Student Roll No.: _____

National Institute of Technology, Jamshedpur
End-Semester Examination (December 2024)

Branch: MCA (3rd Sem)
Course Name: DBMS
Course Code: CS3301

Max. Marks: 50
Time: 3 Hours
Course Instructor: Dr. Dinesh Kumar

Instructions:

1. *All questions are compulsory. All the subparts of a question are to be attempted together.*
2. *Refer the University Database Schema given below to solve the questions related to University Database*

classroom(*building*, *room_number*, *capacity*)
department(*dept_name*, *building*, *budget*)
course(*course_id*, *title*, *dept_name*, *credits*)
instructor(*ID*, *name*, *dept_name*, *salary*)
section(*course_id*, *sec_id*, *semester*, *year*, *building*, *room_number*, *time_slot_id*)
teaches(*ID*, *course_id*, *sec_id*, *semester*, *year*)
student(*ID*, *name*, *dept_name*, *tot_cred*)
takes(*ID*, *course_id*, *sec_id*, *semester*, *year*, *grade*)
advisor(*s_ID*, *i_ID*)
time_slot(*time_slot_id*, *day*, *start_time*, *end_time*)
prereq(*course_id*, *prereq_id*)

Q 1 a) Write the following queries in relational algebra, using the university schema: [5] Marks

- Find the names of all instructors in the Physics department.
- Find the set of all courses taught in the Spring 2018 semester and Fall 2017 Semester.
- Find the courses taught by instructors in the Physics department.
- Find all the courses taught in the Fall 2017 semester but not in Spring 2018 semester
- Find the ID and name of those instructors who earn more than the instructor whose ID is 12121.

b) Design a database for an automobile company to provide to its dealers to assist them in maintaining customer records and dealer inventory and to assist sales staff in ordering cars [5]

Each vehicle is identified by a vehicle identification number (VIN). Each individual vehicle is a particular model of a particular brand offered by the company (e.g., the XF is a model of the car brand Jaguar of Tata Motors). Each model can be offered with a variety of options, but an individual car may have only some (or none) of the available options. The database needs to store information about models, brands, and options, as well as information about individual dealers, customers, and cars.

Your design should include an E-R diagram, a set of relational schemas, and a list of constraints, including primary-key and foreign-key constraints.

Q 2 a) Write the following SQL queries using the university schema: [5]

- Find the number of instructors in each department who teach a course in the Spring 2018 semester.
- For each course section offered in 2017, find the average total credits (tot cred) of all students enrolled in the section, if the section has at least 2 students.
- Find the names of all instructors whose salary is greater than at least one instructor in the biology department.
- Write the equivalent SQL query for the following relational algebra.

$$t1 \leftarrow \sigma_P(r_1 \times r_2 \times \cdots \times r_m)$$
$$\Pi_{A_1, A_2, SumA3}(\sigma_{countA4 > 2}({}_{A_1, A_2}\gamma_{sum(A_3)\ as\ SumA3,\ count(A_4)\ as\ countA4}(t1)))$$

- What is the problem with the following query?

```
select dept_name, ID, avg (salary)
from instructor
group by dept_name;
```

b) Write the following SQL queries using the university schema: [5]

- List the names of students along with the titles of courses that they have taken. (Use *Natural Join*).
- Find all students who have not taken a course (use *left outer join*).
- What is the output of the following SQL query?
  *select * from takes natural right outer join student;*
- Display a list of all students in the Comp. Sci. department, along with the course sections, if any, that they have taken in Spring 2017; all course sections from Spring 2017 must be displayed, even if no student from the Comp. Sci. department has taken the course section.
- Create a view that lists all course sections offered by the Physics department in the Fall 2017 semester with the building and room number of each section

Q 3 a) Define the following terms with atleast one example for each: [5]

- Lossless decomposition
- Functional dependency preserving
- Non-Trivial Functional Dependency
- Transitive Functional Dependency
- Partial Dependency

b) Consider the following Relations, associated FDs and their decomposition: [5]

- A relation $R(ABC)$: $\{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$ is decomposed in two relations $R_1(AB)$ and $R_2(BC)$
- A relation $R(ABCD)$: $\{A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow B\}$ is decomposed in three relations $R_1(AB)$, $R_2(BC)$, $R_3(BD)$
- A relation $R(ABCD)$: $\{AB \rightarrow CD, D \rightarrow A\}$ is decomposed in two relations $R_1(AD)$ and $R_2(BCD)$

Explain which of the above decompositions are FD preserving and why?

