COMP 5 – 5 (RC)

T.E. (Comp.) (Semester – V) (RC) Examination, November/December 2015

DATABASE MANAGEMENT SYSTEMS

Duration : 3 Hours

Total Marks : 100

Instructions : *1) Answer **five** questions in **all**, selecting at least **one** question from **each** Module.*

*2) Assume necessary data, **wherever** required.*

MODULE – I

1. a) The organisers of the EXAM 2015 international multi-conference need to keep track of a large collection of workshops associated with the event. Initial requirements analysis brings out the following information about what needs to be recorded. 10

- Each workshop has a name, and happens on a particular date – or dates, as some workshops last more than one day.
- There are several participants, each of which may sign up to one or more workshops.
- For each participant, it is important to record their name, email address, and the workshops which they wish to attend.
- There are a number of meeting rooms at the conference venue, each of a fixed capacity. Meetings rooms are identified by a floor and room number.
- Every workshop needs an allocated meeting room; where a workshop lasts for two days, it will use the same room on both days.

A) Draw an entity-relationship diagram suitable for representing this information, in particular the connections between participants, workshops, rooms and dates.

P.T.O.

Scanned by CamScanner

B) Further analysis reveals additional requirements. However, not all of these can be captured easily in an ER diagram.

- Each workshop must have an identified organiser among the conference participants.
- No participant may register for two workshops on the same day.
- Every participant may register for at least one workshop.

i) Identify two of these which can be captured in an ER diagram.

ii) For those two, show the additions required to your diagram.

b) What are the various levels of schema abstraction in DBMS ? Why is it required ? 4

c) Explain the responsibilities of the following subsystems of database (any 2) : 4

i) Query processor

ii) Storage manager

iii) Transaction manager.

d) List any two advantages of DBMS. 2

2. a) Explain the distinction between condition defined and user defined constraints. Which of these constraints can the system check automatically ? Explain your answer. 6

b) For each of the following concepts give a brief description of what it means, how it is represented in the ER diagram and provide an example. 6

i) Composite key

ii) Total participation

iii) Key constraint.

c) What do you mean by identifying relationship ? 2

d) Explain the following terms with the help of an example. Also discuss why are they required ? 6

I) Aggregation

II) Specialization.

Scanned by CamScanner

MODULE – II

3. a) Differentiate between a relation schema and a relation instance. Define the term relation degree. What are domain constraints ? 3

b) Let the following relational schema be given : 10

EMPLOYEE (SSN, Name, Age, salary, Dno)

WORK_ON (Project#, SSN)

PROJECT (Project#, pname, plocation)

For each of the following queries, give an expression in relational algebra :

i) Display the names of employees working on project at "Delhi".

ii) Retrieve the name of the working on maximum projects.

iii) For each project retrieve the number of employees working on that project.

iv) Find the names of the employees who work on all the projects that "Mr. ABC" works on.

v) For each department that has more than 5 employees, retrieve the department number and the number of its employees who are making more than 50,000/-.

c) What do you mean by safety of expressions in DRC ? 2

d) Check whether the following set of functional dependencies are equivalent or not. 5

F1 = {A->C, AC->D, E->AD, E->H} F2 = {A->CD, E->AH}

4. a) Consider the following relations : 10

Student (snum: integer, sname: string, major: string, level: string, age: integer)

Class (name: string, meets at: time, room: string, fid: integer)

Enrolled (snum: integer, cname: string)

Faculty (fid: integer, fname: string, deptid: integer)

Scanned by CamScanner

The meaning of these relations is straightforward; for example, Enrolled has one record per student-class pair such that the student is enrolled in the class. Write the following queries in SQL. No duplicates should be printed in any of the answers.

i) Find the names of all Juniors (Level = JR) who are enrolled in a class taught by Prof. 'AB'.

ii) Find the age of the oldest student who is either a History major or is enrolled in a course taught by Prof. BB.

iii) Find the names of all classes that either meet in room R128 or have five or more students enrolled.

iv) Find the names of all students who are enrolled in two classes that meet at the same time.

v) Find the names of students who are enrolled in the maximum number of classes.

b) Define a view. How is it different from a table ? Why are they used ? Provide an example of a view which cannot be updated and explain why it cannot be updated ? 5

c) Consider the following set of functional dependencies :

F = {A->BCE, AB->DE, BI->J}. Compute the canonical cover for F. 5

MODULE – III

5. a) How does hybrid hash join improve on the basic hash join algorithm ? How many buffer blocks are required to implement hash join and hybrid hash join ? Explain. 7

b) Define the term functional dependency. Why are some functional dependencies called trivial ? 6

Scanned by CamScanner

c) Consider a relation R with five attributes ABCDE. You are given the following dependencies : 7

$A \rightarrow B$, $BC \rightarrow E$ and $ED \rightarrow A$

i) List all keys for R.

ii) Is R in 3NF ?

iii) Is R in BCNF ?

