Computer Engineering Department, S V N I T, Surat
End-Semester Examinations, April 2019
B. Tech. II (CO) – 4th Semester
Course: Database Management System (CO204)

Date: 29th April 2019 | Time: 14:30 hrs to 17:30 hrs | Max Marks: 100

Instructions: 1. Be precise and clear in answering the questions.
2. Write your B. Tech. Admission No. and other details clearly on the answer books and B. Tech. Admission No. on the question paper.
3. Assume and write necessary data with proper justifications, if any.
4. Calculator is allowed.

Q. 1 **Suppose you are given the following information about a database for a chain of drug stores:** **[16]**

- a drugstore sells drugs prescribed by doctors to patients
- each drugstore in the chain is identified by a store name, address, and a phone number
- patients are identified by a patient id, and their names, addresses, and ages must be recorded
- doctors are identified by a doctor id. Each doctor's name, specialty, and years of experience must be recorded
- each drug is made by a pharmaceutical company and sold to the drugstore
  The drug's trade name identifies the drug uniquely from among the products of that company. For each drug, the trade name and formula must be recorded
- each pharmaceutical company is identified by name and has a phone number
- every patient has a primary doctor
- every doctor has at least one patient
- each drugstore sells several drugs and has a price for each. A drug could be sold at several drugstores, and the price could vary from one drugstore to another
- doctors prescribe drugs for patients. A doctor could prescribe one or more drugs for several patients, and a patient could obtain prescriptions from several doctors
- pharmaceutical companies have long-term contracts with drugstores. A pharmaceutical company can contract with several drugstores, and a drugstore can contract with several pharmaceutical companies. For each contract, you have to store a start date, an end date, and the text of the contract
- drugstores appoint a supervisor for each contract. There must always be a supervisor for each contract, but the contract supervisor can change over the lifetime of the contract

if a pharmaceutical company is deleted, you need not keep track of its products any longer

a) Draw an ER diagram that captures the above information and identify any constraints that are not captured by the diagram **[08]**

b) Define the relational schema corresponding to the entity sets and relationship sets. Underline the primary keys. **[08]**

Q. 2 **Answer following [Any Three]:** **[24]**

a) Suppose that extendable hashing is being used on a database file that contains records with the following search key values: 4, 2, 3, 5, 25, 7, 11, 17, 19, 23, 29, 31, 14, 16, 30, 18 **[08]**

i) Construct the hash structure for this file if the hash function is $h(x) = x \bmod 7$ and each bucket can hold three records

ii) Show how the structure from part a) changes after inserting a record with the search key value of 32 and then deleting the record with the search key value of 11.

b) Given Relations R= { A,B,C,D,E,F,G,H } and functional dependencies S as follows: **[08]**

S = {A → CD, ACF → G, AD → BEF, BCG → D, CF → AH, CH→ G, D → B, H → DEG}

Find all keys for R.

c) Given Relations R= { A,B,C,D,E,F,G,H } and functional dependencies S as follows: **[08]**

d) Consider the relation schema Membership for a library database as follows: [08]

Membership (MID, Name, Address, PhoneNum, ParentMID, ISBN, Title, Authors, BorrowDate, ReturnedDate, FineDue, FinePaid).

where ParentMID may have the values Null, Father_Name, Mother_Name or both.

Set of functional dependencies held in Membership table:

F = { MID → Name, Address, PhoneNum, ParentMID; ISBN → Title, Authors;
(MID, ISBN, BorrowDate) → ReturnedDate, FinePaid, FineDue }

Normalize the Membership schema to 3NF and show the steps.

Q. 3 For the three relations and their statistical information are given as below: [10]

Relation	Number of Records	Number of Pages
Student (sid, name, age, address)	10,000	1,000
Book (bid, title, author)	50,000	5,000
Checkout (sid, bid, date)	3,00,000	15,000

- There are 500 different authors and Student age ranges from 7 to 24.

Given the following SQL query: SELECT S.name FROM Student S, Book B, Checkout C WHERE S.sid = C.sid AND B.bid = C.bid AND B.author = 'Olden Fames'

a) Show a physical query plan for this query, assuming there are no indexes and data is not sorted on any attribute. Use nested loop Join. [Any one plan out of possible many plans] [04]

b) Compute the cost of this query plan and the cardinality of the result. [06]

OR

Q. 3 For the given two relations, answer the following: [10]

Relation	Record Size	Number of Tuples
Employee (emp_id, salary, age, dept_id)	20 bytes	20,000
Department (dept_id, budget, status)	40 bytes	5,000

- The dept_id attribute in employee is a foreign key of the department relation.
- The file system supports a page size of 4000 bytes and there are 12 buffer pages available to the database.
- Assume the number of page I/O's as the measure of a query's cost is used. The following indices exist:
  - a clustering (i.e. primary) index on the dept_id attribute in employee
  - a non-clustering (i.e. secondary) index on the *age* attribute in employee
  - a clustering index on the dept_id attribute in department

a) Consider the SQL query: [04]

select * from employee where age > 30.

