NEW SCHEME

CS751

Reg. No. 4VM02CS751

Seventh Semester B.E. Degree Examination, January/February 2006

Computer Science/Information Science and Engineering

Advanced DBMS

Time: 3 hrs.) (Max.Marks : 100

Note: *Answer any FIVE full questions.*

1. (a) What are mixed files used for? What are the other types of primary file organization? **(5 Marks)**

(b) Consider a disk with the following characteristics: block size is 512 bytes, interblock gap size is 128 bytes, number of blocks per track is 20; number of tracks per surface is 400. A disk pack consists of 15 double size disks.

i) What is the total capacity of a track, and what is its useful capacity (Excluding interblock gaps)?

ii) How many cylinders are there?

iii) What are the total capacity and the useful capacity of a cylinder?

iv) What are the total capacity and the useful capacity of a disk pack?

v) Suppose that the disk drive rotates the disk pack of a speed of 2400 rpm; what are the transfer rate in bytes/m sec. and the block transfer time in m sec? What is the average rotational delay in ms sec? What is the bulk transfer rate? **(15 Marks)**

2. (a) How does a B tree differ from a B^+ tree? Why is a B^+tree usually preferred as an access structure to a data file? **(5 Marks)**

(b) What is a grid file? What are its advantages and disadvantages? **(5 Marks)**

(c) Consider a disk with block size 512 bytes. A block pointes is 6 bytes long and a record points is 7 bytes long. A file has 30000 employee records of fixed length. Each record has ihc following fields: Name (30 bytes), SSN (9 bytes), department code (9 bytes), address (40 bytes), phone (9 bytes), birthdate (8 bytes), sex (1 byte), job code (4 bytes), salary (4 bytes, real number); an additional byte is used as a deletion mark.

I) Calculate the record size R in bytes

ii) Calculate the blocking factor and the number of file blocks, assuming an unspanned organization.

iii) Suppose the file is ordered by the key field SSN and we want to construct a primary index on SSN. Calculate :

(I) The index blocking factor

(II) The number of first-level index entries and the number of first - level index blocks.

(III) The number of block accesses needed to search for the retrieval of a record from the file-given an SSN value-using the primary index. **(10 Marks)**

Contd.... 2

c. Discuss how time is represented in temporal databases and compare the different time dimensions. **(05 Marks)**

d. How do spatial databases differ from reguiar databases? **(05 Marks)**

6 a. What are the main reasons for and potential advantages of distributed databases? **(10 Marks)**

b. Discuss the factors that do not appear in centralized systems that affect concurrency control and recovery in distributed systems. **(10 Marks)**

7 a. What are goals or tasks that data mining attempts to facilitate? **(05 Marks)**

b. What are five types of knowledge produced from data mining? **(05 Marks)**

c. Describe the characteristics of a data ware house. Divide them into functionality of a ware house and advantages users derive from it. **(10 Marks)**

8 Write notes on:

a. Mobile databases

b. GIS

c. Multimedia databases

d. Genome data management. **(20 Marks)**

*****

CS751

USN

NEW SCHEME

Seventh Semester B.E. Degree Examination, Dec. 06 / Jan. 07
CS / IS
Advanced DBMS

Time: 3 hrs.] [Max. Marks:100

Note: ***Answer any FIVE full questions.***

1 a. Describe the different types of records and their characteristics. **(08 Marks)**
b. Explain different file organizations from database perspective. **(06 Marks)**
c. Explain external hashing for disk files using extendible hashing. **(06 Marks)**

2 a. Explain the difference between:
i) Primary Vs secondary indexes
ii) Dense Vs sparse indexes and
iii) Clustered Vs unclustered indexes. **(12 Marks)**
b. Describe the B+ tree insertion algorithm. **(08 Marks)**

3 a. Describe the steps involved in query processing in DBMS. **(05 Marks)**
b. Explain the components of query cost and the way the query cost is quantified. **(05 Marks)**
c. Describe the options available to implement SELECT operation. **(10 Marks)**

4 a. Discuss the important decisions to be made during physical database design. **(06 Marks)**
b. Describe the ODMG object model of OODBMS. **(14 Marks)**

5 a. What are temporal databases? How are they different from conventional databases? **(10 Marks)**
b. What is deductive database? Explain prolog and data log notation. **(10 Marks)**

6 a. What is distributed databases? What are advantages of distributed database? **(10 Marks)**
b. What is need of data fragmentation, replication and allocation techniques in distributed databases? **(10 Marks)**

7 a. What is data mining? What are goals of data mining and knowledge discovery? **(10 Marks)**
b. What are mobile databases? Explain mobile computing architecture with diagram. **(10 Marks)**

8 Write short notes on:
a. Data ware housing
b. Geographic information systems
c. Multimedia databases
d. 3-Tier client server architecture. **(20 Marks)**

