

Name :

Roll No. :

Invigilator's Signature :

**CS/M.TECH (CSE)/SEM-2/MCS-202/2012
2012**

ADVANCED DATABASE MANAGEMENT SYSTEM

Time Allotted : 3 Hours

Full Marks : 70

The figures in the margin indicate full marks.

*Candidates are required to give their answers in their own words
as far as practicable.*

Answer any five questions.

5 × 14 = 70

1. a) State the correctness rule of fragmentation. 2
b) Consider the following global, fragmentation and allocation schema :

Global Schema : STUDENT (NUMBER, NAME, DEPT)

Fragmentation Schema : STUDENT₁ = $\sigma_{DEPT = 'EE'}$ (Student)

STUDENT₂ = $\sigma_{DEPT = 'CS'}$ (Student)

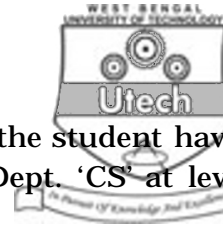
Allocation Schema : STUDENT₁ at sites 1, 2

STUDENT₂ at sites 3, 4

(Assume that 'EE' and 'CS' are the only possible values for DEPT)

- i) Write an application that requires the student number from the terminal and outputs the name and department at levels 1, 2, 3 of transparency.

3



- ii) Write an application that moves the student having number 232 from dept 'EE' to Dept. 'CS' at levels 1, 2 and 3 of transparency. 3
- iii) Write 2-phase commit protocol, when blocking problem will take place. 4 + 2
2. a) Describe the objectives of data distribution. 3
- b) Describe distributed serializability. Describe distributed 2PL. 2 + 2
- c) A country-wide drug-supplier chain operates from five different cities in the country has the following tables in the database :
- Drug Shop Mstr (DS-ds, DS-city, DS-contact no.)
- Medicine Mstr (med-id, med-name, manu-id)
- Manufacturer Mstr (manu-id, manu-name, manu-city)
- Other (med-id, DS-id, Qty)
- Suggest a fragmentation and allocation schema keep in mind following queries :
- i) List the manufacturer's names who belong to the same city in which the drug shop that has placed on order resides.
- ii) How many orders are generated from a city, say 'X' ?
- Justify your design and state your assumption clearly. 7
3. a) Consider the following schema :
- EMP (ENO, ENAME, TITLE)
- PROJ (PNO, PNAME, BUDGET)
- ASG (ENO, PNO, RESP, DUR)
- The relation PROJ is horizontally fragmented as
- PROJ 1 = $\sigma_{PNO \leq 'P3'} (PROJ)$
- PROJ 2 = $\sigma_{PNO > 'P3'} (PROJ)$



Transform the following query into a reduced query on fragments :

Select budget from PROJ, ASG

where PROJ . PNO = ASG . PNO

AND ASG . PNO = 'P4'

8

- b) Simplify the following query using the idempotency rules :

select ENO from ASG where

(NOT (TITLE = "PROGRAMMER")

and (TITLE = "PROGRAMMER" OR TITLE = "ELECT.ENG")

and not (TITLE = "Elect. ENG"))

OR ENAME = "J. DAS"

4

- c) How is the parametric query processed ?

2

4. a) Write 3-phase commit. How blocking problem is handled in 3-phase commit ?

4 + 2

- b) Write the different 2-phase locking strategy in distributed database in case of redundant copies.

3

- c) How deadlock is detected in distributed database ?

5

5. a) When is semijoin preferred in join queries ?

3

- b) Why query optimization is important in distributed databases ? Explain with example.

4

- c) Define Data Warehouse. Explain each term.

3

- d) Differentiate between OLTP and OLAP.

3 + 1

6. a) Why data warehouse is kept separate from OLTP database system ?

2

- b) Suppose that the data warehouse consists of the three dimensions time, item and location and two measures amount-sales and quantity-sales. Draw a star-schema for the above data warehouse.

4

- c) Define ROLAP, MOLAP, HOLAP.

3

- d) Discuss different OLAP operations.

5



7. a) Explain support and confidence. 2
 b) Find out strong association rule with support 20% and confidence 50% from the following transactions : 7

<i>TID</i>	<i>List of items</i>
T100	I_1, I_2, I_5
T200	I_2, I_4
T300	I_2, I_3
T400	I_1, I_2, I_4
T500	I_1, I_3
T600	I_2, I_3
T700	I_1, I_3
T800	I_1, I_2, I_3, I_5
T900	I_1, I_2, I_3
T1000	I_2, I_3

- c) Explain the architecture of data mining. 3
 d) Differentiate between classification and clustering. 2
8. a) What are the different methods of computing best split ? What is gini index ? What are entropy gain and gain ratio ? 3 + 2 + 2
 b) Describe K-means Algorithm. 4
 c) How noisy data is smoothed using binning method ? 3
9. Write notes on any *two* of the following : 2 × 7
 a) Snow flake schema
 b) Data mart
 c) Hierarchical method of clustering.
