

BITS, PILANI – DUBAI
International Academic City, Dubai
Second Semester 2009 – 2010
Database Systems CS C352 (III year)
Comprehensive Exam

Duration : 3 Hours
24.5.10

Weightage : 40%
Max Marks: 80

Note: Answer all questions, this question paper contains three printed pages.

1. With reference to distributed database systems answer the following questions.
 - a. Explain the need for a two phase commit protocol and explain the two phases.
 - b. Describe the three data fragmentation strategies and give examples for each
 - c. What is data replication, give the different data replication strategies and mention the problems of data replication.
(4 + 3 + 3 M)

2. With reference to XML answer the following queries.
 - a. What does DTD stand for ? Explain the significance of using DTD to define XML documents.
 - b. Explain clearly the difference between an HTML document to an XML document.
 - c. Given the following DTD document construct the corresponding XML document. (3+3+4 M)

```
<!DOCTYPE note
[
<!ELEMENT note (to,from,heading,body)>
<!ELEMENT to (#PCDATA)>
<!ELEMENT from (#PCDATA)>
<!ELEMENT heading (#PCDATA)>
<!ELEMENT body (#PCDATA)>
]>
```

3. With reference to query execution answer the following queries.
 - a. Given two identical tables A and B write the SQL statement to find all the elements that are in table A but not in table B and all the elements that are in table B but not in table A.
 - b. Express the above query as a query execution plan, is it possible to have an alternate query execution plan, if so show the representation.

- c. Represent diagrammatically the different plans that can be generated by a typical optimizer to represent the different plans that can be generated to represent the join operation of four tables A,B,C,D. (3+4+3M)
4. With reference to indexes answer the following queries
- Explain clearly how a search engine like Google makes use of inverted indexes for document retrieval. What are stopping words and stemming words.
 - Diagrammatically give the organization of a i) dense index ii) sparse index iii) multiple levels of index. (4 + 6M)
5. With reference to concurrency control answer the following queries.
- In a database system why is concurrency control important, what is the technique followed to implement concurrency control ?
 - What is the significance of the lock table? Explain its organization.
 - What are increment locks ? how are they used ? give the compatibility matrix for an increment lock. (3+ 3 + 4M)
6. a. Consider the given relation schema Sales(sales_transaction_no, item_no, item_price, item_qty_sold, seller, seller_district) and the functional dependencies are shown below. What is the key of this relation ? Transform this relation to 3NF.
- sales_transaction_no, Item_no → item_qty_sold
 item_no → item_price
 sales_transaction_no → seller
 seller → seller_district
- b. How is 3NF different from BCNF ? Give an example to illustrate the difference clearly. (6 + 4M)
7. A General Hospital consists of a number of specialized wards (such as Maternity, Paediatrics, Oncology, etc). Each ward hosts a number of patients, who were admitted on the recommendation of their own GP and confirmed by a consultant employed by the Hospital. On admission, the personal details of every patient are recorded. A separate register is to be held to store the information of the tests undertaken and the results of a prescribed treatment. A number of tests may be conducted for each patient. Each patient is assigned to one leading consultant but may be examined by another doctor, if required. Doctors are specialists in some branch of medicine and may be leading consultants for a number of patients, not necessarily from the same ward.
- represent the above specifications as an ER diagram.

- b. convert the ER diagram to the relational model and show the attributes of each table.
- c. Give the SQL syntax to create the patients table. (3 + 4 + 3M)

8. Consider the given relational tables

Employee(empid, name)

Project(projid, name)

WorkLoad(empid, projid, duration)

Answer the following queries in SQL. Answer query iv) in both SQL and relational algebra.

- i) List the number of projects that each employee (empid) is working on.
- ii) List, in descending order, employee names that are working on the project called "Databases".
- iii) List the employee (empid) who spent the longest duration working on a project.
- iv) List the employees (name) who have not worked on any project.

(2 X 5 =10M)

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

BITS, PILANI – DUBAI
International Academic City, Dubai
Second Semester 2009 – 2010
Database Systems CS C352 (III year)
Test – 2 (Open Book)

