BITS, PILANI – DUBAI
International Academic City, Dubai
Year III – Semester II   2008– 2009
Course No.: CS 352   Course Title: Database Systems

COMPREHENSIVE EXAM (Closed Book)

Date: May 24, 2009   Time: 3 Hours Max. Marks = 40 Weightage : 40%

Instructions : There are six questions in this paper answer all questions sequentially.

1. The table below lists customer/car-hire data. Each customer may hire cars from various outlets. A car is registered at a particular outlet and can be hired out to a customer on a given date.
a) The data in the table is subject to *update anomalies*. Provide examples of how *deletion, and modification* anomalies could occur on this table. [2 marks]
b) Identify *the functional dependencies* represented by the data in the table. State any assumptions you make about the data. [2 marks]
c) Using the functional dependencies identified, describe and illustrate the process of *normalization* by converting the table to Third Normal Form (3NF) relations. [2 marks]
d) Draw an *Entity-Relationship model* for the data in the table. Show all the entities, relationships, and attributes.[2 marks]

CarReg	Make	Model	CustNo	CustName	HireDate	OutletNo	OutletLoc
W 565	CDC	Ford Escort	C100	Smith.J.	14/5/01	21	Woodstock
W565	CDC	Ford Escort	C222	Patel V	15/5/01	21	Woodstock
V734	HSB	Nissan Sunny	C100	Smith.J.	14/5/01	21	Woodstock
W104	RSM	Ford Escort	C303	Brown.F.	14/5/01	24	Denham
W104	RSM	Ford Escort	C100	Smith.J.	16/5/01	24	Denham
W611	SBH	Nissan Sunny	C222	Patel.V	15/5/01	24	Denham

2. The following relations are part of a relational database.
Resort (**ResortID**, ResortName, Address)
Hotel (**HotelID, ResortID**, HotelName, Rating, Price)
Customer (**CustomerID**, CustomerName, Address)
Reservation (**ResortID, CustomerID, FromDate**, ToDate, HotelID)
The primary key in each table is shown in bold.
Write expressions in **BOTH relational algebra AND in SQL** to retrieve each of the following:
a) a list of the names of all the resorts.   [1 mark]
b) a list of all three star hotels that are under £60 per night. You may identify each hotel by its HotelID   [2 x 1.5 marks]
c) the names and addresses of all customers who have reservations at the Grand Hotel Slais, which has a HotelID of SLAIS.   [2 x 1.5 marks]
d) a list of all hotels, showing the reservations for the 5 June 2001. For each reservation show the CustomerID. You should include in your list hotels that have no reservations.   [2 x 1.5 marks]

Also write an expression in SQL to retrieve
e) a count of the hotels in the resort of Datamania. [1 mark]

3. Answer the following three parts.
a. Define the notion of serializability. Give an example illustrating why serializability is an important concept. [2 marks]
b. Give an example of a two-phase locking history and briefly explain the notion of two-phase locking. [2 marks]
c. What are the advantages and disadvantages of the two-phase locking protocol? [2 marks]

4. a. Suggest why XML is an important standard and how it can affect the evolution of the World-Wide-Web. [1 mark]
b. How would you compare the relational data model and the DTD used in XML? Use an example where appropriate. [2 marks]

5. Distributed database management systems (DDBMS) are more complex than centralized ones for a number of reasons. When a database is both *fragmented* and *replicated* by a DDBMS, query processing and updating are more complex activities than they are for a centralized system.
a) Explain the terms *fragmentation and replication.* [2 marks]
b) Explain through the use of an example how the choice of distributed query processing strategy can have a significant effect on the query response time. [2 marks]
c) State FOUR reasons why an organization might move from using centralized to distributed databases, and FOUR disadvantages of so doing. [2 marks]

