

PSG COLLEGE OF TECHNOLOGY, COIMBATORE - 641 004

Department of Department of Applied Mathematics and Computational Sciences

M. Sc (TCS) – Semester : V

CONTINUOUS ASSESSMENT TEST 1 Date: 29.08.2025

23XTE8 / 23SA63 - Big Data and Modern Database Systems

Time: 1 Hour 30 min.

Maximum Marks: 40

INSTRUCTIONS:

1. Answer **ALL** questions. Each question carries 20 Marks.
2. Subdivisions (a)(i) and (a)(ii) carries 2 marks each, subdivision (b) carries 6 marks each and subdivision (c) carries 10 marks each.
3. Subdivisions (a) and (b) will be with no choice and Subdivision (c) may be with choice but not in more than 1 question.

4. Course Outcome Table :

Qn.1	CO 1	Qn.2	CO 2
------	------	------	------

1. a)i) Analyze the key differences between Relational Database Management Systems (RDBMS) and Object-Relational Database Management Systems (ORDBMS) in terms of data modeling and query capabilities. Provide examples where necessary. **BTL 4**

ii) Write a Python snippet using geopy to convert the address of your current location into its latitude and longitude. **BTL 3**

b) i) Differentiate between the raster and vector data models used for representing geographic information. Provide a suitable example for each model to illustrate their respective strengths and use cases. (2 marks) **BTL 4**

ii) A GIS database stores city road networks (vector). Discuss how appropriate spatial indexing techniques can be applied to speed up range queries and nearest-neighbor queries on spatial data. Illustrate with examples. (4 Marks) **BTL 2**

c) A university is transitioning from Relational Database to an Oracle Object-Relational Database Management System (ORDBMS) to manage its student, faculty, and course data. Explain and demonstrate how traditional relational database concepts are implemented within this new paradigm. Provide a detailed explanation of how a composite attribute and multi-valued attributes can be effectively modeled using ORDB with example. Illustrate how one-to-one relationships and one-to-many relationships can be established using object references with suitable examples. **BTL 3**

2. a)i) A relation consists of 10,000 pages stored on disk. The available main memory has only three buffer pages. The relation is to be sorted using the External Sort-Merge algorithm. Determine

the number of initial sorted runs generated after the first pass. Calculate the total number of passes required to sort the relation completely, producing a single sorted run. **BTL 3**

ii) Consider the following SQL query on Online Book database. Transform this SQL query into relational algebra expression. Draw the initial query tree for this expression, and then derive their optimized query tree after applying heuristics on them. **BTL 3**

BOOK

ISBN	Book_title	Category	Price	Copyright_date	Year	Page_count	P_ID
------	------------	----------	-------	----------------	------	------------	------

PUBLISHER

P_ID	Pname	Address	State	Phone	Email_ID
------	-------	---------	-------	-------	----------

AUTHOR

A_ID	Aname	City	State	Zip	Phone	URL
------	-------	------	-------	-----	-------	-----

AUTHOR_BOOK

A_ID	ISBN
------	------

REVIEW

R_ID	ISBN	Rating
------	------	--------

Figure Database Schema for *Online Book Database*

SELECT ISBN, Book_title, Year, Page_count, Price FROM BOOK B, AUTHOR A, AUTHOR_BOOK AB WHERE B.ISBN = AB.ISBN AND AB.A_ID = A.A_ID AND Aname = 'Charles Smith';

b) Discuss the two ways of evaluating a query. Which one is better and why? Does the available buffer space affect the query evaluation speed? If yes, how? **BTL 2**

c) Consider two relations R and S stored on disk. Relation R occupies M blocks, Relation S occupies N blocks and the available main memory can hold B blocks at a time. The relations are to be joined using the Hash Join algorithm. a) Analyze the working of the Hash Join algorithm by writing its pseudo code and explain the partitioning and probing phases. b) Evaluate the efficiency of the Hash Join algorithm by estimating the total I/O cost (in terms of block transfers). Discuss how overflow situations are handled during hashing, and justify with a suitable example. **BTL 4**