Duration : 50 minutes
25.4.10

Weightage : 20%
MAX MARKS: 40

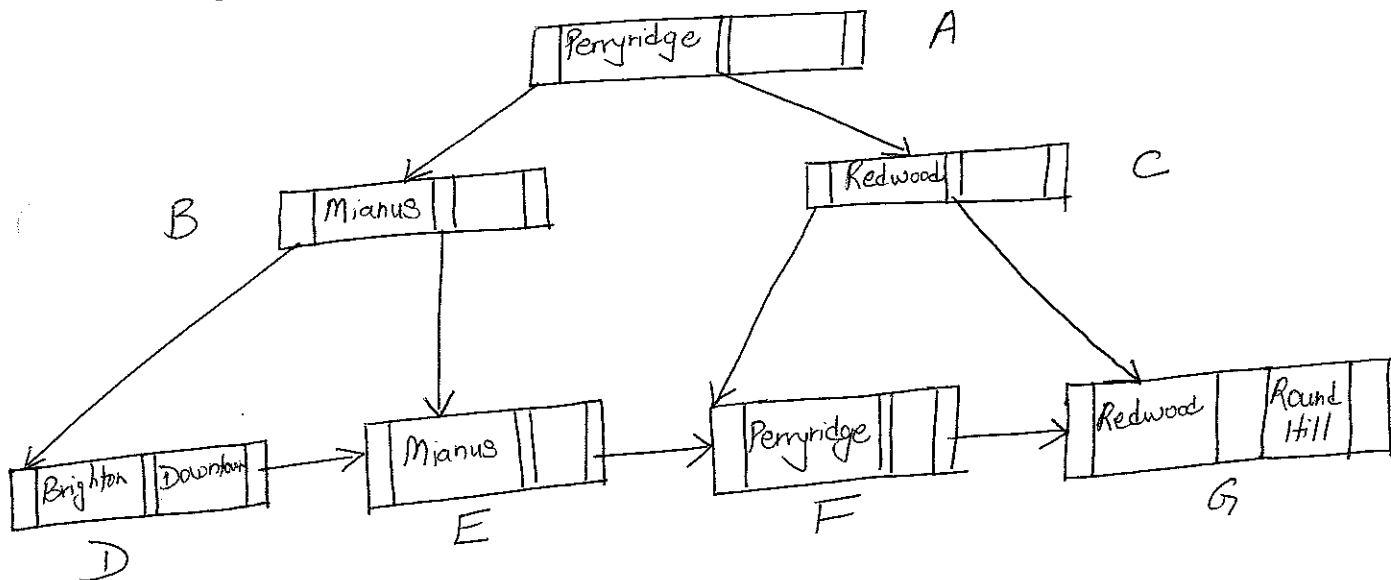
Note: Only prescribed text book and handwritten class notes are allowed

1. Consider a record having the following fields in order: A character string of length 15, an integer of 2 bytes, an SQL date and an SQL time. How many bytes does the record take if:

- Fields start at any byte
- Fields must start at a byte that is a multiple of 4.
- Fields must start at a byte that is a multiple of 8.

If we wish to pack as many records as we can into a block of 4096 bytes using a block header that consists of ten integers each integer requiring 4-bytes. How many records can we fit into a block in each of the cases a, b, and c, specified above. 9M

2. If multiple granularity locking policy is followed in the below given tree, clearly indicate and justify the type of lock obtained and whether the lock is held or given up in the next stage/level when each of the following happens: 6M



- Details of Perryridge are read.
- When data Astonville is inserted into the tree.
- When data Redwood is deleted

3. Show the creation of an index using extensible hashing when data records with the following search key values 2,3,5,7,11,17,19,23,29,31 are added to the table. The hashing function used is $h(x) = x \bmod 8$ and the bucket can hold three records. 10M

4. Consider the following two transactions:

T1: read(A);
 read(B);
 if A = 0 then B = B + 1;
 write(B);

T2: read(B);
 read(A);
 if B = 0 then A = A + 1;
 write(A);

Write lock and unlock instructions for transactions T1 and T2, so that they observe the two phase locking protocol. Give an example where the execution of these instructions leads to a deadlock. Consider the use of shared and exclusive locks. 3 + 2 M

5. For each of the given schedules state whether the schedule is serializable, conflict-serializable or both, justify your answer.

a. T1:R(X), T2:R(X), T1:W(X), T2:W(X)

b. T1:W(X), T2:R(Y), T1:R(Y), T2:R(X) 2+2M

6. For the sequence of actions given below, assume that shared locks are requested before every read action, and exclusive locks are requested immediately before every write action. Also, unlocks occur immediately after the final action that a transaction executes. Clearly indicate which actions are denied and whether deadlocks occur. Draw the waits for graph and indicate if a deadlock occurs. Which action would you pick in the event of a deadlock and indicate how the sequence of actions proceed.

