B.E./B.Tech. DEGREE EXAMINATION, NOVEMBER/DECEMBER 2004.

Fourth Semester

Information Technology

IF 254 — DATABASE MANAGEMENT SYSTEMS

Time : Three hours                                    Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1.  Draw the flow diagram which shows the interaction of DBS between user and OS.

2.  Mention two main differences between file processing system and database management system.

3.  What is a derived attribute? Give an example.

4.  What is the difference between weak and strong entity set?

5.  Define the commit point.

6.  How does multilevel indexing improve the efficiency of searching an indexed file?

7.  Why is a relation with many NULLs considered to be bad?

8.  Define — repeating group and set occurrences with respect to network model.

9.  Explain two user defined control parameters used in association rule mining activity.

10. Explain semi–join with respect to distributed database.

PART B — (5 × 16 = 80 marks)

11. (i) Give three schema architecture of DBS. Why do we need mappings between different schema levels? How do different schema definition languages support this architecture? (8)

(ii) A database is being constructed to keep track of the teams and games of a sports league. A team has number of players. All of them do not participate in each game. It is desired to keep track of players participating in each game for each team, the position they played in that game, and the result of that game. Assuming reasonable attributes and relationships, construct E–R diagram, relational schemas for above problem. (8)

12. (a) (i) Consider the relational database given below :

LIVES (person–name, street, city)

WORKS (person–name, company–name, salary)

LOCATED–IN (company–name, city)

MANAGES (person–name, manager–name)

Give an expression in

(1) Relational Calculus and

(2) SQL for the following queries :

(A) Find all employees who live in the same city as the company they work for.

(B) Find all employees who does not work for bank XYZ. (6)

(ii) Give the algorithm to check whether the decomposition of the relation in lossless or not. Let $R = (A, B, C, D, E, F)$ and FDs are $A \rightarrow BC$, $C \rightarrow A$, $D \rightarrow E$, $F \rightarrow A$, $E \rightarrow D$. Is the decomposition of $R$ into $R1(A, C, D)$, $R2(B, C, D)$ and $R3(E, F, D)$ lossless? (10)

Or

(b) (i) Briefly explain the guidelines for good database design by taking suitable examples. (10)

(ii) Define Boyce–Codd Normal Form (BCNF). How does it differ from 3NF? (6)

13. (a) (i) Compare and contrast the methodology relating to hash method and join index method query evaluation. Give one example of each. (6)

(ii) Describe the timestamp ordering protocol for concurrency control. Comment on the schedule generated by this protocol. (10)

Or

N 1067

(b) (i) Consider a disk with block size $B = 512$ bytes. A block pointer is $P = 6$ bytes, and record pointer is 7 bytes long. A file has 40,000 employee records of fixed length. Each record has size 125 bytes with emp–number as a key field whose size is 10 bytes.

    (1) Consider the blocking factor, assume unspanned organization.

    (2) Construct primary index and secondary index on emp–number field.

    (3) Calculate index blocking factor, number of blocks and number of block accesses needed to access arbitrary record in above cases.

(10)

(ii) What are the different techniques for allowing a hash file to expand and shrink dynamically? Explain one of them. (6)

14. (a) (i) Explain different insertion and retention constraints in a network model. (10)

(ii) Discuss the techniques of data replication, fragmentation used in distributed database design by taking suitable examples. (6)

Or

(b) (i) Explain how recovery is handled in DDBMS. (10)

(ii) What are the similarities and differences between parallel and distributed DBMS? (6)

15. (a) (i) What are the two kinds of new data types supported in object–database systems? Given an example for each. (8)

(ii) What is versioning? Why is it important? What is the differences between versions and configurations? (8)

Or

(b) (i) Give Apriori algorithm to find the frequent item–sets in a transaction database. Explain the main characteristics of this algorithm. (10)

(ii) Explain the concept of data warehousing. (6)

———