*****

CS/IS751

USN

NEW SCHEME

Seventh Semester B.E. Degree Examination, May / June 2006

CS / IS

Advanced DBMS

Time: 3 hrs.] [Max. Marks:100

Note: 1. ***Answer any FIVE full questions.***

1 a. Discuss the advantages and disadvantages of using
i. an unordered file ii. an ordered file
iii. a static hash file with buckets and chaining. **(10 Marks)**

b. A PARTS file with Part# as hash key includes records with the following Part# values: 2369, 3760, 4692, 4871, 5659, 1821, 1074, 7115, 1620, 2428, 3943, 4750, 6975, 4981, and 9028. The file uses eight buckets, numbered 0 to 7. Each bucket is one disk block and holds two records. Load these records into the file in the given order, using the hash function $h(k) = k \bmod 8$. Calculate the average number of block accesses for a random retrieval on Part#. **(10 Marks)**

2 a. How does multilevel indexing improve the efficiency of searching an index file? **(05 Marks)**

b. How does a B-tree differ from a B^+-tree? Why is a B^+-tree usually preferred as an access structure to a data file? **(05 Marks)**

c. A PARTS file with Part# as key field includes records with following Part# values: 23, 65, 37, 60, 46, 92, 48, 71, 56, 59, 18, 21, 10, 74, 78, 15, 16, 20, 24, 28, 39, 43, 47, 50, 69, 75, 8, 49, 33, 38. Suppose that the search field values are inserted in the given order in a B^+-tree of order $p = 4$ and $p_{leaf} = 3$; show how the tree will expand and what the final tree will look like. **(10 Marks)**

3 a. What is meant by the term heuristic optimization? Discuss the main heuristics that are applied during query optimization. **(10 Marks)**

b. What is the difference between pipelining and materialization? **(05 Marks)**

c. What is meant by semantic query optimization? How does it differ from other query optimization techniques? **(05 Marks)**

4 a. What are the origins of the object-oriented approach? **(05 Marks)**

b. What are the differences and similarities between objects and literals in the ODMG object model? **(10 Marks)**

c. What are main differences between designing a relational database and an object database? **(05 Marks)**

a. What are the differences between row-level and statement-level rules? **(05 Marks)**

b. Discuss some applications of active databases. **(05 Marks)**

Contd... 2

3. (a) What is a query execution plan?

(b) What is meant by cost-based query optimization? (5 Marks)

(c) Given a query on an EMPLOYEE database to find the lost names of employee born before 1957, who work on the project name 'Aquarius'; write the initial query tree. Write the following steps to convert the given query from initial query tree using heuristic-optimization. (5 Marks)

i) Moving SELECT operation down the tree.

ii) Applying more restrictive SELECT operation.

iii) Replacing Cartesian product and SELECT with join operation.

iv) Moving project operation down the query tree.

```
SELECT   LNAME
FROM     EMPLOYEE, WORKS-ON,    PROJECT
WHERE    PNAME='Aquarius'       AND
         PNUMBER =PNO           AND
         ESSN=SSN               AND
         BDATE= '1957-12-31';
```

(10 Marks)

4. (a) What are the differences and simiperities between objects and literls in the ODMG object model? (6 Marks)

(b) Describe the steps of the algorithm for object database design by EER to 00 mapping. (10 Marks)

(c) How would a view integration tool work? (4 Marks)

5. (a) What are the differences among immediate, deferred and detached consideration of active rule conditions? (8 Marks)

(b) Consider the relational schema shown in fig.5(b), write active rules for keeping the SUM_COMMISSIONS attribute of SALESPERSON equal to the sum of the COMMISSION attribute in SALES for each person.

SALES

S_ID	V_ID	COMMISSION

Sales_PERSON

SALESPERSON_ID	NAME	TITLE	PHONE	SUM_COMMISSIONS

Fig. 5(b)

Your rules should also check if the SUM_COMMISSIONS exceeds 100000; if it does, call a procedure NOTIFY_MANAGER(S_ID). Write both statement_ level rules in STARBVRST notation and ROW_LEVEL rules in Oracle. (12 Marks)

6. (a) What is a fragment of a relation? What are the main types of fragments? Why is fragmentation a useful concept in distributed database design? (10 Marks)

(b) Compare the two tier and three-tier client server architectures. (5 Marks)

(c) Discuss the naming problem in distributed database. (5 Marks)

Contd.... :

7. (a) What types of Indexes are built for a warehouse? Illustrate the uses for each with an example. **(8 Marks)**

   (b) What are the five types of knowledge produced from data mining? **(5 Marks)**

   (c) Describe the steps of building a warehouse. **(7 Marks)**

8. Write short notes on :

   (a) Association rules in data mining

   (b) Multimedia database application

   (c) Mobile database management issues

   (d) Distributed databases **(5×4=20 Marks)**

** * **

*****