Q 4 a) A B+ tree of order d has the following properties: [5]

1. Each internal node contains between d and $2d$ key values.
2. An internal node with M key values has $M + 1$ children.
3. The root, if it is an internal node, contains between 1 and $2d$ key values.
4. The distance of a node from the root is the length of the path from the root to that node.
5. All leaf nodes are equidistant from the root, meaning they share the same distance.
6. The height of the tree is defined as the distance from the root to any leaf node.

Answer the following questions based on these properties:

a) What is the total number of key values in the internal nodes of a B+ tree with l leaves ($l \geq 2$)?

b) What is the maximum number of internal nodes in a B+ tree of order 4 with 52 leaves?

c) What is the minimum number of leaves in a B+ tree of order d and height h ($h \geq 1$)?

b) Consider the database with the following parameters: [5]

The Block size is 1000, records are of size 100B of which 12B are the key field and pointer is of size 8B. The file consists of 100000 records.

Determine the following:

- The number of index blocks required for a sparse primary index on this file.
- The number of index blocks required for a secondary index on this file.

Q 5 a) Explain the basic time stamp protocol for transaction executions? Then determine which of the following Transaction in the below schedule will be rollback according to the Basic Time stamp protocol? [5]

$R1A, R2B, W1C, R3B, R1C, W2B, W3A$

b) Determine the equivalent view serializable schedule for the following schedule: [5]

$R2A, R2B, W1A, W1B, W2B, W3B$

**********End of Question Paper **********

Student Roll No.: Kapse.14

National Institute of Technology, Jamshedpur
Third Semester (Mid-Sem) Examination - Oct 2023

Branch: MCA (2nd Year)
Course Name: DBMS
Course Code: CS3301

Max. Marks: 30
Time: 2 Hours
Course Instructor: Dr. Dinesh Kumar

Instructions:

1. *All questions are compulsory. All the subparts of a question are to be attempted together.*

			Marks
Q 1	a)	What is an unsafe query? Give an example and explain why it is important to disallow such queries.	[2]
	b)	What factors might prompt you to opt for a database system instead of just saving data in operating system files? In what situations would it be advisable to avoid using a database system?	[2]
	c)	Discuss, in short, the difference between logical and physical data independence. Does the relational model, as seen by an SQL query writer, provide physical and logical data independence? Explain.	[2]
	d)	For the relation R(ABCDEFGH) with FD's= {CH->G, A->BC, B->CHF, E->A, F->EG such that F+ is exactly the set of FDs that hold for R.}.Find the highest normal form of R.	[2]
	e)	Let R1 (A, B, C) and R2 (D, E) be two relation schema, where the primary keys are shown underlined, and let C be a foreign key in R1 referring to R2. Suppose there is no violation of the above referential integrity constraint in the corresponding relation instances r1 and r2. Write an relational algebra expressions that would necessarily produce an empty relation?	[2]
Q 2	a)	Consider the following information about a university database:	[5]

- *Professors have an SSN, a name, an age, a rank, and a research specialty.*
- *Projects have a project number, a sponsor name (e.g., NSF), a starting date, an ending date, and a budget.*
- *Graduate students have an SSN, a name, an age, and a degree program (e.g., M.S. or Ph.D.).*
- *Each project is managed by one professor (known as the project's principal investigator).*
- *Each project is worked on by one or more professors (known as the project's co-investigators).*
- *Professors can manage and/or work on multiple projects.*
- *Each project is worked on by one or more graduate students (known as the project's research assistants).*
- *When graduate students work on a project, a professor must supervise their work on the project. Graduate students can work on multiple projects, in which case they will have a (potentially different) supervisor for each one.*
- *Departments have a department number, a department name, and a main office.*
- *Departments have a professor (known as the chairman) who runs the department.*
- *Professors work in one or more departments, and for each department that they work in, a time percentage is associated with their job.*
- *Graduate students have one major department in which they are working on their degree.*
- *Each graduate student has another, more senior graduate student (known as a student advisor) who advises him or her on what courses to take.*

Design and draw an ER diagram that captures the information about the university. Use only the basic ER model here; that is, entities, relationships, and attributes. Be sure to indicate any key and participation constraints.