$r_1(A)$; $r_2(B)$; $w_1(C)$; $r_3(D)$; $r_4(E)$; $w_3(B)$; $w_2(C)$; $w_4(A)$; $w_1(D)$; 6M

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

BITS PILANI – DUBAI
International Academic City, Dubai
Second Semester 2009 – 2010
Database Systems CS C352 (III year)
Test – 1 (Closed Book)

Duration : 50 minutes
14.3.10

Weightage : 25%
MAX : 50 Marks

1. Use an Entity-Relationship Diagram to depict the following requirements for a restaurant:
 - The restaurant employs a number of chefs. A record is kept of each chef's name, address, phone number and salary.
 - Each chef can prepare a number of meals. The name of the meal and the price of the meal is recorded.
 - Each meal consists of a number of ingredients. The name of the ingredient and the quantity required for that particular meal is recorded.
 - These meals are ordered by customers. A record is kept of the customers name, address and phone number. A record is kept of the time and date the meal is ordered. State any assumptions made in the design of the E-R diagram. 8M
2. Given the following two tables clearly indicate how you would determine if a customer could afford a particular item(an item is affordable only if the price is less than the credit of the customer) . 5 X 2 = 10M
 - a. mention which relational operator is used to answer the query and why.
 - b. show the step wise contents of the table at each state of answering the query.

Product

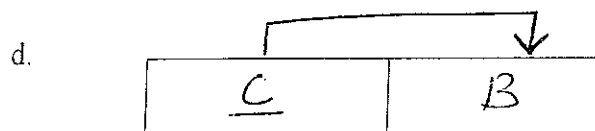
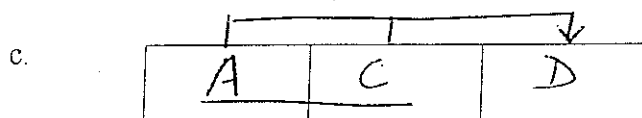
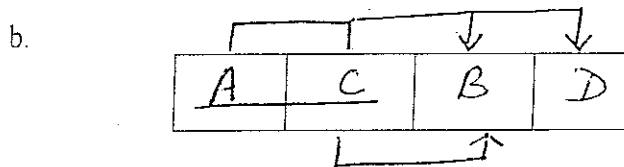
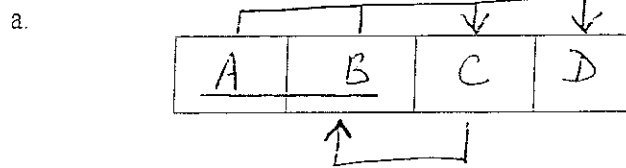
Product ID	Name	Price
123	Basketball	125
343	Bike	550
489	Golf clubs	980

Customer

Customer ID	Credit
C3920	1500
C563	350
C332	200

3. Consider the functional dependencies depicted in the figures below. Considering 1NF,2NF,3NF,BCNF mention which is the highest normal form each of these tables are in and clearly mention which normal form they violate and why ? 3 X 4 = 12M

P.T.O



4. Given the schema of the tables, use relational algebra to answer the following queries.

Car(RegNo, Make, ModelYear, Color)

Inspection(RegNo, DateInspected, Period, Evaluation),

Problem(RegNo, DateInspected, ProblemCode)

Driver(RegNo, Name, Accidents)

- Information about cars of year 1996 model, where faults have been found in the inspection for year 1999.
- Driver's name for the model year 1995 or older cars that have not been inspected for year 2000. $5 \times 2 = 10M$

5. a. Compare the usage of a file system to store data as compared to a database. What would be more advantageous and why?

- Explain the fourth normal form, what are its requirements and how do you bring table to the 4NF? $5 \times 2 = 10M$

***** GOOD LUCK *****

BITS PILANI – DUBAI
International Academic City, Dubai
Second Semester 2009 – 2010
Database Systems CS C352 (III year)
Quiz (Closed Book)

Duration : 20 minutes
30.3.10

Weightage : 8%
MAX : 16 Marks

SET B

Name : _____ **ID No:** _____

1. Give the time stamp of noon June 2nd 2007. 2M

2. BLOB stand for _____. 1M

3. Given the following tables

Flights (aid : integer, from: string, to: string, distance: integer, departs: time, arrives: time)

Aircraft (aid: integer, aname: string, cruisingrange: integer)

Certified (eid: integer, aid: integer)

Employees (eid: integer, ename: string, salary: integer, age : integer)

1. Answer the following queries in SQL In the above table cruising range indicates the maximum distance a flight can fly without refueling and all details in the employee table are regarding pilots.

a. Find the maximum salary of a pilot. 2M

b. Give the syntax for creating the table Employees with the condition that the age of the pilot should be in the range of 22-60 years(inclusive of 22 and 60). 3M

c. Find the *names* of pilots certified for some Boeing aircraft. 2M

d. Find the *aids* of all aircraft that can be used on non-stop flights from Delhi to New York. 3M

e. For each flight find the names of the pilots who are certified to fly it. 3M