6. a. How does the one pass algorithm work for a tuple at a time operation. [2 marks]
   b. What is meant by pinned records and how are they handled? [2 marks]
   c. Suppose in a B+-tree with three levels, every index node can have up to 200 children (half full means 100 children), and every leaf node can store up to 20 records. What is the maximum and minimum number of records that the tree can hold? For an insertion (assume the record did not exist in the tree beforehand), what is the maximum and minimum number of logical I/O requests to the buffer manager [2 marks]

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

Course No.: CS C352   TEST II (Open Book)   Course Title: Database Systems

Date: April 12, 2009 Time: 50 Minutes   Max. Marks = 20

1. Consider the following relational schema. An employee can work in more than one department; the pct time field of the Works relation shows the percentage of time that a given employee works in a given department.

Emp(<u>eid: integer</u>, ename: string, age: integer, salary: real)
Works(<u>eid: integer, did: integer</u>, pct time: integer)
Dept(<u>did: integer</u>, budget: real, managerid: integer)

Write SQL integrity constraints (domain, key, foreign key, or CHECK constraints; or assertions) or SQL triggers to ensure each of the following requirements, considered independently.

a. Employees must make a minimum salary of $1000.
b. A manager must always have a higher salary than any employee that he or she manages.
c. The total percentage of time an employee works in must be under 100%.
                (2 + 2 + 3)

2. Why are NULL values used in a database ? How are NULL values handled during a join operation. Explain clearly with an example.   2M

3. Consider the relation shown in the figure

X	Y	Z
X1	Y1	Z1
X1	Y1	Z2
X2	Y1	Z1
X2	Y1	Z3

a.   List all the functional dependencies that this relation instance satisfies.

b. Assume that the value of attribute Z of the last record in the relation is changed from $z3$ to $z2$. Now list all the functional dependencies that this relation instance satisfies.　　　3M

4. Suppose you are given a relation R with four attributes $ABCD$. For each of the following sets of FDs, assuming those are the only dependencies that hold for R, do the following: (a) Identify the candidate key(s) for R. (b) Identify the best normal form that R satisfies (1NF, 2NF, 3NF, or BCNF).
1. $C \rightarrow D, C \rightarrow A, B \rightarrow C$
2. $B \rightarrow C, D \rightarrow A$
3. $ABC \rightarrow D, D \rightarrow A$　　　　　　　4.5M

5. Given that a block can store 4096 bytes of data , and that the header information of a block requires 20 bytes, calculate the number of blocks required to store 30 records each of the same size of 512 bytes. Consider cases of a. Storing records completely in one block. b. Storing records over different blocks
　　　　　　　3.5M

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

Course No.: CS C352   TEST I (Closed Book)   Course Title: Database Systems

Date: March 01, 2009   Time: 50 Minutes   Max. Marks = 25

1.a. Define what is a transaction. Why are transaction managers needed in database systems.?   3M

b. Define the following terms clearly with reference to a student.
   i) domain
   ii) primary key
   iii) composite attribute
   iV) multivalued attribute   4M

2. The Motor Vehicle Branch administers driving tests and issues driver's licenses. Any person who wants a driver's license must first take a learner's exam at any Motor Vehicle Branch in the province. If he/she fails the exam, he can take the exam again any time after a week of the failed exam date, at any branch. If he passes the exam, he is issued a license (type = learner's) with a unique license number. A learner's license may contain a single restriction on it. The person may take his driver's exam at any branch any time before the learner's license expiry date (which is usually set at six months after the license issue date). If he passes the exam, the branch issues him a driver's license. A driver's license must also record if the driver has completed driver's education, for insurance purposes.

