VIT
Vellore Institute of Technology

Final Assessment Test – November 2024

Course: PMCA503L - Database Systems
Class NBR(s): 3140/3199/3244 Slot: A2+TA2
Time: Three Hours Max. Marks: 100

- KEEPING MOBILE PHONE/ANY ELECTRONIC GADGETS, EVEN IN 'OFF' POSITION IS TREATED AS EXAM MALPRACTICE
- DON'T WRITE ANYTHING ON THE QUESTION PAPER

Answer ALL Questions

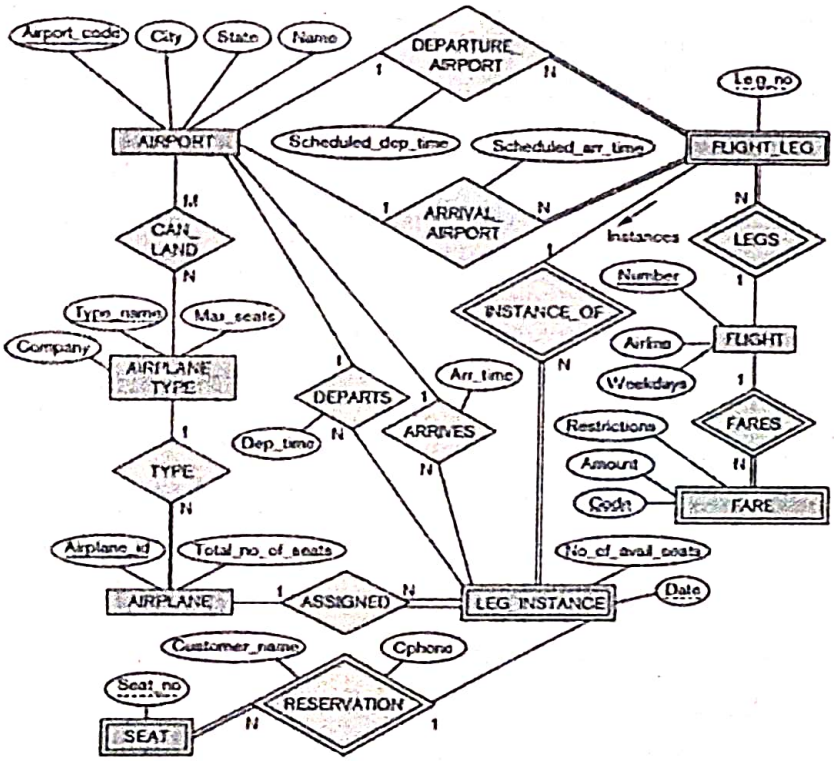
(10 X 10 = 100 Marks)

1. Mr. X is running a finance business. He lends money to customers on interest basis per month. He was storing everything in an excel sheet and calculating profit and loss. He wants to extend his business to another city and moves to a centralized database software now. What are the new features and advantages he will get now, compared to the old system he was using?

2. Design a database to keep track of information for an art museum. Assume that the following requirements were collected: The museum has a collection of ART_OBJECTS. Each ART_OBJECT has a unique Id_no, an Artist (if known), a Year (when it was created, if known), a Title, and a Description. The art objects are categorized in several ways, as discussed below. ART_OBJECTS are categorized based on their type. There are three main types: PAINTING, SCULPTURE, and STATUE, plus another type called OTHER to accommodate objects that do not fall into one of the three main types. A PAINTING has a Paint_type (oil, watercolor, etc.), material on which it is Drawn_on (paper, canvas, wood, etc.), and Style (modern, abstract, etc.). A SCULPTURE or a statue has a Material from which it was created (wood, stone, etc.), Height,Weight, and Style. An art object in the OTHER category has a Type (print, photo, etc.) and Style. ART_OBJECTs are categorized as either PERMANENT_COLLECTION (objects that are owned by the museum) and BORROWED. Information captured about objects in the PERMANENT_COLLECTION includes Date_acquired, Status (on display, on loan, or stored), and Cost. Information captured about BORROWED objects includes the Collection from which it was borrowed, Date_borrowed, and Date_returned. Information describing the country or culture of Origin (Italian, Egyptian, American, Indian, and so forth) and Epoch (Renaissance, Modern, Ancient, and so forth) is captured for each ART_OBJECT. The museum keeps track of ARTIST information, if known: Name,

DateBorn (if known), Date_died (if not living), Country_of_origin, Epoch, Main_style, and Description. The Name is assumed to be unique. Different EXHIBITIONS occur, each having a Name, Start_date, and End_date. EXHIBITIONS are related to all the art objects that were on display during the exhibition. Information is kept on other COLLECTIONS with which the museum interacts, including Name (unique), Type (museum, personal, etc.), Description, Address, Phone, and current Contact_person. Draw an EER schema diagram for this application. Discuss any assumptions you make and justify your EER design choices.

3.(a) Derive the Implementational Design (Relational Schema) for the given conceptual diagram.



