Subject: DATABASE MANAGEMENT SYSTEMS

Time: 3 Hours   **JUNE 2009**   **Max. Marks: 100**

NOTE: There are 9 Questions in all.

- Question 1 is compulsory and carries 20 marks. Answer to Q. 1. must be written in the space provided for it in the answer book supplied and nowhere else.
- Out of the remaining EIGHT Questions answer any FIVE Questions. Each question carries 16 marks.
- Any required data not explicitly given, may be suitably assumed and stated.

Q.1 **Choose the correct or the best alternative in the following:** (2×10)

a. The _____ command is used to allocate an extent for the table in SQL

(A) MODIFY ALLOCATES   **(B)** MODIFY TABLE
(C) ALTER TABLE   **(D)** REDEFINE TABLE

b. The physical location of a record is determined by a mathematical formula that transforms a file key into a record location in

(A) a tree file   **(B)** an indexed file
(C) a hashed file   **(D)** a sequential file

c. Updating a database means

(A) revising the file structure
(B) reorganizing the database
(C) modifying or adding record occurrences
(D) all of the above

d. Goals for the design of the logical schema include

(A) avoiding data inconsistency
(B) being able to construct queries easily
(C) being able to access data efficiently
(D) All of the above

e. The way a particular application views the data from the database that the application uses is a :

(A) module   **(B)** relational model
(C) schema   **(D)** subschema

f. Two files may be joined into a third file if

(A) they have a row in common
(B) they have a field in common
(C) they have no record with the same value in the common field
(D) both **(B)** and **(C)**

g. An indexing operation

(A) sorts a file using a single key   **(B)** sorts file using two keys

(C) establishes an index for a file   **(D)** both **(B)** and **(C)**

h. The database environment has all of the following components except:

  (A) users   **(B)** separate files
  (C) database   **(D)** data-base administrator

i. Report generators are used to

  (A) store data input by a user   **(B)** retrieve information from files
  (C) answer queries   **(D)** both **(B)** and **(C)**

j. Consider the following set of functional dependencies on the scheme (A, B, C)

$$A \rightarrow BC$$
$$B \rightarrow C$$
$$A \rightarrow B$$
$$AB \rightarrow C$$

The canonical cover for this set is

  (A) $A \rightarrow BC$ and $B \rightarrow C$ **(B)** $A \rightarrow BC$ and $AB \rightarrow C$
  (C) $A \rightarrow BC$ and $A \rightarrow B$ **(D)** $A \rightarrow B$ and $B \rightarrow C$

Answer any FIVE Questions out of EIGHT Questions.
Each question carries 16 marks.

Q.2   a. List five responsibilities of a database management system. For each responsibility, explain the problems that would arise if the responsibility were not discharged. **(8)**

    b. List at least two reasons why database systems support data manipulation using a declarative query language such as SQL, instead of just providing a library of C or C++ functions to carry out data manipulation. **(4)**

    c. List some applications that most likely use a database system to store persistent data. **(4)**

Q.3   a. Consider the following relational database. Give a relational-algebra expression for each of the following queries: **(2+4+4)**

          employee (person-name, street, city)
          works (person-name, company-name, salary)
          company (company-name, city)
          manages (person-name, manager-name)

    (i)   Find the company with the most employees.
    (ii)   Find the company with the smallest payroll.
    (iii)   Find those companies whose employees earn a higher salary, on average, than the average salary at First Bank Corporation.

    b. List two reasons why null values might be introduced into the database. **(6)**

Q.4   a. Let the following relation schemas be given:
          $R = (A,B,C)$
          $S = (D,E, F)$
    Let relations $r(R)$ and $s(S)$ be given. Give an expression in SQL that is equivalent to each of the following queries.
    (i) $\Pi_A(r)$

(ii) $\sigma_{B=17}(r)$

(iii) $r \times s$

(iv) $\Pi_{A,F}(\sigma_{C=D}(r \times s))$ **(8)**

b. Consider the following relational database:

employee (employee name, street, city)
works (employee name, company name, salary)
company (company name, city)
manages (employee name, manager name)

Using SQL, define a view consisting of manager name and the average salary of all employees who work for that manager. Explain why the database system should not allow updates to be expressed in terms of this view. **(8)**

Q.5 a. Given a relational schema $r(A,B,C,D)$, does $A \rightarrow\rightarrow BC$ logically imply $A \rightarrow\rightarrow B$ and $A \rightarrow\rightarrow C$? If yes prove it, else give a counter example. **(8)**

b. Use Armstrong's axioms to prove the soundness of the decomposition rule. **(8)**

Q.6 a. What are the causes of bucket overflow in a hash file organization? What can be done to reduce the occurrence of bucket overflows? **(8)**

b. Why might the leaf nodes of a B$^+$ tree file organization lose sequentiality? Suggest how the file organization may be reorganized to restore sequentiality. **(8)**

Q.7 a. Pipelining is used to avoid writing intermediate results to disk. Suppose you need to sort relation r using sort–merge and merge-join the result with an already sorted relation s. Describe how the output of the sort of r can be pipelined to the merge join without being written back to disk. **(8)**

b. Why is it not desirable to force users to make an explicit choice of a query processing strategy? Are there cases in which it is desirable for users to be aware of the costs of competing query-processing strategies? Explain your answer. **(8)**

Q.8 a. Explain the distinction between the terms serial schedule and serializable schedule. **(5)**

b. What is a recoverable schedule? Why is recoverability of schedules desirable? Are there any circumstances under which it would be desirable to allow nonrecoverable schedules? Explain your answer. **(8)**

c. In multiple-granularity locking, what is the difference between implicit and explicit locking? **(3)**

Q.9 a. Explain the following terms: **(8)**

(i) Transaction rollback
(ii) Shadow paging

b. Explain the purpose of the checkpoint mechanism. How often should checkpoints be performed? How does the frequency of checkpoints affect
(i) System performance when no failure occurs
(ii) The time it takes to recover from a system crash
(iii) The time it takes to recover from a disk crash **(8)**