b) Consider the university database and the ER diagram you designed in Q2 (a). Write SQL statements to create the corresponding relations and capture as many of the constraints as possible. If you cannot capture some constraints, explain why. [5]

c) Consider the following Schema: [5]

Suppliers(*sid:* `integer`, *sname:* `string`, *address:* `string`)
Parts(*pid:* `integer`, *pname:* `string`, *color:* `string`)
Catalog(*sid:* `integer`, *pid:* `integer`, *cost:* `real`)

Write the following queries in relational algebra:

i. Find the *sids* of suppliers who supply every part.
ii. Find the *sids* of suppliers who supply every red or green part.
iii. Find the *sids* of suppliers who supply every red part or supply every green part.
iv. Find the *pids* of the most expensive parts supplied by suppliers named Yosemite Sham.
v. Find the *pids* of parts supplied by at least two different suppliers.

d) Consider the instance of the Sailors relation shown in below Figure [5]

sid	*sname*	*rating*	*age*
18	jones	3	30.0
41	jonah	6	56.0
22	ahab	7	44.0
63	moby	null	15.0

Figure: An Instance of *Sailors*

i. If you divide the sum just computed by the count, would the result be the same as the average? How would your answer change if these steps were carried out with respect to the ***age*** field instead of ***rating***?

ii. Consider the following query: Find the names of sailors with a higher rating than all sailors with *age* < 21. The following two SQL queries attempt to obtain the answer to this question. Do they both compute the same result? If not, explain why. Under what conditions would they compute the same result?

```
SELECT S.sname
FROM   Sailors S
WHERE  NOT EXISTS ( SELECT *
                    FROM   Sailors S2
                    WHERE  S2.age < 21
                           AND S.rating <= S2.rating )
```

Query I

```
SELECT *
FROM   Sailors S
WHERE  S.rating > ANY ( SELECT S2.rating
                        FROM   Sailors S2
                        WHERE  S2.age < 21 )
```

Query II

iii. Let us define instance S1 of Sailors to consist of the first two tuples, instance S2 to be the last two tuples, and S to be the given instance. Show the left outer join of S with itself, with the join condition being *sid=sid*.

iv. Show the right outer join of S with itself, with the join condition being *sid=sid*.

v. Show the full outer join of S with itself, with the join condition being sid=sid.

**********End of Question Paper **********

Student Roll No.: _____

National Institute of Technology, Jamshedpur

3rd Semester (Mid-Sem) Examination (Sept-Oct 2024)

Branch: MCA (2nd Year)
Course Name: DBMS
Course Code: CS3301

Max. Marks: 30
Time: 2 Hours
Course Instructor: Dr. Dinesh Kumar

Instructions:

1. *All questions are compulsory. All the subparts of a question are to be attempted together.*

			Marks
Q 1	a)	What are the advantages of Physical Data Independence and Logical Data Independence in DBMS design? Also give atleast one example for each.	[3]
	b)	Consider the below query: *"Find the information about courses taught by instructors in the physics department with salary greater than 90,000"* Write three different but equivalent queries in Relational algebra that produces the above result.	[3]
	c)	Write SQL query with or without *case* statement to do the following change in database: *"Increase salaries of instructors whose salary is over $100,000 by 3%, and all others by a 5%"*	[3]
	d)	Suppose we have three relations r(A, B), s(B, C), and t(B, D), with all attributes declared as not null. • Give instances of relations r, s, and t such that in the result of ***(r natural left outer join s) natural left outer join t*** attribute C has a null value but attribute D has a non-null value • Are there instances of r, s, and t such that the result of ***r natural left outer join (s natural left outer join t)*** has a null value for C but a non-null value for D? Explain why or why not.	[3]
	e)	Weak entities in a database are those entities whose existence depends on another entity set, called its identifying entity set. • Identify three examples of weak entities from different real-world domains. • For each example, explain the following: What constitutes the weak entity and the strong entity in the relationship? • How the primary key of the weak entity is constructed using the strong entity's key.	[3]
Q 2	a)	In the context of a University schema, consider the purpose of having the ***section_id*** attribute within the ***Section*** entity in an ER model. After mapping to the relational model: • How does the inclusion of the ***section_id*** help in differentiating sections? Illustrate this with an example. • What is the impact of omitting the ***course_id*** attribute from the Section table? Illustrate this with an example. Based on the problems identified in both scenarios, propose solutions that would resolve the issues arising from changing the schema of ***section*** table. Use relevant examples to support your explanation.	[5]

Q 2 b) Reduce the below ER diagram for university enterprise into relational model by defining the unique relational schema for each entity set and each relationship set. Explain this reduction by following: [10]

- Representation of Strong Entity Sets
- Representation of Strong Entity Sets with Complex Attributes
- Representation of Weak Entity Sets
- Representation of Relationship Sets
- Redundancy of Schemas
- Combination of Schemas

Also draw the Schema diagram for the university database.

