M.C.A. Sem-IV (CBGS)
Sub : ADTA.
(Advanced Database Theory & Application).

Date : 12|5|16

QP Code : 26709

[Total Marks : 80

(3 Hours)

N.B. : 1) Question No.1 is **compulsory**.
2) Attempt any **four** from the remaining **six** questions.

Q1 : a) Write short notes on (any two) : 10
1. Distributed Catalog management
2. Web Mining
3. Bitmap and Join Index

b) Compare followings (any two) 10
1. Semi join and Bloom join
2. Parallel database and Distributed database
3. Data Mining and OLAP

Q2 : a) Explain Knowledge Discovery Process (KDD) in detail. What is the role 08
of data mining in the KDD process.

b) Explain various operations of OLAP 07

Q3 : a) What are the various complex data types available in Object Relational 08
DBMS ? Explain with suitable examples.

b) What are frequent itemsets ? Describe an algorithm for finding frequent 07
itemsets.

Q4: a) Explain the concept of parallel query evaluation. 08

b) Explain ORDBMS Implementation challenges in detail. 07

Q5 a) Explain the features of XML and also differentiate between DTD and XML 08
Schema

b) What is datawarehouse and why it is needed ? Explain ETL (07
Extraction, Transformation and Loading)process in datawarehouse .

Q6 a) Use K-means algorithm to cluster the following 8 points into 3 clusters. 7 08
A1(2, 10) A2(2, 5) A3(8, 4) A4(5, 8) A5(7, 5) A6(6, 4) A7(1, 2),
A8(4, 9). Suppose that initial cluster centers (seeds) are: A1(2, 10),
A4(5, 8) and A7(1, 2). Run K-means algorithm for **three** iterations.
(The distance function between two points $a=(x1, y1)$ and $b=(x2, y2)$
is defined as: $d(a, b)$
$= |x2 - x1| + |y2 - y1|$)

[TURN OVER

b) What is the fundamental difference between MOLAP and ROLAP systems? Explain with suitable diagrams. 07

Q7 a) What is the difference between synchronous and asynchronous replication? Why has asynchronous replication gained in popularity? 08

b) Discuss in brief Time-in database and Spatial database. 07