



Name :

Roll No. :

Invigilator's Signature :

CS/M.Tech (SE, IT, CSE)/SEM-2/PGSE-201, PGIT-201, PGCSE-203A/2013

2013
ADVANCED DBMS

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.

Question No. 1 is compulsory and any four from the rest.

1. Comment critically on the following statements and justify your opinion (any four) :

$$4 \times 3\frac{1}{2} = 14$$

- a) Distributed database is essentially a variant of spatial database only.
- b) With increasing number of faults, the availability of a distributed database system degrades gracefully as compared to a centralized database system.
- c) A data warehouse stores non-volatile, historical data used for transactional processing.
- d) Allocation schema for a distributed database system is site independent.
- e) Semi-structured data can be best represented by relational data model.



2. a) How is distribution transparency achieved for distributed databases in multiple levels ?
- b) 'Redundancy control using normalization is not practiced for designing distributed database systems' – evaluate the statement and justify your opinion.
- c) What is a prime implicant ? How is it related for better design of a distribute database system ?
- d) Explain the significance of join graphs for distributed database design. 2 + 6 + 4 + 2
3. a) Explain the role of semi-join for query optimization in a distributed environment using as example.
- b) M/s Travello Resorts Private Limited maintains a chain of serviced apartment houses in important tourist places across India. The information to be stored is given below :
- i) Apartment information : apartment id, city, number of rooms, room tariff.



- ii) Guest information : guest id, name, address, phone number
- iii) Booking information : guest id, apartment id, number of rooms, period of booking.

Design a distributed database solution for apartment reservation system to be accessed from multiple terminals in multiple cities considering the following frequent queries/operations :

- i) Request for reservation of rooms. If a guest wants to book multiple apartments, for different period, separate reservation is to be done for each apartment and period of booking.
- ii) Availability of rooms - the system is to check for availability of room(s) in a particular apartment for the requested period.
- iii) Revenue query – the system is to be used to enquire about the revenue generated in a single apartment or in the entire chain for a period.



Your design should include the definition of global schema, fragmentation schema and allocation schema.

State your assumptions, if any.

8 + 6

4. a) Consider the following global, fragmentation and allocation schema :

Global schema :

Student (Roll, Name, Department)

Fragmentation schema :

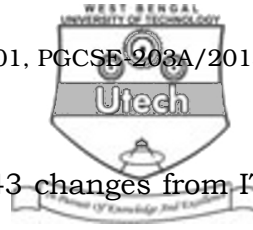
$Student_1 = SL_{Department = "IT"} Student$

$Student_2 = SL_{Department = "CSE"} Student$

Allocation schema : $Student_1$ at sites 1, 2

$Student_2$ at sites 3, 4

- i) Write a query that accepts the roll number of student from the terminal and outputs the name and department of the student, at levels 1, 2, and 3 of transparency.



- ii) A student having roll number 243 changes from IT to the CS department. Write a query at levels 1 and 2 of transparency to update the database accordingly.
- b) Consider an additional relation marks (Roll, Paper_code, Total) that store total marks obtained by a student in each paper. Define a derived fragmentation schema to split the marks table in manner such that records for students belonging to the two departments are kept separately in two different fragments. (5 + 5) + 4
5. a) What is virtual data warehouse ?
- b) A data warehouse consists of the four dimensions time, product, location, and one fact – profit. Draw and enumerate a cube data model for quarterly data on 4 products, and 3 locations for a total period of 1 year. Make your own assumption on the values.
- c) Propose the Star Schema for the above warehouse. State your assumption, if any. 2 + 6 + 6



6. a) Define global state.
- b) Why is it challenging to record the global state of a distributed system ?
- c) Explain in brief the 2-phase commit protocol for transaction management in a distributed database environment.
- d) Explain how the atomicity of a transaction is ensured under 2-phase commit protocol when a ready message from a site is lost before it reaches the coordinator.
- e) What would be the impact of different granularity of lock size for concurrency control in a distributed database environment ?
- f) Define user mobility, and name transparency.

1 + 2 + 4 + 2 + 2 + 3



7. a) Name five different classification technique.
- b) What kind of data is the decision tree method most suitable for ? Briefly outline the major steps of the algorithm to construct a decision tree.
- c) What is the advantage of using decision tree based classification.

3 + 2 + 6 + 3