Create a E-R diagram following these steps.
a. Find out the entities in the specification.
b. Find out the relationships among the entities.
c. Find out attributes of the entities and (if any) of the relationships.
d. Find out constraints between entities and relationships.
e. Check to see if you don't miss anything in specification.   6M

3. a) What do you mean by an aggregation explain clearly with an example.
   b) Define the terms owner entity type, week entity type, identifying relationship type and partial key. Explain clearly all the constraints associated with a week entity taking an example of an employee and his dependents.   3 + 3M

(PTO)

4. Relational schemas for five relations in a movie database are depicted below.

    star(starName, age)
    studio(studioName, where)
    produces(studioName, movieName)
    starsIn(starName, movieName)
    movie(movieName, whenMade)

Answer the following queries for the given relational schema.

    a) Give the names of stars who do not appear the the movie Haunted .
    b) Give the names of all the movies starring Dustin Hoffman.
    c) Give the names of people who star in atleast 2 movies.   3X2=6M

Course No : CS C352   **Course Title : Database Systems**
Date : 5/5/09   **Weightage : 5%**

1. The data to be stored in memory is given   10,15,20,25,25,25,30,35,35,40,45
   create a sparse index for this data.
     a. Show the changes to the tables when data 23 followed by 20 is added.
     b. When data 35, 45 is deleted
   Assume that the index block can store 5 records and that the data blocks can store
   3 records. Also note that overflow blocks are not allowed.   3M

2. Discuss the use of a secondary index. Create a secondary index for the records
   5,5,10,20,30,10,17,19,20,30,33 which are stored in this specified order.   2M

3. Create a linear hash table to store the data 15,17,12,6,3,4 ,11,8 show the step by
   step creation of the tables as each data element is added. Assume that two records
   can be stored in each block.   3M

4. What types of queries are B trees best suited to answer? explain why ?   1M

Course No : CS C352

Date : 30/3/09

**Course Title : Database Systems**

Weightage : 5%

Name : _____   ID NO: _____

1. Write an assertion to specify the condition that the salary of an employee must not be greater than the salary of the manager of the department that the employee works for in the given relational schema.

EMPLOYEE (Name, SSN, Address, DNO, SuperSSN)

DEPARTMENT (DName, DNO, MgrSSN)   3M

2. What is a candidate key ? Explain clearly with an example. Consider a relation R (A, B, C, D, E) with the following dependencies AB -> C, CD -> E, DE -> B. Identify the candidate key of the relation? Explain your answer.   2M

3. Which normal form is this relation in, and in which form is it not in, justify your answer with valid explanations and definitions.   2M
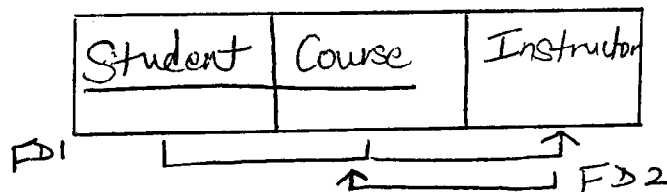
4. Consider the universal relation R = { A, B, C, D, E, F, G, H, I, J } and the set of functional dependencies G = { {A, B} -> {C }, {B, D} -> {E, F} , {A,D} -> {G, H}, {A} -> {I}, {H} -> {J}}. What is the key of G ? Decompose G into 2NF and than 3NF relations.   3M

Course No : CS C352   **Course Title : Database Systems**

Date : 10/3/09   **Weightage : 5%**

Name : _____   ID NO: _____   _____

Consider the following tables

EMPLOYEE (Name, SSN, Address, DNO, SuperSSN)
DEPARTMENT (DName, DNO, MgrSSN)
DEPT_LOCATIONS (DNO, DLOCATION)
PROJECT (PNO, PNAME, PLOCATION, DNO)
WORKS_ON (SSN, PNO, Hours)
DEPENDENT (SSN, DEPENDENT_NAME, Bdate)

1. Write the SQL statements to create the tale DEPT_LOCATIONS 3M

2. Write the SQL statements for the given queries
a. For each project, retrieve the project number, the project name and the number of employees from department 5 who work on the project. 3M

b. List the names of managers who have atleast one dependent (answer as a correlated query) 3M

c. List the names of employees who live anywhere in Europe. 1M