**Total No. of Questions : 12]** **[Total No. of Pages : 4**

**P1510** **[3764]-432**

# B.E. (IT)

## ADVANCED DATABASE MANAGEMENT

## (414442) (2003 Course)

***Time : 3 Hours]*** ***[Max. Marks : 100***

***Instructions to candidates :***

1) ***Answers to the two sections should be written in separate books.***
2) ***Neat diagrams must be drawn wherever necessary.***
3) ***Assume suitable data, if necessary.***
4) ***Section I : Q.1 or Q.2, Q.3 or Q.4, Q.5 or Q.6.***
5) ***Section II : Q.7 or Q.8, Q.9 or Q.10, Q.11 or Q.12.***

## SECTION - I

***Q1)*** a) Explain speed up and scale up with Parallel System. **[5]**

b) Describe the benefits and drawbacks of pipelined parallelism. **[6]**

c) Explain any two Parallel Database Architectures. **[6]**

OR

***Q2)*** a) Describe different approaches to handle cache coherency problem in Parallel Databases. **[9]**

b) Write a short note on:

i) Hash Partitioning.

ii) Fragment and Replicate Join. **[8]**

***Q3)*** a) Explain the use of reduction techniques to generate and optimized query in distributed databases using different types of fragmentation with suitable examples. Draw relational algebra trees. **[5]**

b) Write a short note on Persistent Messaging in Distributed Transaction Processing. **[6]**

c) Explain Heterogeneous distributed databases. **[6]**

OR

***Q4)*** a) Compute semi-join r $\alpha$ s for the relations r and s. **[5]**

| Relation r | | |
|---|---|---|
| A | B | C |
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 1 | 2 | 4 |
| 5 | 3 | 2 |
| 8 | 9 | 7 |

| Relation s | | |
|---|---|---|
| C | D | E |
| 3 | 4 | 5 |
| 3 | 6 | 8 |
| 2 | 3 | 2 |
| 1 | 4 | 1 |
| 1 | 2 | 3 |

b) Describe the voting and read-any-write-all approaches to synchronous replication. **[6]**

c) Explain Optimistic methods for Distributed Concurrency Control. **[6]**

***Q5)*** a) Consider Relations for bibliography. **[12]**

Book (title, author, year, publisher, place)

Article (title, author, journal, year, number, volume, pages)

Author (lname, fname)

Create DTD and XML Schemes.

Write queries in XQuery on the bibliography fragment.

i) Find all authors who have authorized a book and an article in the same year.

ii) Display books and articles sorted by year.

iii) Display books with more than one author.

iv) Find all books that contain the word “database” in their title and the word “Korth” in an author’s name.

b) Explain advantages and disadvantages of the Web-DBMS approach.**[4]**

OR

***Q6)*** a) Explain XML Applications for storing and communicating data and for accessing Web services. **[8]**

b) Describe the various issues for efficient evaluation of XML Queries.**[8]**

## SECTION - II

***Q7)*** a) Discuss the activities associated with a data warehouse for Financial Services with the help of following points. **[10]**

- Business processes.
- Business Questions expected in data warehouse environment.
- Design schemes.
- Failures and Backup strategies.

b) Explain Kimball's nine steps design for Data Warehouse. **[7]**

OR

***Q8)*** a) Write short notes on: **[10]**

i) Warehouse Manager.

ii) Materialized View.

b) Explain different indexing techniques in Data Warehouse. **[7]**

***Q9)*** a) Define with suitable example. **[12]**

i) Entropy.

ii) Information Info(T).

iii) Information Gain.

iv) Gain Ratio.

v) GINI Index.

Write ID3 algorithm with student data set in details. Construct a decision tree for at least 15 records in data set using ID3.

b) Write a short note on Text Mining. **[5]**

OR

***Q10)***a) Explain data preprocessing in Data Mining. **[6]**

b) Explain clustering major approaches. **[6]**

c) Describe Bayesian classification approaches for Fraud detection application. **[5]**

***Q11)***a) Define Information Retrieval System. Describe how it is differ from database system. **[6]**

b) Write short notes on: **[10]**

i) Signature Files.

ii) Ranking Document Similarity.

OR

***Q12)***a) Describe distinct ways a user can find information on the web. **[6]**

b) Write short notes on: **[10]**

i) Web Crawler.

ii) Retrieval Effectiveness.