

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2025

Course Name: M.Tech. (QR&OR) II

Subject: Database Management Systems

Date: 17/11/2025

Maximum Marks: 50

Duration: 3 Hours

Note: Answer all questions.

Question 1

A socio-economic survey has been conducted in a set of remote villages in eastern India to study their agricultural practices. According to the survey, villages are identified by names only. Each village has a certain number of houses identified by a house number. Each house belongs to a certain caste. In each house, the number of persons involved in agriculture may vary. Some houses may also have some non-agricultural income.

Every village has three types of lands, classified as – *highly fertile*, *moderately fertile*, and *infertile*. Infertile land is used only to grow fruit trees. In each village, each plot of land belongs to a single type mentioned above. Each plot in a village is identified by a plot number and a cropping practice — either *mono-cropping* or *multi-cropping*. Each plot is owned by one of the houses in that village.

The village authority of every village keeps an account of the crops cultivated in that village in a calendar year and the average yield of each of them in tons/acre.

Tasks:

- (i) Draw a neat ER/EER diagram for the above scenario.
- (ii) Map the ER/EER diagram to proper relational schemas.

6+6=12

Question 2

Answer any two:

- (i) Explain how the **system log** helps in database recovery from a crash.
- (ii) What do you mean by **equivalence of two relational expressions**? State any 5 equivalence rules that can be used on such expressions?
- (iii) Explain different types of **indexing** with examples. Explain **hashing**, and compare it with indexing. A hash file uses the hash function $h(k) = K \bmod 10$, where K is the key value. There are 10 buckets (0–9). Insert the following keys: 12, 25, 36, 20, 30, 42. Find in which bucket each record will go.
- (iv) Write different types of query optimisation techniques and bring out the technique involved in '*Selection Operation*' - such as File scan, Index scan, complex selection.

5x2=10

Question 3(a)

Consider the following relation schemas:

Sailors(sid, sname, rating, age)

Boats(bid, bname, color)

Reserves(sid, bid, day_of_the_week, date)

Answer the following queries:

- (i) Find the sailor(s) details who have reserved **all** the boats.
- (ii) Find the sailor(s) who **never reserve boats on Saturdays**.

Tasks:

- (a) Express the above queries using **Relational Algebra**. (Assume basic operations: Selection, Projection, Cross-product, Set-difference, Union, and Intersection.)
- (b) Specify the above queries (i) and (ii) in **Domain Relational Calculus**.
- (c) Specify the above queries (i) and (ii) in **Tuple Relational Calculus**.

Question 3(b)

Choose the correct alternative(s) from the following:

- (i) A prime attribute of a relation schema R is an attribute that appears:
 - (a) in all candidate keys of R .
 - (b) in some candidate keys of R .
 - (c) in a foreign key of R .
 - (d) only in the primary key of R .
- (ii) In SQL, the following expression $100 < \text{NULL}$ evaluates to:
 - (a) True
 - (b) 0
 - (c) False
 - (d) Unknown

- (iii) What is the difference between **INNER JOIN** and **OUTER JOIN**?
- (a) INNER JOIN only returns matched rows, while OUTER JOIN returns all rows.
 - (b) INNER JOIN is used for tables with unique keys, while OUTER JOIN is used for non-unique keys.
 - (c) INNER JOIN is used for self-joins, while OUTER JOIN is used for joins with different data types.
 - (d) INNER JOIN and OUTER JOIN are interchangeable terms in SQL.
- (iv) What is the purpose of the **HAVING** clause in SQL?
- (a) To filter rows before the JOIN operation.
 - (b) To specify the order of the result set.
 - (c) To filter rows **after** the GROUP BY clause based on aggregated values.
 - (d) To filter rows based on a condition in a subquery.

(2+2)×3+4=16

Question 4

(a) Explain the concept of **Normalization** in Database Management Systems with suitable example. Discuss its importance in minimizing redundancy and maintaining data integrity. Define **First (1NF)**, **Second (2NF)**, and **Third Normal Form (3NF)** with suitable examples.

(b) Consider the following unnormalized relation for a company database:

Employee_ID	Employee_Name	Department	Dept_Location	Project_Code	Project_Name	Hours_Worked
E01	Alice	HR	London	P101	Recruitment Drive	20
E01	Alice	HR	London	P102	Employee Training	15
E02	Bob	IT	New York	P103	Software Upgrade	25

1. Identify the **functional dependencies** that exist in the relation.
 2. Decompose the table step-by-step into **1NF**, **2NF**, and **3NF**, showing intermediate relations at each stage.
- (c) Compare **Boyce-Codd Normal Form (BCNF)** with **Third Normal Form (3NF)**. Explain a situation where a relation in 3NF may still have anomalies that BCNF eliminates. Support your answer with an example relation and functional dependency analysis.

3+5+4=12

— End of Paper —