Total No. of Questions : 12] [Total No. of Pages : 3

P 1220 [3664] - 339

B.E. (Computer Engineering)

ADVANCED DATABASES

(2003 Course) (410445) (Elective - I) Sem 1

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

Instructions to the candidates:

1) ***Answers to the two sections should be written in separate books.***
2) ***Neat diagrams must be drawn wherever necessary.***
3) ***Figures to the right indicate full marks.***
4) ***Assume suitable data, if necessary.***

SECTION - I

Q1) a) "Some parallel database systems store an extra copy of each data item on disks attached to a different processor" **[4]**

i) Explain the reason for above strategy.

ii) What is the advantage if you partition the copies of data items of a processor across multiple processors.

b) Explain the following with suitable example **[6]**

i) Pipelined parallelism.

ii) Independent parallelism.

c) What is meant by horizontal partitioning? Explain any two partitioning techniques. **[6]**

OR

Q2) a) In a range selection on a range - partitioned attribute, it is possible that only one disk may need to be accessed. Describe the benefits and drawbacks of this property. **[4]**

b) What is interquery parallelism? Explain cache coherency problem and protocol available to guarantee cache coherency. **[6]**

c) Explain parallel External sort-Merge. **[6]**

Q3) a) Explain Data Replication in distributed databases. **[6]**

b) Explain the different system failure modes in distributed transaction model. How the two-phase commit protocol handles these failures. **[10]**

c) What is multimaster replication? **[2]**

OR

P.T.O.

Q4) a) Explain Data Fragmentation in distributed databases. **[6]**

b) Explain the following with respect to robustness of distributed databases.

i) Read one, write all available protocol. **[10]**

ii) Co-ordinator selection.

iii) Majority-based approach.

c) What is Persistent messages? **[2]**

Q5) a) Explain the following with respect to web architecture **[8]**

i) Web server.

ii) Common gateway interface.

iii) Cookie.

iv) Uniform Resource Locator.

b) Explain the components of an XML document with suitable example. **[8]**

OR

Q6) a) Explain the following terms and describe what they are used for **[6]**

i) XPath.

ii) SOAP.

b) Write short notes on **[8]**

i) Client - Server architecture.

ii) Advantages of N-tier architecture.

c) What is XSL? **[2]**

SECTION - II

Q7) a) What is a Data warehouse? Explain the key features of Data warehouse. **[6]**

b) Write short notes on : **[12]**

i) Handling missing values in Data cleaning.

ii) Three-Tier data warehouse architecture.

iii) Data cubes.

OR

Q8) a) Explain the following OLAP operations on multidimensional data **[6]**

i) Roll-up.

ii) Slice and dice.

b) Write short notes on **[12]**

i) Data Transformation.

ii) Data Reduction.

iii) Fact constellation.

Q9) a) Explain the following with suitable example [8]

i) Data characterization.

ii) Data discrimination.

b) State and explain the algorithm to generate a decision tree from training tuples. [8]

OR

*Q10)*a) Explain in detail classification and prediction. [8]

b) What is outlier analysis? [4]

c) Write the k-means algorithm for partitioning. [4]

*Q11)*a) What is relevance feedback? Explain in brief. [4]

b) Explain in detail Web Search Engines. [6]

c) What is the difference between a false positive and a false drop? [2]

d) Explain the following terms [4]

i) Homonyms.

ii) Ontologies.

OR

*Q12)*a) What is meant by search engine spanning? Explain in brief. [4]

b) Explain in detail Information retrieval and structured data. [6]

c) What is the difference between precision and recall? [2]

d) Explain the following terms : [4]

i) Synonyms.

ii) Inverted index.