date (at most one payment per loan per day). The lenders share the repayment based on how large a part of the loan they are responsible for.

P.T.O

- If the loan is not repaid before the agreed upon deadline, a new date is agreed. The database must not delete the old deadline, but save the history (the deadline can be overridden multiple times).
- Each lender can for each borrower save a “trust”, which is a number between 0 and 100 that determines the lender’s evaluation of the risk of lending money to that person. The number must only be saved for the borrowers, for whom there has been made such an evaluation.

a) Make an ER/EER model for the data described above. If you make any assumptions about data that doesn’t show from the problem, they must be described. Put an emphasis on having the model express as many properties about the data as possible, for instance participation constraints.

b) Make a relational data model for micro loans:

- Describe at least two of the relations using SQL DDL (make reasonable assumptions about data types), and
- state the relation schemas for the other relations.

The emphasis is if there is a correlation between the relational model and the E-R diagram from a), along with primary key and foreign key constraints being stated for all relation.

It is not necessary to state CHECK constraints and the like. [7+7M]

6. Consider the given table and give one example for each of the following types of constraints. Kindly note that the underlined keys denote Primary keys.

a. Key Constraint

b. Entity Integrity

c. Domain [2+2+2M]

<u>Title</u>	<u>Year</u>	Length	3D	StudioName	Producer
Titanic	2006	360	Yes	Paramount	James Cameroon
Inception	2010	148	No	Universal	Christopher Nolan
Avatar	2009	162	No	Paramount	Sam
Speed	2002	250	Yes	20 th Century Fox	
Rio	2010	96	Yes	Disney	Carlos Saldana

*****ALL THE BEST*****

BITS PILANI, DUBAI CAMPUS

Dubai International Academic City
Second Semester 2017 – 2018

Quiz – 1 (Closed Book)

SET A

No. of Questions: 4

No. of Pages : 2

Year : II
Course No : CS F212
Course Title : Database Systems

Date: 13-3-18
MAX Marks: 10(5%)
Duration: 20 minutes

ID NO: _____ **Name:** _____
Faculty: _____

Consider the given two tables, where underlined key shows primary key

Executive (name, eid, address, netWorth)

Company (CompanyName, ceo, ceoID)

1. Write an assertion to ensure that no company has as its CEO as someone who's net worth is less than \$10,000,000. 2M

```
CREATE ASSERTION TooRich CHECK (  
(NOT EXISTS  
(SELECT company-name  
FROM Company, Executives  
WHERE ceo = eid AND networth < 10000000  
))  
);
```

2. Write a trigger which will foil (by undoing) any update on the attribute (netWorth), if a lower value for netWorth is entered. 2M

```
CREATE TRIGGER NetWorthTrigger  
AFTER UPDATE of netWorth on MovieExec  
REFERENCING  
OLD ROW as OldTuple,  
NEW ROW as NewTuple  
FOR EACH ROW  
WHEN (OldTuple.netWorth > NewTuple.netWorth)  
UPDATE MovieExec  
SET netWorth = OldTuple.netWorth  
WHERE cert = NewTuple.cert;
```

3. Is the above trigger a row level or a statement level trigger? Why 1M

Row level trigger because it has to execute for each row when update must be foiled

4. Consider the following Project Information System related data:

Studentmaster

idno	name	dateofbirth
001	AAA	27/10/63
004	DDD	5/6/63
002	BBB	18/5/63
005	EEE	8/3/63
003	CCC	20/4/63

Studentproject

idno	projectmarks	projectgrade
001	85	A
004	70	B
002	75	B
005	85	A
003	73	B

Write SQL queries in standard form for the following:

a) Create the above two tables.

[1 M]

b) List the *name* and *dateofbirth* for all students.

[1 M]

c) List *idno*, *name*, *projectgrade* for students who have obtained 'A' grade.

[1 M]

d) List the *idno*, *name* of the student who have got the **maximum** *projectmarks*.

[2 M]

Question Paper
BITS PILANI, DUBAI CAMPUS
DUBAI INTERNATIONAL ACADEMIC CITY, DUBAI
SECOND SEMESTER 2017 - 2018
COURSE: DATABASE SYSTEMS (CS F212)
COMPONENT: Test 2 (OPEN BOOK) DURATION: 50 Minutes
WEIGHTAGE: 15% (30 Marks) DATE: 10.04.2018 TIME: 8.30am-9.20am.
This question paper has 2 pages.