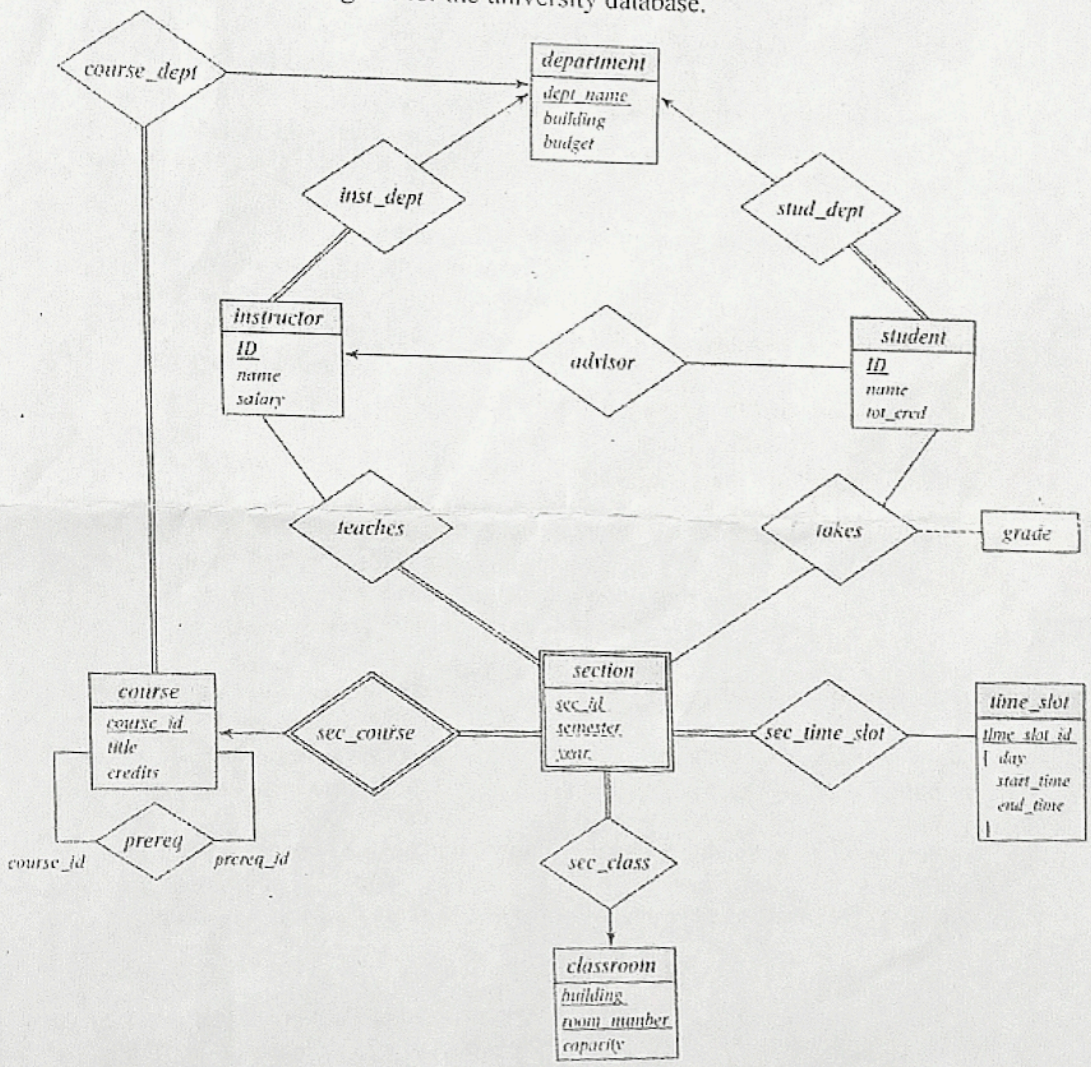

Figure 1: ER Diagram for University Enterprise

**********End of Question Paper **********