

JK LAKSHMIPAT UNIVERSITY, JAIPUR

INSTITUTE OF ENGINEERING AND TECHNOLOGY

End Term Examination (Regular), November-December 2015

M. Tech. in Computer Science & Engg.

Semester I

Roll No.....

MTCS 103: Advanced DBMS

Time: 3 hours

Max. Marks: 100

Instructions to students:

1. Do not write anything other than your roll number on question paper.
2. Assume suitable data wherever required and mention it clearly.
3. Writing appropriate units, nomenclature, and drawing neat sketches/schematics/ flow chart wherever, required are an integral part of the answer.

Q.1A	Objective type multiple choice questions (10 questions @ 1 mark)	10
i.	Which of the following is true concerning an ODBMS? (a) They have the ability to store complex data types on the Web. (b) They are overtaking RDBMS for all applications. (c) They are most useful for traditional, two-dimensional database table applications. (d) All of the above.	
ii.	Which of the following is typical use of a parallel database? (a) Query Optimization (b) Indexing (c) Parallel query processing (d) None of the above	
iii.	Which of the following is a spatial indexing method? (a) R-tree (b) B-tree (c) X-tree (d) None of the above	
iv.	Which of the following is a horizontal mobile application? (a) Mail-enabled applications (b) Image editing (c) Weather forecasting (d) None of these	
v.	The application program typically communicates with a database server, through _____, or other protocols, in order to get or store data. (a) JDBC (b) ODBC (c) CGI (d) None of the mentioned	
vi.	Which of the following is an example of a database application where content-based retrieval is useful? (a) Medical imaging (b) Trademarks (c) Fashion and Fabric design (d) All of the above	
vii.	_____ is concerned with the data cleaning process during which unnecessary information is removed. (a) Data selection (b) Data preprocessing (c) Data transformation (d) None of these	
viii.	A data warehouse is said to contain a 'time-varying' collection of data because _____. (a) Its contents vary automatically with time (b) Its life-span is very limited (c) It	

	contains historical data (d) Its content has explicit time-stamp	
ix	Which step in Apriori algorithm uses Apriori property? (a) Join step (b) Prune step (c) Candidate generation step (d) None of these	
x	Which of the following schema(s) is best suitable for the data warehouse? (a) Snowflake (b) Star (c) Both(a) and (b) (d) Fact constellation	
Q.2	Short answer type questions (4 questions @ 2.5 marks)	10
i.	State comparison between RDBMS and ORDBMS.	
ii.	Explain with example nested relation in ORDBMS.	
iii.	What is the automatic analysis and manual identification approaches to 'multimedia indexing?	
iv.	Give the differences between OLAP and data mining	
Q.3A	What is the use of online analytical processing? Explain how the data is analyzed using this process?	05
Q.3B	Explain how the concept of OID differs from that of primary key and tuple identifier in the relational model.	05
Q.4A	What are multimedia databases? Describe its data management issues and application.	05
Q.4B	What is a cursor? Define types of cursor. Explain attributes of implicit and explicit cursor.	05
Q.5A	In real world data, tuples with missing values for some attributes are a common occurrence. Describe various methods for handling this problem.	05
Q.5B	What is Database Trigger? Give syntax for creating and deleting a trigger.	05
Q.6A	What is data reduction? Explain the various strategies for data reduction.	05
Q.6B	Name any three advantages of the STAR schema. Can you think of any disadvantages of the STAR schema?	05
	OR	
Q.6B	What are the different advantages of a distributed database management system? Discuss in detail the different functions of a DDBMS to manage the complexity of data distribution.	05
Q.7A	Discuss the architecture of parallel database. What form of parallelism (inter-query, interoperation, or intra-operation) is likely to be the most important for each of the following tasks? (a) Increasing the throughput of a system with many small queries (b) Increasing the throughput of a system with a few large queries, when the number of disks and processors is large.	10
	OR	

Q.7A Cluster the following eight points (with (x, y) representing locations) into three clusters A1(2,10) A2(2,5) A3(8,4) A4(5,8) A5(7,5) A6(6,4) A7(1,2) A8(4,9). Initial cluster centers are: A1(2,10), A4(5,8) and A7(1,2). The distance function between two points $a = (x_1, y_1)$ and $b = (x_2, y_2)$ is defined as: $\rho(a, b) = |x_2 - x_1| + |y_2 - y_1|$. Use k-means algorithm to find the three cluster centers. 10

Q.8A Write short notes on following: 10

- Spatial database
- Mobile database
- Simple Object Access Protocol (SOAP)
- Knowledge base system

Q.9A Write a PL/SQL block that will display the customer name, the fixed deposit number and the fixed deposit amount of the first 5 customers holding the highest amount in fixed deposits. 10

Q.9B The weather attributes are Outlook, Temperature, Humidity, and Wind speed. They can have the following values:
 Outlook= {sunny, overcast, rain}
 Temperature = {hot, mild, cool}
 Humidity= {high, normal}
 Wind={weak, strong} 10

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Construct decision tree using classification algorithm. Decide on which day you can play tennis.

OR

Q.9B Create a XML schema documents and XML DTDs to correspond to the hierarchies shown in figures: 10

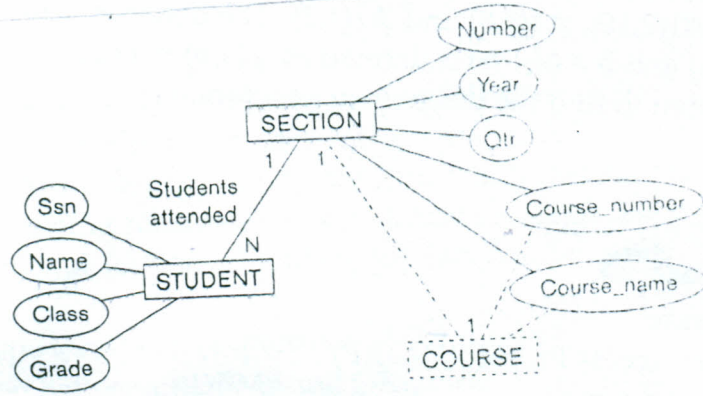


Fig:1

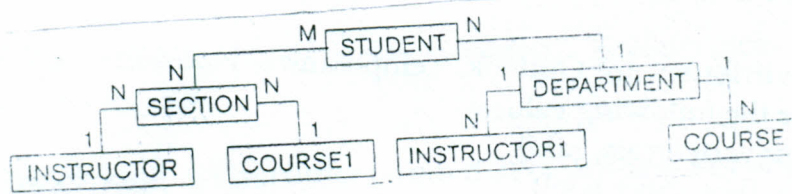


Fig:2