Name : ....

Roll No. : ....

Invigilator's Signature : ....

CS/M.Tech (CSE)/SEM-2/CSEM-202/2013
2013
ADVANCED DATABASE MANAGEMENT SYSTEMS

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. Explain reference architecture with diagram. Define different levels of transparency. Discuss best fit, all beneficial site and additional beneficial site strategy for allocation of fragments.   6 + 3 + 5

2. a) Consider the following two allocation level designs of fragments $R1$, $R2$ and $R3$

   i) Allocation Design 1 : $R1$ at site 1; $R2$ at site 2; $R3$ at site 3;

   ii) Allocation Design 2 : $R1$ and $R2$ at site 1; $R2$ and $R3$ at site 3;

   With the following applications (all with same frequency of execution)

   i) $A1$, issued at site1, reads 5 records of $R1$ and update 5 records of $R2$.

30307 (M.Tech)   [Turn over

ii) $A2$, issued at site 3, update 5 records of $R3$ and update 5 records of $R2$.

iii) $A3$, issued at site 2, reads 10 records of $R1$, update 8 records of $R2$.

A single update to read cost ratio is 2.

If we take locality of reference as the objective of allocation, which of the above two designs of allocation schema is better ? Justify your answer.

b)   Determine common sub-expressions in the following global query based on global relations :

SUPPLY (**suppno,** prdno, eptno, quant), DEPT (**deptno,** dname, area, mgrno). Do step by step transformations by showing the operator tree of each step along with total no. of tuples processed in each operator tree, in order to achieve an optimized query.

(SL $_{deptno\ =\ 10}$ Dept NJN (SL $_{prdno\ =\ 'P1'}$ SUPPLY DF SL $_{deptno\ =\ '20'}$ SUPPLY)

UN (SL $_{deptno\ =\ 10}$ Dept NJN SL $_{prdno\ =\ 'P1'}$ SUPPLY)

(Here SL, NJN, DF and UN are select, natural join, difference and union operator).   7 + 7

3. Explain the algorithm of 2PC protocol in distributed environment for both coordinator and participants. Explain TimeStamp based concurrency control protocol. How are distributed 2-phase locking protocol and 2PC protocol related to each other ? 7 + 4 + 3

4. Explain concurrency control based on locking in DDBMS using the notion of primary copy and majority locking protocol. What is the additional threat to handle deadlock in centralized database compared to DDBMS ? What is false deadlock ? 7 + 4 + 3

5. a) What is query decomposition ? Is it same for centralized and distributed database ?

   b) Consider two relations PROPERTY (**pno**, type, price, mgr, bno) and BRANCH (**bno**, street, city). Now, consider the following SQL query :

   SELECT ∗

   FROM BRANCH b, PROPERTY p

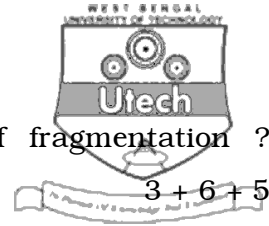   WHERE b.bno = p.bno AND p.type = 'flat';

   Assume that relations Property and Branch are horizontally fragmented as follows :

   $p1 : \sigma_{bno = B001 \wedge type = "House"}$ (PROPERTY) $b1 : \sigma_{bno = B001}$ (BRANCH)

   $p2 : \sigma_{bno = B001 \wedge type = "Flat"}$ (PROPERTY) $b1 : \sigma_{bno\,!\,= B001}$ (BRANCH)

   $p3 : \sigma_{bno\,!\,= B001}$ (PROPERTY)

   Draw the generic tree for the above query. Draw the reduced tree using reduction technique.

c) What are the correctness rules of fragmentation ? Explain each rule. $3 + 6 + 5$

6. What is distributed serializability ?

Let object x and y be stored at site 1, and objects z and k be stored at site 2. Determine, for each of the following executions, whether the execution is serializable or not. If the answer is affirmative, determine all possible total orders of transactions. If answer is negative, prove that there is no total order possible.