1. A Recruitment Agency called RECRUIT_MANPOWER assigns part-time/temporary staff to perform cleaning services, at all universities in the UAE.

The following information is to be stored in a database.

<temporary_staff_name, emiratesid_number_of_temporary_staff,
contract_number, hours_per_week, university_id, university_location>

The functional dependencies are stated as follows:

{emiratesid_number_of_temporary_staff, contract_number} \rightarrow hours_per_week

emiratesid_number_of_temporary_staff \rightarrow temporary_staff_name

contract_number \rightarrow {university_id, university_location}

university_id \rightarrow university_location

You required to **design the database in 3NF**. (comprising of requisite number of tables or relations along with primary keys, foreign keys and non-key attributes). [7 M]

2. Consider a relation with schema R(A,B,C,D) and FDs
{AB \rightarrow C, C \rightarrow D, D \rightarrow A}.

Find the $(AB)^+$ (i.e. attribute closure of AB). [4 M]

3. Using only Armstrong's Axioms and the FDs AB \rightarrow C, A \rightarrow BE, C \rightarrow D,
give a complete derivation of the FD A \rightarrow D. [4 M]

4. Disk Characteristics of the following Block-addressable disk drive:

Size of track = 20,000bytes

Nondata overhead per block (including inter block gaps) = 300 bytes

Size of track = 20,000bytes

Average seek time = 8msec

Average rotational delay = 3msec

Maximum rotational delay = 6msec

Spindle speed = 10,000rpm

Sectors per track = 170 sectors

Sector size = 512 bytes

File Characteristics:

Record Size: 100 bytes

How many records can be stored per track for the following blocking factors?

a. Blocking factor=10

b. Blocking factor=60

c. What is the average time to read one sector? [2+2+3M]

5. Consider the following tables with the specified attributes. Underlined attributes indicate the primary key

Airport (airportID, name, city)

Flight (FlightNo, FlightCompany, depAirport, arrAirport)

Booking (ticketNo, name, nationality, FlightNo, seatNo)

Seat (SeatNo, FlightNo, class)

Write queries in Tuple Relational Calculus and Domain Relational Calculus

a. Retrieve details of all bookings by British and French passengers. The schema of the output table should be same as that of the Booking table.

b. Retrieve the flight number, Departure and Arrival airports of all British Airways flights [2X4=8M]

*****ALL THE BEST*****

BITS PILANI, DUBAI CAMPUS
Dubai International Academic City, Dubai
Second Semester 2017-18
Comprehensive Exam (Closed Book)

No. of Questions:
Part A: 8
Part B: 6
No of Pages: 5

Course Number& Title : CS F212 – Database Systems Marks : 80 Weightage : 40%

Duration: 3Hours Date: 23- 5-2018 Time: 12.30PM – 3.30PM Year: II year

Note: This paper consists of two parts. Answer Part A and Part B in separate booklets

PART A

1. List **any three** significant differences between a file processing system and a DBMS. (3M)

2. Consider a **B+ tree** of order 3 (i.e. 3 data elements and 4 pointers at each node)

a) Show the step by step construction of a **B+ tree** where the elements to be inserted into the tree are in the following sequence:

20, 23 ,86 ,09, 92 ,35 ,38, 01, 24 ,07 ,50, 53 ,06 ,39, 22.

b) Search for the element 86 in the tree as mentioned in a). i.e. You specify the nodes visited until you get the element 86.

c) Delete the element 39 from the tree as mentioned in a) and redraw the final tree.
(4 +1+2 =7M)

3. Create an **inverted index** for the following document collection:(assume that each document contains six words as given below) (5M)

Doc 1	MERCURY VENUS EARTH MARS JUPITER MERCURY
Doc 2	URANUS NEPTUNE PLUTO SATURN MERCURY SATURN
Doc 3	VENUS EARTH MARS JUPITER URANUS VENUS
Doc 4	NEPTUNE PLUTO SATURN MERCURY VENUS SATURN
Doc 5	EARTH MARS JUPITER URANUS NEPTUNE MARS
Doc 6	SATURN MERCURY VENUS EARTH MARS EARTH
Doc 7	EARTH MARS SATURN MERCURY VENUS SATURN

