

BITS PILANI, DUBAI CAMPUS

Dubai International Academic City

Second Semester 2013 – 2014

Comprehensive Exam

Year : III/ II

Date : 01-6-14

Course No : CS C352 / CS F212

MAX Marks: 80 (40%)

Course Title : Database Systems

Duration : 3Hours

Note: Answer all nine questions. Clearly explain any calculations done.

1.a. Consider a sample database. Underlined attributes indicate primary key.

Car(RegNo, Make, ModelYear, Color)

Inspection(RegNo, Car, DateInspected, Period, Evaluation),

Problem(RegNo, DateInspected, Inspection, ProblemCode)

Driver(RegNo, Car, Name, Accidents)

Answer the following queries using **relational algebra expressions**.

- Display information about cars of year 1996 model, where faults have been found in the inspection for year 1999. 2M
- Display driver's name for the model year 1995 or older cars that have not been inspected for year 2000. 2M
- Display what makes have provided similarly colored cars as models 1999 and 2000? 2M

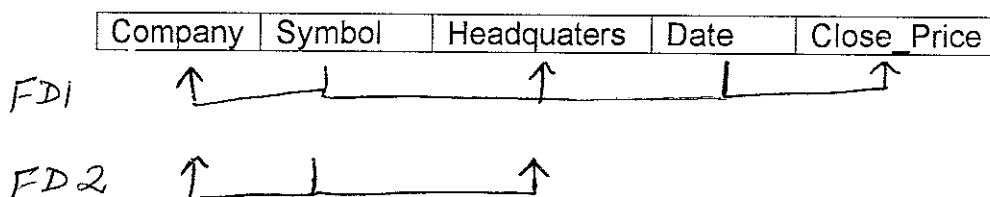
b. Write SQL statements for the following

- Have a constraint to ensure that the ModelYear of a car is greater than 1985. 2M
- Write a trigger to count all entries made for cars which are manufactured in the year 2005 when inserting records into the car table. 2M

2. a. What is the need to normalize tables in the Relational model? What would happen if you leave un normalized tables? 1M

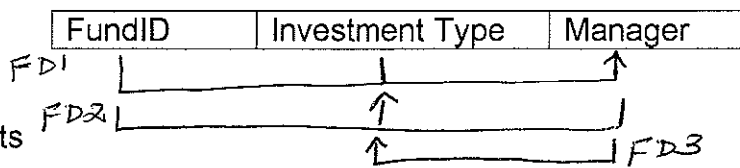
b. Consider the following relational tables. Which is the highest Normal form each table satisfies? Explain why they violate a specified normal form and bring them to the highest normal form. The functional dependencies are indicated.

- Stocks 2M



ii) Funds

2M



iii) Students

2M

StudID	Major	Activities
--------	-------	------------

In the given tables StudID → Major and StudID → Activities

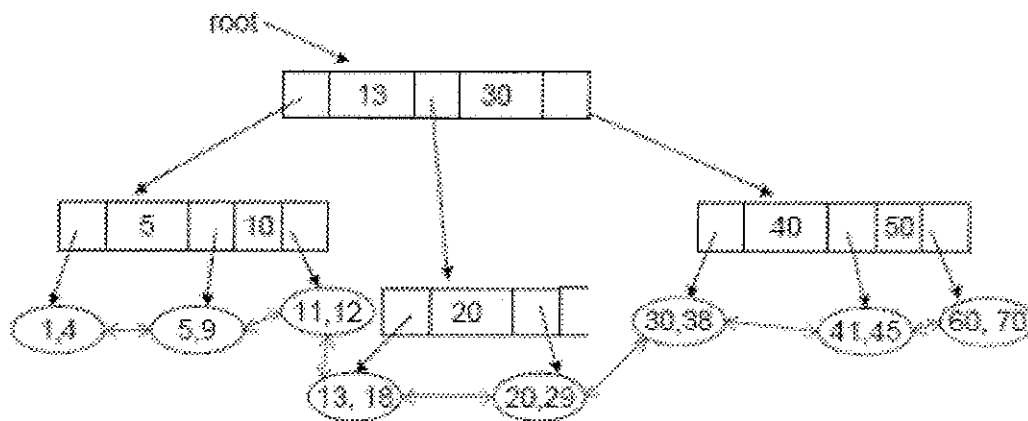
3. Consider a disk with a sector size of 512 bytes, 2000 tracks per surface, 50 sectors per track, five double sided platters, and average seek time of 10msec.