OR

3.(b) Consider the universal relation R = {A, B, C, D, E, F, G, H, I, J} and the set of functional dependencies F = {{A, B} →{C}, {A} → {D, E}, {B}→{F}, {F} → {G, H}, {D} → {I, J}}.

i) What is the key for R? State which normal it is.

ii) Decompose R such that it satisfies the highest normal form. Discuss the algorithms used in decomposition.

4. a) Discuss insertion, deletion and modification anomalies. Why are they considered bad? Illustrate with examples.

b) 5NF is also called project-join normal form (PJNF). Justify your answer.

5. Consider the following relations

Customer (No, Name, Address, Phone)

Loan (Num, Cus_No, Amount, Ltype, Duration, sanctioned_date)

Account (Num, Cus_No, Balance)

Write **SQL Queries** to

a) Find the customer names who live in Vellore.

b) List the number of loans for each customer.

c) Find the customers who do not have any loan.

d) List the customer names who have all loan types.

e) Find the customers who have two loans and three accounts.

6.(a) Consider the following two ways of computing the names of employees who earn more than $100,000 and whose age is equal to their manager's age.

First, a nested query:

```
SELECT E1.ename FROM Emp E1
WHERE E1.sal > 100 AND E1.age = ( SELECT E2.age FROM Emp E2, Dept D2
WHERE E1.dname = D2.dname AND D2.mgr = E2.ename )
```

Second, a query that uses a view definition:

```
SELECT E1.ename FROM Emp E1, MgrAge A
WHERE E1.dname = A.dname AND E1.sal > 100 AND E1.age = A.age ;
CREATE VIEW MgrAge (dname, age)
AS SELECT D.dname, E.age FROM Emp E, Dept D
WHERE D.mgr = E.ename
```

i) Describe a situation in which the first query is likely to outperform the second query.

ii) Describe a situation in which the second query is likely to outperform the first query.

iii) Can you construct an equivalent query that is likely to beat both these queries when every employee who earns more than $100,000 is either 35 or 40 years old? Explain briefly.

OR

6.(b) Thirty Thousand student records are stored in unspanned manner in a double-sided disc pack of 10 disc with 20 tracks on each surface, 20 blocks of 512 bytes in each track. If the structure of student record is as follows

Student (Regno: varchar (10), Name: varchar(50), DOB: Date, Batch: Number(4), Programme: Varchar (8), School: Varchar(28))

Compare and contrast the various indexes if the following exist

i) Primary index on Regno

ii) Clustered index on School

iii) Secondary index on DOB

Note: Date occupies 8 bytes.

7. What is the two-phase locking protocol? How does it guarantee serializability? Illustrate with example.

8. Consider the three transactions *T*1, *T*2, and *T*3, and the schedules *S*1 and *S*2 given below. Draw the serializability (precedence) graphs for *S*1 and *S*2, and state whether each schedule is serializable or not. If a schedule is serializable, write down the equivalent serial schedule(s).

*T*1: *r*1 (*X*); *r*1 (*Z*); *w*1 (*X*);

*T*2: *r*2 (*Z*); *r*2 (*Y*); *w*2 (*Z*); *w*2 (*Y*);

*T*3: *r*3 (*X*); *r*3 (*Y*); *w*3 (*Y*);

*S*1: *r*1 (*X*); *r*2 (*Z*); *r*1 (*Z*); *r*3 (*X*); *r*3 (*Y*); *w*1 (*X*); *w*3 (*Y*); *r*2 (*Y*); *w*2 (*Z*); *w*2 (*Y*);

*S*2: *r*1 (*X*); *r*2 (*Z*); *r*3 (*X*); *r*1 (*Z*); *r*2 (*Y*); *r*3 (*Y*); *w*1 (*X*); *w*2 (*Z*); *w*3 (*Y*); *w*2 (*Y*);

9.

```
{
 "address": {
   "building": "1007",
   "coord": [ -73.856077, 40.848447 ],
   "street": "Morris Park Ave",
```

```
    "zipcode": "10462"
  },
  "borough": "Bronx",
  "cuisine": "Bakery",
  "grades": [
    { "date": { "$date": 1393804800000 }, "grade": "A", "score": 2 },
    { "date": { "$date": 1378857600000 }, "grade": "A", "score": 6 },
    { "date": { "$date": 1358985600000 }, "grade": "A", "score": 10 },
    { "date": { "$date": 1322006400000 }, "grade": "A", "score": 9 },
    { "date": { "$date": 1299715200000 }, "grade": "B", "score": 14 }
  ],
  "name": "Morris Park Bake Shop",
  "restaurant_id": "30075445"
}
```


For the above given restaurants collection structure, answer the following queries in MongoDB.

a) Display the next 5 restaurants after skipping first 5 which are in the borough Bronx.

b) Find the restaurants who achieved a score more than 90.

c) Find the restaurants that do not prepare any cuisine of 'American' and their grade score more than 70 and latitude less than -65.754168.

d) Find the restaurant Id, name, borough and cuisine for those restaurants which contain 'Reg' as three letters somewhere in its name.

e) Find the count of restaurants that received a grade of 'A' for each cuisine.

10. a) What are the differences between valid time, transaction time, and bitemporal relations?

b) Derive the mathematical notation for distributed fragmentation.

⇔⇔⇔ Z/K/TX⇔⇔⇔