PTO

4. Draw an **E-R Diagram** and **Covert** it into **Relational Tables** for the following problem:

VENUS PLAYSCHOOL is a child day-care centre. A parent registers their child or children at the school using a registration form. A parent can submit more than one registration form. Each room in the day-care is assigned an age group. For example an infant is under 1 year of age and toddlers are from 1 to 3 years of age. A child is assigned to a room based on their age and availability of space. A room may be assigned one or more employees. An employee can only be assigned to one room. The minimum number of employees required for a room is determined by the number of children assigned to the room and the child : staff ratio identified by the government. For example one employee can care for 5 infants or 8 toddlers. (7M)

5. Consider the following relational schema:

author (<u>authorid</u> , authorname)	has 4000 records
magazine (<u>magazineid</u> , magazinename)	has 1000 records
writearticle (<u>magazineid</u> , <u>authorid</u> , no_of_articles_written)	has 40000 records

The following query is required over the above schema:-

“Find the names of all Authors who have written more than 20 articles in the magazine ‘LINUX MAGAZINE’.”

a) Write the **optimized query** in R.A (Relational Algebra) using appropriate equivalence rule(s).

b) **draw the initial expression tree** in R.A.

c) **draw transformed expression tree** (in R.A.). (2+2+2M)

6. Diagrammatically give the organization of Extensible Hashing with an example .Consider arbitrary four bit binary values as the data elements to be inserted. Show till the bucket size expands to three bits and consider the most significant n bits for distribution of data elements to the bucket. (6M)

7. The Main Hard Disk of a computer system has the following characteristics:

- There are ten surfaces, with 20,000 tracks each. Data is recorded on all surfaces.
- Tracks hold an average of 800 sectors of 512 bytes each.
- 10% of each track is used for gaps.
- The disk rotates at 9600 rpm.
- The time it takes the head to move n tracks is $1 + 0.006n$ milliseconds.

Answer the following questions about the disk.

a) What is the capacity of the disk in GB?

b) What is the maximum seek time in milliseconds?

c) What is the maximum rotational latency in milliseconds? (2+2+2M)

8. A Block header in a disk describes the data elements in a disk block. What are the contents of a block header? (2M)

PART B

1. Consider the following relational schema and set of functional dependencies. List all superkey(s) for this relation. Which of these superkeys form a key (i.e. a minimal superkey) for this relation? Justify your answer in terms of functional dependencies and closures.
 $R(A, B, C, D, E)$ with functional dependencies $AB \rightarrow E$ and $D \rightarrow C$.

a. Answer (Find all the superkeys and keys)

b. Decompose R into BCNF. Show your work for partial credit. Your answer should consist of a list of table names and attributes and an indication of the keys in each table (underlined attributes). (3+3M)

2. Consider the following relation:

$TRIP(trip_id, sta$

2. Consider the following relation:

$TRIP(trip_id, start_date, cities_visited, cards_used)$

This relation refers to business trips made by salesmen in a company. Suppose the trip has a single $start_date$ but involves many cities and one may use multiple credit cards for that trip. Make up a mock-up population of the table.

a. Discuss what FDs and / or MVDs exist in this relation.

b. Show how you will go about normalizing it.

(2+2M)

3.a. Consider the following relations A, B, C. Show the working of the evaluation of this expression. How many tuples does the result of the following relational algebra expression contain? Assume that the schema of $A \cup B$ is the same as that of A.

$(A \cup B) \bowtie_{A.Id > 40 \vee C.Id < 15} C$

Table A

Id	Name	Age
12	Arun	60
15	Shreya	24
99	Rohit	11

Table B

Id	Name	Age
15	Shreya	24
25	Hari	40
98	Rohit	20
99	Rohit	11

Table C

Id	Phone	Area
10	2200	02
99	2100	01

P.T.O

b) Consider the above tables A, B and C. How many tuples does the result of the following SQL query contains? Explain

```
SELECT A.id
FROM   A
WHERE  A.age > ALL (SELECT B.age
                    FROM   B
                    WHERE  B.name = "arun")
```

c) Discuss two scenarios where a view can be helpful in a database management system.

