CS4.301: Data and Applications (Monsoon '23)

End-Semester Examination

Maximum Marks: 50 Time: 2 hours

Instructions:

- This question paper consists of 7 questions. All questions are **compulsory**.
- If any question is ambiguous, state your assumptions clearly and proceed to answer. No clarifications will be provided during the exam.
- Clearly specify the attributes, entities that you are using in your queries, any ambiguity may lead to deduction of marks.

Q1) Consider the following populated tables which are a part of university management system:

Major

Name	Department
Computer Science	Engineering
Mathematics	Science
Physics	Science
Biology	Science
Chemistry	Science

Student

ID	Name	Age	GPA	Major
101	Alice	19	3.5	Computer Science
102	Bob	20	3.2	Mathematics
103	Charlie	21	3.8	Physics
104	David	18	2.9	Biology

Constraints on the database:

- Major Table:
  - Name: Primary key
  - Department: NOT NULL

- Student table:
  - ID: Primary key
  - Name: NOT NULL
  - Age: Should be > 0
  - GPA: Should be between 0 and 4
  - Major: References Major.Name

Mention **all** the database constraints that are violated when the following statements are executed (Treat each statement as independent): [2 + 2 + 2]

-
  ```
  INSERT INTO Student VALUES (104, 'Eve', 22, 4.1,
  'Chemistry');
  ```

-
  ```
  UPDATE Student SET Major = 'History' WHERE ID = 103;
  ```

-
  ```
  DELETE FROM Major WHERE Name = 'Physics';
  ```

Q2) Consider the following database schema for a restaurant management system:

- **Customers:** (c_id (Primary Key), customer_name, phone_number, email)
- **Restaurants:** (r_id (Primary Key), restaurant_name, location, cuisine_type)
- **MenuItems:** (item_id (Primary Key), item_name, price, restaurant_id (Foreign Key references Restaurants))
- **Orders:** (order_id (Primary Key), customer_id (Foreign Key references Customers.c_id), restaurant_id (Foreign Key references Restaurants.r_id), order_date, amount)
- **Chefs:** (chef_id (Primary Key), chef_name, specialty, years_of_experience, restaurant_id (Foreign Key references Restaurants.r_id))

Assumptions:

- The order_date column in Orders table has datatype DATE in the format (YYYY-MM-DD)
- The total_amount column in the Orders table stores the amount in Rupees.

Write SQL statements for the following 5 queries. You can also provide an explanation for your statement which will be evaluated in case your query is incorrect **(there is no partial marking for SQL statement)**. You can receive up to 1.5 marks for correct explanation.

a) Find the total revenue generated(which is the sum of the amount of all orders placed) by each restaurant in the October month of the current year. Display the results(restaurant name, restaurant location and revenue) in descending order of revenue. [3]

b) Find the customer who has spent the most money on orders. Display the customer's name and the total amount spent. [3]

c) Find the top 3 most frequent customers(based on the number of orders placed) of the restaurant having r_id = 90881. Display the customer's name and number of orders. [3]

d) List the names of restaurants where the average order amount is higher than the overall average order amount. [3]

e) Find the names of restaurants that have chefs with an average experience of more than 7 years. [3]

Q3) Consider the following three tables:

Students:

student_id	student_name	major	age
1	Alice	Computer Science	21
2	Bob	Physics	22
3	Charlie	Physics	20
4	David	Mathematics	23
5	Emily	Computer Science	22

Courses:

course_id	course_name	instructor
101	Database Design	Dr. Smith
102	Physics I	Prof. Johnson
103	Calculus II	Dr. Davis
104	Genetics	Prof. White
105	Organic Chem	Dr. Brown

Registrations:

registration_id	student_id	course_id	grade
1	1	101	A
2	2	102	B
3	3	103	C
4	4	104	B
5	5	105	A

Show the output of the following queries (You have to write the entire output that you would expect when the query is executed): [2.5 + 2.5]

a)

```
SELECT students.student_id, student_name, course_name, grade FROM
students JOIN registrations ON students.student_id =
registrations.student_id JOIN courses ON registrations.course_id =
courses.course_id WHERE grade = 'A' AND major = 'Computer
Science';
```

b)

```
SELECT students.major, AVG(students.age) as average_age FROM
students
RIGHT OUTER JOIN registrations ON students.student_id =
registrations.student_id
GROUP BY students.major ORDER BY average_age;
```

Q4) Consider a relational scheme R(A, B, C, D, E) with the following FDs:

- A → DE
- B → A
- D → C

a) Identify the primary key for R. (Show how it is a primary key) [2]

b) What do you understand by the second normal form? Is the above relation in 2NF? If yes, justify your answer. If not, decompose it so that it is in 2NF after decomposition. [2 + 2]

c) What do you understand by the third normal form? Is the above decomposed relation (the answer to Q2) in 3NF? If yes, justify your answer. If not, decompose it so that it is in 3NF after decomposition. [2 + 2]

Q5) Given a relation R(P, Q, R, S, T, U, V, W, X, Y) and Functional Dependency set FD = { PQ → R, PS → VW, QS → TU, P → X, W → Y }, determine whether the given R is in 2NF (Note that the key for R is PQS). If not, convert it into 2 NF. [5]

Q6) Consider the following definition of dependency preserving decomposition:

If we decompose a relation R into relations R1 and R2, all dependencies of R must be part of ***either R1 or R2 or must be derivable from combination of functional dependencies(FD) of R1 and R2***

a) Mention any two advantages & two disadvantages of decomposing a table into smaller tables in a database. [2]

b) Consider a relation scheme R(A, B, C, D) with the following functional dependencies:
A → B, B → C, C → D
Mention whether the following decompositions are dependency preserving and justify your answer: [2 + 2]

- R1(A, B, D), R2(B, C)
- R1(A, B), R2(B, C), R3(C, D)

R1
R2
4 3
2

21.5

1

23

Q7) Suppose there is a ternary relationship called **Teaches** between three entity types: Professor, Course, and Student. The relationship indicates which professor teaches which course to which student. Assume the following constraints:

- A course has to be taught by at least 1 professor and at most 2 professors and should be attended by at least 1 student and at most 100 students.
- A professor can teach no course or at most 3 courses.
- A student can attend no course or at most 6 courses.

Write the (min, max) participation constraint for each of the participating entity types and justify your answer (Please note that you have to write exact numbers instead of using N in your answer). [3]