Let N = the number of tuples retrieved with this query. For what values of N would a sequential table scan of the employee relation be cheaper than processing the query using the index? Explain your answer.

b) Consider the SQL query: [06]

select * from employee, department where employee.dept_id = department.dept_id

What evaluation plan would a query optimizer likely

Q. 4 **Answer the following [Any Five]:** **[10]**

1) Write the language types of Relational algebra and Structured Query Language (SQL).
2) There are m rows in relation S (A,...) and n rows are in relation T (F,...). Attribute A is the primary key of S and attribute F is a foreign key reference to S(A). Write the maximum number of rows that can be output as a result in case of a) left outer join and b) inner join.
3) Identify the transaction lock used to avoid the search of entire tree to determine whether the tree node/leaf can lock or not.
4) State the isolation level which disallows both dirty reads and nonrepeatable reads, but allows phantom reads.
5) Enlist three failure classifications that affects the consistency of DBMS.
6) Name the preemptive deadlock prevention technique and also write what will happen, if transaction T_4 requests a data item held by T_3 where the transactions T_2, T_3, T_4 having time-stamps 5, 10 and 15 respectively,

Q. 5 **Answer the following [Any Five]:** **[20]**

1) Explain the file organization which stores records of several different relations in one file.
2) Explain with example how the strict two-phase locking protocol avoids cascading rollback.
3) Justify the statement: Multiple-granularity locking protocol uses multiple lock modes to ensure serializability.
4) Draw and explain the data structure used to achieve freedom from starvation for lock requests.
5) Explain any one type of cursor used in PL/SQL.
6) Identify and explain the type of indexing which uses a) only sparse index and b) sparse and dense index both.
7) Bank is having database that contain 60,000 records and each record size is 200 bytes. Bank is having a system with memory block size of 2048 bytes. Calculate how many records can be stored in one block of memory and find how many blocks are required to store total bank records by following un-spanned mapping technique. If, we want to utilize system memory space then which mapping is suitable and why justify your answer?

Q. 6 Consider following twitter application schema that allows to post picture with tweet text as well as retweets and liking rate of other user's pictures: **[04]**

Tweet (uid, pid, liking_rate, txt), Users (uid, name), Picture (pid, author, img) Where Tweet.uid is a foreign key to Users, Tweet.pid is a foreign key to Picture, Picture.author is a foreign key to Users, Tweet.liking_rate ranges from 1 to 10, All attributes are NOT NULL.

a) Write SQL query to discover users' ID and name who have given a liking_rate of 8 or higher to 50 pictures or more. **[02]**

b) Write Relational Algebra expression for the following SQL query: **[02]**

select x.pid from picture x where not exists (select * from tweet y where x.pid = y.pid and y.liking_rate <5)

OR

a) Write SQL query to discover users' ID and name who have received at least one liking_rate of 10, from a user other than the author. **[02]**

b) Write Relational Algebra expression for the following SQL query: **[02]**

select distinct u.uid from Users u, Picture x, Tweet y where u.uid = x.author and x.pid = y.pid and

Q. 7 **Answer the following for the given schedule:** [08]

1	W3 (A)
2	R1 (A)
3	W1 (B)
4	R2 (B)
5	W3 (C)
6	R2 (C)

a) Write the conflicting instructions from the given schedule.

b) Identify the given schedule is conflict serializable or not? Also, write serial equivalent schedule, if the schedule is conflict serializable schedule.

c) Identify schedule is view serializable or not? Draw the relation of view and conflict serializable schedules.

d) Identify is there any deadlock or not?

Q. 8 **The DBMS crashes with the given undo log. Answer the following for the DBMS recovery:** [08]

1	<START T11>
2	<T11, W, w>
3	<T11, X, x>
4	<START T12>
5	<T12, Y, y>
6	<START T13>
7	<T13, Z, z>
8	<COMMIT T11>
9	<START CKPT (T12, T13)>
10	<T12, P, p>
11	<START T14>
12	<T14, Q, q>
13	<T13, R, r>
14	<COMMIT T13>
15	<COMMIT T12>
16	<END CKPT>
17	<COMMIT T14>

a) Write the recovery steps for the log recovery with checkpoint.

b) Which log entries are read?

c) Which transactions are undone and why?

d) Which data do the DBMS change and why?

* * *