- What is the capacity of the track in bytes? What is the capacity of each surface? What is the capacity of the disk? 3M
- How many cylinders does the disk have? 1M
- Give examples of valid block sizes. Is 256 bytes a valid block size? how about 2048? 51200? 1M
- If the disk platters rotate at 5400rpm(revolutions per minute) what is the maximum rotational delay? 2M
- If one track of data can be transferred per revolution, what is the transfer rate? 2M

4.a. Show the effect of constructing an index when the data elements are 5,10,15,15,20,25,35,35,40,45,50,55,60,60,70. We can assume that 3 index records are put on a block and 5 data records are put on a block.

- Create a sparse index 3M
- Create two levels of indexing. 3M

b. Consider the given B+ tree



Show the following actions. Consider each action to follow the other i.e. ii) will work on the tree obtained by i) and likewise.

- i) Insert element 19 to the B+ tree 2M
- ii) Insert element 27 to the B+ tree 2M
- iii) Delete element 5 3M

5. a. Draw two different query evaluation plans for this given SQL Statement.

SELECT name FROM ships, classes AS C1
 WHERE ships.class = C1.class
 AND numGuns = (SELECT MAX(numGuns)
 FROM classes AS C2
 WHERE C1.bore = C2.bore);

2+2M

b. c. Model the following requirements as an ER / EER diagram and convert the same to the relational Model. You can consider the relevant attributes for each class.

i) All varieties of fruits and vegetables have certain attributes like color, shape, size, taste, all fruits and vegetables grow in nature. 2M

ii) A taxi driver could work independently or work for an organization. His pay is different in these two roles. He could own a vehicle or drive around a rented vehicle. Any driving fines are recorded against the vehicle he drives. 3M

6.a. Suppose you have a table called HR.EMPLOYEES which stores information about the department and the job id of employees .You are required to create bit map indexes for the appropriate tables in order to answer the following queries.

HR.EMPLOYEES (dept_id, job_id, job description,years of service)

<u>dept id</u>	<u>job id</u>	<u>job description</u>	<u>years of service</u>
10	MGR	XXX	10
15	DBA	YYY	12
20	Developer	AAA	3
30	Developer	BBB	2

1. SELECT * FROM HR.EMPLOYEES WHERE DEPARTMENT_ID=20 AND JOB_ID='MGR'; 2M
2. SELECT * FROM HR.EMPLOYEES WHERE DEPARTMENT_ID=20 OR JOB_ID='Developer'; 2M
3. SELECT * FROM HR.EMPLOYEES WHERE JOB_ID='DBA' AND DEPARTMENT_ID<>20 ; 2M

b. Give the run length code for the following data

a. 100000001000

b. 000100001001

4M

7. a. Consider the following schedule of transactions T1, T2 and T3.

R1(A);R2(B);R3(C);W1(B);W2(C);W3(D)

R1(A);R2(B);R3(C);R1(B);R2(C);R3(D);W1(C);W2(D);W3(E)

- i) Show how each of these transactions would work if they had a scheduler which supports shared and exclusive locks. 2M
- ii) Show how each of these transactions would work if they had a scheduler which supports shared , exclusive and update locks. 2M
- b. What is the difference between 2PL and strict2PL ? 2M
- c. Give the compatibility matrix for shared, exclusive and increment locks. 3M

8.a. What is the main objective of a scheduler which executes concurrent jobs, why cannot all jobs be executed in a sequential manner ? 1M

b. What do you mean by a conflict in a schedule? List the three conditions which can lead to a conflict. 3M

c. Given the following schedule, justify if this schedule is conflict equivalent to a serial schedule.

r1(A); r2(A);w1(B);r2(B);w2(B);r2(B);r1(A);w1(A);w1(B)

perform the above operation by i) Considering a series of swaps of the actions ii) By using the precedence graph. 2+2M