d) Explain why it is not always possible to perform SQL UPDATE/DELETE/INSERT statements on top of a view. (3+3+2+2M)

4. a. Consider the following transactions with data items P and Q initialized to zero:

```
T1: read (P) ;
    read (Q) ;
    if P = 0 then Q := Q + 1 ;
    write (Q) ;
T2: read (Q) ;
    read (P) ;
    if Q = 0 then P := P + 1 ;
    write (P) ;
```

Explain with an explanation if any non-serial interleaving of T1 and T2 for concurrent execution leads to

- (A) A serializable schedule
- (B) A schedule that is not conflict serializable
- (C) A conflict serializable schedule

b. Which of the following concurrency control protocols ensure both conflict serializability and freedom from deadlock? (A) 2-phase locking (B) Time-stamp ordering. Explain your answer

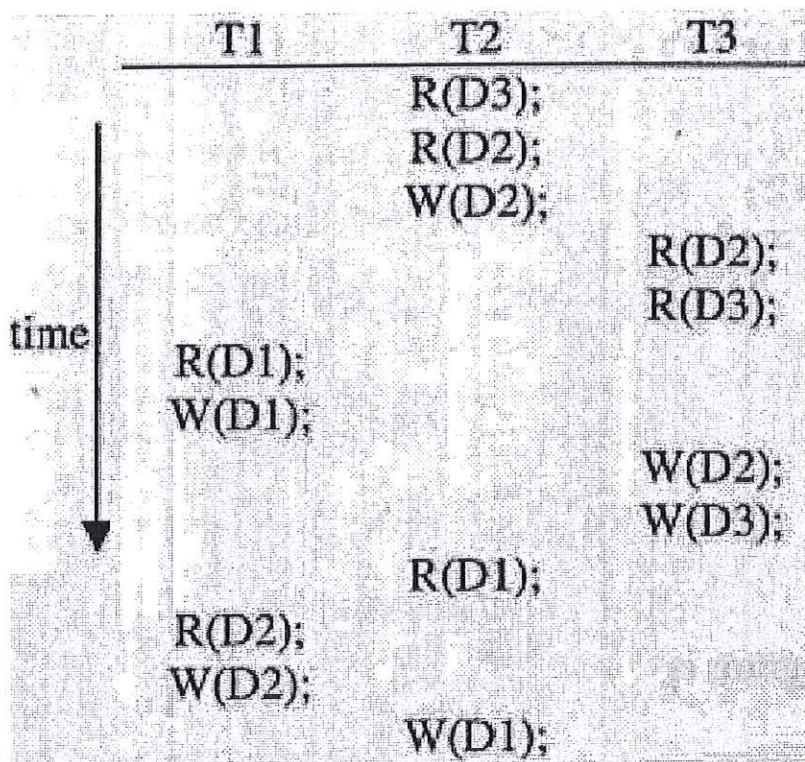
c. Consider the given situation

T ₃	T ₄
lock-X(B) read(B) B:=B-50 write(B) lock-X(A)	 lock-S(A) read(A) lock-S(B)

What is this condition called? How can it be resolved?

(2+2+2M)

5. a. Consider the given schedule, is it a serializable schedule? If so give the order. Use a precedence graph to justify your answer



b. Consider the following log sequence of two transactions on a bank account, with initial balance 12000, that transfer 2000 to a mortgage payment and then apply a 5% interest.

1. T1 start
2. T1 B old=12000 new=10000
3. T1 M old=0 new=2000
4. T1 commit
5. T2 start
6. T2 B old=10000 new=10500
7. T2 commit

Suppose the database system crashes just before log record 7 is written. When the system is restarted, what are the different methods of retrieving the database to a consistent state? Explain with this example, Till what point does the recovery have to be done? (4+4M)

6.a. Consider the following transaction involving two bank accounts x and y.

```
read(x); x := x - 50; write(x); read(y); y := y + 50; write(y)
```

The constraint that the sum of the accounts x and y should remain constant is which of the following constraints of ACID i.e. (Atomic, Consistent, Isolation, Durable)? Explain

b. Explain what is meant by a multiple granularity locking protocol with an example.

(2+2M)

*****ALL THE BEST*****