DISTRIBUTED DATABASES
(CSEN 4261)

Time Allotted : 3 hrs **Full Marks : 70**

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and any 5 (five) from Group B to E, taking at least one from each group.

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

Group – A
(Multiple Choice Type Questions)

1. Choose the correct alternative for the following: **10 × 1 = 10**

(i) Vertical fragmentation of global relation, is subdivision of
(a) It's attributes (b) It's rows
(c) It's records (d) It's relation.

(ii) A homogenous distributed database is which of the following?
(a) The same DBMS is used at each location and data are not distributed across all nodes
(b) The same DBMS is used at each location and data are distributed across all nodes
(c) A different DBMS is used at each location and data are not distributed across all nodes
(d) A different DBMS is used at each location and data are distributed across all nodes.

(iii) "Users view each data object as logically unique." Which property of DDBMS supports this given statement?
(a) Fragmentation Transparency (b) Replication Transparency
(c) Location Transparency (d) Allocation Transparency.

(iv) Which operation is used to reconstruct the global relation from vertical fragments?
(a) Union (b) Selection (c) Natural join (d) Semi join.

(v) Which of the following is NOT a component of a DDBMS?
(a) Data communication component (b) Data Dictionary
(c) Distributed database component (d) Data security component.

(vi) A serially connected topology of all the sites of a network is known as
(a) Star (b) Ring (c) Bus (d) Daisy chains.

(vii) A Semi-join is which of the following?
(a) Only the joining attributes are sent from one site to another and then all of the rows are returned
(b) Only the joining attributes are sent from one site to another and then only the required rows are returned
(c) All of the attributes are sent from one site to another and then only the required rows are returned.
(d) None of the above.

(viii) _____ property will check whether all the operation of a transaction are completed or none.
(a) Atomicity (b) Consistency
(c) Isolation (d) Durability.

(ix) In order to represent run-time simplification in operator trees, the operator that is used is
(a) Cut (b) Union (c) Join (d) Intersection.

(x) The size of a fragment, size(R_i) is given by
(a) The sum of the sizes of the tuples in the fragment
(b) The size of the largest attribute of the fragment
(c) The sum of the sizes of all the attributes of the fragment
(d) None of the above.

Group – B

2. (a) What do you understand by a Distributed database? What are its two important aspects?

(b) What were the main motivations behind the development of distributed databases?

(2 + 2) + 8 = 12

3. (a) Define natural join and semi join operations.
What will be the result of (*Emp* **SJ**$_{\text{Emp.Dno=Dept.Dno}}$ *Dept)* based on the following relations Emp and Dept?

Emp		
3215	A	D1
2243	B	D2
3800	C	D1
1723	D	D3

Dept	
D2	X
D3	Y
D2	Z

(b) What are the parameters that characterize the basic functions of a computer network?

(4 + 2) + 6 = 12

Group – C

4. (a) What do you understand by Mixed Fragmentation? Explain with an appropriate example.

(b) Given are the relations:
Supplier(supNo,sname,code) where code can be either 'indian' or 'foreign'.
Parts(partNo,partName,supNo,price) with supNo being a foreign key in Parts referencing Supplier.supNo.
Fragment the relation *Parts* based on supNo. What type of fragmentation needs to be done? Justify if all the correctness rules are met or not.

6 + (2 + 1 + 3) = 12

5. Consider the following schemata:
Global schema: EMP (Enum, EName, Dept)
(i) Assume EMP has a fragmentation by Dept and 'Developer', 'Tester' and 'Marketing' are the only possible values of Dept. What type of fragmentation is this? Define the fragmentation.
(ii) Assume the allocation of the fragments as follows:
Allocation schema: EMP1 at site 1, EMP2 at sites 2, 4 and EMP3 at site 3. Write an application that requires the employee number (Enum) from the user and outputs the Ename and dept at levels 1, 2 and 3 of transparency.
(iii) Write an application that transfers the employee having Enum = 100 from department "Developer" to "Marketing", at levels 1 and 2 of transparency.

2 + (3 × 2) + (2 × 2) = 12

Group – D

6. (a) Given is the following algebraic expression. Draw the equivalent operator tree and apply simplification rules, if applicable, to get an optimized tree.
(**SL**$_{\text{DEPTNUM=10}}$ *DEPT* **NJN** (**SL**$_{\text{PNUM="P1"}}$ *SUPPLY* **DF** SL$_{\text{PNUM="P2"}}$ *SUPPLY*))
UN (**SL**$_{\text{DEPTNUM=10}}$ *DEPT* **NJN SL**$_{\text{PNUM="P1"}}$ *SUPPLY*)

Where the global schema is as follows:
DEPT = (DEPTNUM, NAME, AREA, MGRNUM)
SUPPLY = (SNUM, PNUM, DEPTNUM, QUAN)

(b) What are parametric queries? Why are temporaries used in parametric queries? Explain your answer with suitable example(s).

6 + (3 + 3) = 12

7. (a) What are the problems that need to be addressed during query optimization?

(b) Consider a given relation *Student* as follows:

Attribute	Size	No. of distinct values
Roll	5	120
Name	12	100
Dept	3	12
Phn	10	120

There are 120 tuples in the relation *Student* and Roll is the primary key.
(i) Draw the database profile for *Student*.
(ii) Suppose there are two horizontal fragments of *Student,* namely *Student1 (allocated at site 1)* and *Student2 (allocated at site 3).*

If *Student1* has 50 tuples and the number of distinct values of the attributes in the fragments is given as follows, draw the database profiles for the fragments.

Attributes	*Student 1*	*Student 2*
Name	42	58
Dept	8	10
Phn	50	70

3 + (3 + 6) = 12

Group – E

8. (a) Describe the 2PC protocol stepwise. How does the 2PC protocol deal with the failure of a participating site?

   (b) Explain the Majority locking and Primary Copy locking schemes used for concurrency control in DDB?

(5 + 3) + (2 + 2) = 12

9. (a) Describe in brief the contents of distributed database catalogs.

   (b) What are the two ways in which authorization rules are allocated and enforced in distributed database systems?

6 + 6 = 12

Department & Section	Submission Link
CSE A + B + C	https://classroom.google.com/c/MzAwODAwMjMxMDUw/a/MzU2MzU2OTczMjgw/details