9. a. How would you define a Data Warehouse ? 1M
b. What are the 3 V's associated with Big Data ? 1M
c. Where is NOSQL used ? What are the classifications ? 1M
d. Toyota has a number of vehicles of different makes e.g. SUV, coupe, luxury car, sedan etc. under each car it has different models. Represent this form of representation using a Data Model. 2M

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

BITS PILANI, DUBAI CAMPUS

Dubai International Academic City

Second Semester 2013 – 2014

Test – 2 (Open Book)

Year : III/ II

Date : 29-4-14

Course No : CS C352 / CS F212

MAX Marks: 40 (20%)

Course Title : Database Systems

Duration : 50 minutes

Note: Answer all four questions. Only prescribed text book and handwritten class notes are permitted for this component. **Clearly explain any calculations done.**

1. Suppose we are designing a database relation named Enrollment to store student enrollments in different courses during a semester. The table stores basic information about the students and the list of courses in which the student has enrolled. For simplicity, let us assume that a student cannot be enrolled in more than 5 courses in a semester. Here is the description of database column fields:

- RecordHeader CHAR(10)
- StudentID CHAR(9)
- Name CHAR(30)
- Gender CHAR(1)
- Variable number of course names (max 5) each of type CHAR(30)

Let the size of a block be 4096 bytes out of which 96 bytes are used for the block header. Records are not spanned.

(a) Assuming that we used fixed-length records to store the tuples,

i) what is the length (in bytes) of each record 2.5M

ii) what is the number of such tuples that can be stored per block. 2.5M

(b) Using fixed-length records to store tuples is definitely going to result in space wastage. To be more space efficient, we are now going to use variable-length records to store the repeating course fields. We will still use 30 bytes to store each course name. The other fields are still of fixed length. However, we now add one byte to the record header to store the record length (in bytes).

i) Assuming that a student has 3 courses on the average, how many records can we store in a block, on average? Of course, some blocks will hold fewer than this average number of records. 2.5M

ii) In the worst case, what is the smallest number of records that a full block may be holding? 2.5M

P.T.O

2. Consider a Hard Disk with the following specifications:

- Disk does one full revolution in 512 μ sec.
- Usable capacity: 5 Gigabytes (i.e., 5×2^{30} bytes)
- Number of cylinders: 64 + 128 (explained below)
- 1 block = 4 Kilobytes (i.e., 4×2^{10})
- negligible overhead between blocks (gaps)
- Average seek time: 10,000 μ sec
- Transfer time: 1 μ sec

There are 64 inner cylinders and 128 outer cylinders. The outer cylinder has double the density than the inner ones, that is, an outer cylinder has twice the number of blocks than an inner cylinder.

Hint: To compute your answers below, work with powers of 2, and you can leave your answer as a power of 2 if appropriate. For example, if the answer is 3×2^{20} divided by 2^5 , just leave your answer as 3×2^{15} .

- What is the total number of blocks on the disk?
- How many of the blocks are on the inner cylinders?
- How many blocks are on the outer cylinders?
- Once the head arrives at the beginning of an inner block, how much time does it take to transfer a block off the disk?
- What is the expected average time to read a block that resides on an inner cylinder when the disk head can currently be at any location? $2 \times 5 = 10M$

3. a. Construct a B+ tree with data values 1, 4, 7, 10, 17, 21, 31, 25, 19, 20, 28, 42 where each of the non – root nodes can store 3 data elements and 4 pointers. Show clearly the step by step procedure for tree construction.
- b. From the above given tree indicate the effect of deleting data element 25.
- c. Indicate how the search of all data elements between 7 to 20 proceeds in this tree. $(6 + 2 + 2M)$

4. Suppose the hash function is $h(x) = x \bmod 8$ and each bucket can hold at most two records. Show the extendable hash structure after inserting 1, 4, 5, 7, 8, 2, 20. Each of these elements can be represented using 3 bits. Show clearly the step by step procedure. (10M)

Extendable Hashing

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

BITS PILANI, DUBAI CAMPUS

Dubai International Academic City

Second Semester 2013 – 2014

Test – 1 (Closed Book)

Year : III/ II

Course No : CS C352 / CS F212

Course Title : Database Systems

Date : 11-3-14

MAX Marks: 40 (20%)

Duration : 50 minutes

Note : Answer all questions

1. UPS prides itself on having up-to-date information on the processing and current location of each shipped item. To do this, UPS relies on a company-wide information system. Shipped items are the heart of the UPS product tracking information system. Shipped items can be characterized by item number (unique), weight, dimensions, insurance amount, destination, and final delivery date. Shipped items are received into the UPS system at a single retail center. Retail centers are characterized by their type, uniqueID, and address. Shipped items make their way to their destination via one or more standard UPS transportation events (i.e., flights, truck deliveries). These transportation events are characterized by a unique scheduleNumber, a type (e.g, flight, truck), and a deliveryRoute.