Execution 1 :

$$S1 : R_i(x) \; R_j(x) \; W_j(y) \; W_i(x)$$

$$S2 : R_i(k) \; R_j(z) \; W_j(k) \; W_i(k)$$

Execution 2 :

$$S1 : R_i(x) \; R_j(x) \; W_j(y) \; W_i(y)$$

$$S2 : W_i(z)$$

Execution 3 :

$$S1 : R_i(x) \; R_j(x) \; W_i(x) \; W_j(y)$$

$$S2 : R_i(z) \; R_j(z) \; W_j(z) \; W_i(k)$$

Execution 4 :

$$S1 : R_i(y) \; R_j(x) \; W_j(x)$$

$$S2 : W_i(z) \; R_i(k) \; R_j(k) \; W_i(k) \qquad 2 + (3 \times 4)$$

7. Consider the following distributed wait-for graph (DWFG) :
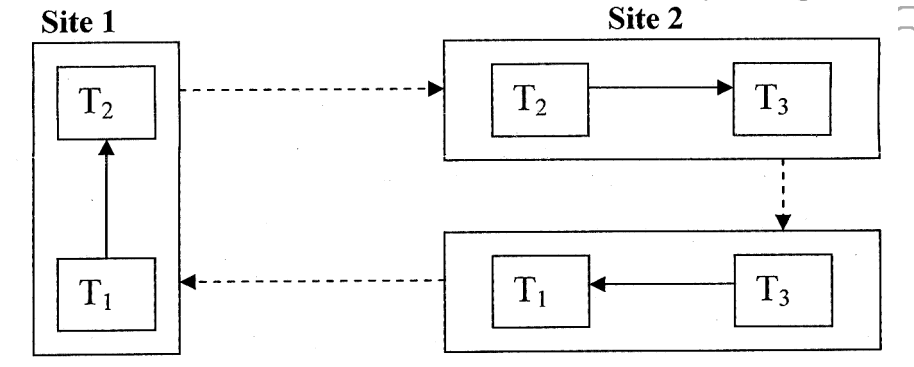

a) Detect the deadlock using distributed deadlock algorithm. 6

b) Describe deadlock detection using 'centralized controller' with example. 4

c) What is the difference between reliability and availability ? What is local autonomy ? 4

8. Justify any *four* of the following statements : $4 \times 3\frac{1}{2}$

a) 3-phase commitment protocol overcomes the limitation of 2-phase commitment protocol.

b) Unique timestamp generation is difficult in DDBMS than the centralized DBMS.

c) Query graph identified redundant relation in an SQL.
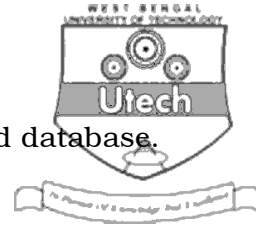
d) Cold restart is very hard in distributed database.

e) Bottom-up approach of distributed data distribution is applicable for integrating existing databases.

9. Consider a time-variant database and answer the following :

a) Explain the term 'valid time' in temporal database with example.

b) Do you think that a 'transaction time' is a 'fact' in the temporal database ? Provide sufficient support on your answer.

c) What are the two components that make a temporal database 'bitemporal' ? Explain them.

d) How does a temporal database differ from a normal database ?

e) Consider the relation S_FROM_TO (sno, sname, status, city, from, to) and SP_FROM_TO (sno, pno, from, to). Discuss why extra constraint is needed other than Primary key-Foreign key constraint.   2 + 2 + 4 + 2 + 4
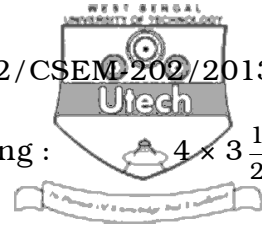
10. Write short notes on any *four* of the following : $4 \times 3\frac{1}{2}$

    a) Cold restart

    b) Distributed wait for graph

    c) Wound wait protocol

    d) Conservation timestamp ordering protocol

    e) False deadlock

    f) Distributed serializability

    g) Majority locking protocol.