- Draw an Entity Relationship diagram that captures this information about the UPS system. Be certain to indicate primary key and cardinality constraints.
- Map this ER diagram to the relational model , clearly showing all foreign key constraints. [6 + 4M]

2. Given a relation $R = (A, B, C, D, E)$ with the following functional dependencies:
 $\{CE \rightarrow D, D \rightarrow B, C \rightarrow A\}$

- Determine $\{C, E\}^+$ and $\{D\}^+$
- Find all candidate keys of R.
- Identify the best normal form that R satisfies (1NF, 2NF, 3NF, or BCNF) and justify.
- If the relation is not in BCNF, decompose it until it becomes BCNF. At each step, identify a new relation, decompose and re-compute the keys and the normal forms they satisfy. [4+2+2+4M]

P.T.O

3. Given relation R(A,B,C,D,E) the instance of this relation is shown in Fig.1 below.

a. State whether the following FD's hold

- i) $A \rightarrow C$ ii) $AC \rightarrow E$ iii) $ABC \rightarrow E$

b. If this relation R is decomposed into two relations R1(A,B,C,D) and R2(A,C,E).

Check whether this decomposition is lossless. (i.e. when the two tables are joined together do you get the same original relation) [3 + 5M]

A	B	C	D	E
'a'	122	1	's1'	'a'
'e'	236	4	'e2'	'b'
'a'	199	1	'b5'	'a'
'b'	213	2	'z8'	'd'

Fig. 1

4. You are given the following relational schema for students.

Student(SID, Sname, Curriculum)

Takes (CourseID, SID, Semester, Grade) .

Write the following queries in relational algebra.

a. Find the id of all students who took the same course in two different semesters.

b. Find the name of all students who never got an F. [5+5M]

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

BITS PILANI, DUBAI CAMPUS

Dubai International Academic City

Second Semester 2013 – 2014

Quiz – 1 (Closed Book)

Year : III/ II

Course No : CS C352 / CS F212

Course Title : Database Systems

Date : 27-3-14

MAX Marks: 10 (5%)

Duration : 20 minutes

Note : Answer all questions

1. Consider a database which consists of the following tables. The relation of the different tables and their attributes is shown in the diagram below. Note that underlined attributes denote primary key.

EMPLOYEE

<u>Fname</u>	<u>Minit</u>	<u>Lname</u>	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	Dno
--------------	--------------	--------------	------------	-------	---------	-----	--------	-----------	-----

DEPARTMENT

<u>Dname</u>	<u>Dnumber</u>	Mgr_ssn	Mgr_start date
--------------	----------------	---------	----------------

DEPT_LOCATIONS

<u>Dnumber</u>	<u>Dlocation</u>
----------------	------------------

PROJECT

<u>Pname</u>	<u>Pnumber</u>	<u>Plocation</u>	<u>Dnum</u>
--------------	----------------	------------------	-------------

WORKS_ON

<u>Essn</u>	<u>Pno</u>	Hours
-------------	------------	-------

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
-------------	-----------------------	-----	-------	--------------

Figure 3.7

Referential integrity constraints displayed on the COMPANY relational database schema.

Write **SQL** statements for the following queries.

1. Retrieve the names of Employees who have no dependents.

1.5M

2. Find the maximum, minimum and average salary of employees in the "Research department" 1.5M

3. For each project on which more than two employees work, retrieve the project number, project name and the number of employees who work on that project. 2.5M

4. For each department that has more than five employees, retrieve the department number and the number of employees who are making more than \$40,000 2.5M

5. Give the instruction to create the table WORKS_ON and ensure that no employee works for more than 100 hours on a project. 2M