6. a) Consider the following relational schema : 10

Student(rollno, name, sem, dno) dept(dno, dname)

Subject(subid, sname, sem) marks_obt(rollno, subid, marks)

Write a query to retrieve the names of Computer Engineering students who obtained marks greater than 50 in the subject 'DBMS'.

Draw the query graph and the initial query tree for the above query and then show (each step) how the query can be optimised. Provide the final optimised query.

b) Suppose you have a file with 20,000 blocks and you have five buffer blocks. Answer the following questions, assuming that general external sorting algorithm is used : 10

i) How many runs will you produce in the first pass ?

ii) How many passes will it take to sort the file completely ?

iii) What is the total I/O cost of sorting the file ?

iv) How many buffer pages do you need to sort the file completely in just two passes ?

Scanned by CamScanner

MODULE – IV

7. a) Explain the following terms using an example : 10

i) Partial schedule

ii) Recoverable schedule

iii) Cascadeless schedule

iv) Strict schedule

v) Consider the following schedule.

r1(x);w1(x);r2(x);w2(x);r2(y);w2(y);r1(z);w1(z);c1;c2;

Check whether the given schedule is cascadeless. If not, write down the minimum changes required to convert it into cascadeless schedule.

b) Consider the following actions taken by transaction T1 on database objects X and Y :

R(X), W(X), R(Y), W(Y) 4

i) Give an example of another transaction T2 that, if run concurrently to transaction T without some form of concurrency control, could interfere with T.

ii) Explain how the use of Strict 2PL would prevent interference between the two transactions.

c) Bring out the differences between the basic two-phase locking, strict two phase locking and rigorous two-phase locking using a appropriate example. 6

Scanned by CamScanner

8. a) What is the difference between conflict serializability and view serializability ? Which one of them is a stricter form ? Under what conditions conflict and view serializability are similar. Give example of the schedule which follows conflict serializability but not view serializability or vice versa. 10

b) Briefly explain why concurrency control is required. Give an example where adding lock/unlock instruction to the transactions (not in 2-phase order) does not give us the serializable schedule but by following 2 phase locking protocol we get a serializable schedule devoid of deadlocks. 6

c) What is the difference between wound wait and wait-die techniques ? 4

Scanned by CamScanner

T.E. (Comp.) (Semester – V) (RC) Examination, May/June 2016
DATABASE MANAGEMENT SYSTEMS

Duration : 3 Hours

Max. Marks : 100

Instruction : *Answer **any five** questions by selecting atleast **one** question from **each** Module.*

MODULE – I

1. a) A university registrar's office maintains data about the following entities : 10

a) Courses, including number, title, credits, syllabus and prerequisites ;

b) Course offerings, including course number, year, semester, section number, instructor(s), timings and classroom ;

c) Students, including student-id, name and program ; and

d) instructors, including identification number, name, department and title. Further, the enrollment of students in courses and grades awarded to students in each course they are enrolled for must be appropriately modeled. Construct an E-R diagram for the registrar's office. Document all assumptions that you make about the mapping constraints. Convert the resulting ER into tables, clearly specifying the constraints.

b) What do you mean by aggregation ? When is it required ? Provide two examples of aggregation.

c) What are advantages of using DBMS over file systems ?

2. a) Explain the following terms with respect to an ER diagram (provide examples).

i) Participation constraint

Scanned by CamScanner

ii) Key constraint

iii) Integrity constraints.

b) What is an identifying relationship ? How is it represented in an ER diagram ? Provide examples.

c) What are the various levels of schema abstraction in DBMS ? Why is it required ?

d) What is data independence ? What role does the data independence play in database design ? How is physical data independence different from logical data independence ?

MODULE – II

3. a) What do you mean by safety of expression in TRC ?

b) What are views ? Can views be updated ? Justify using examples.

c) Suppose that we have the following three tuples in a legal instance of a relation schema S with three attributes ABC (listed in order) : (1, 2, 3), (4, 2, 3) and (5, 3, 3).

i) Which of the following dependencies can you infer does not hold over schema S ?

a) $A \rightarrow B$ b) $BC \rightarrow A$ c) $B \rightarrow C$

ii) Can you identify any other dependencies that hold over S ?

d) Consider the following database :

Faculty (fid, fname, faddress, fage, salary)

Subject (subid, Name, semester)

teacher (fid, subid, year)

Scanned by CamScanner

Answer the following queries in relational algebra.

i) Retrieve the name of the faculties along with the subject they teach.

ii) Retrieve the number of courses taught by each faculty. 4

iii) Retrieve the name of the faculty teaching maximum number of courses in a given year. 4

iv) Retrieve the name of the faculties who have taught the same course more than once.

4. a) Consider the attribute set R = ABCDEG and the FD set $F=\{AB \rightarrow C, AC \rightarrow B, AD \rightarrow E, B \rightarrow D, BC \rightarrow A, E \rightarrow G\}$. 10

i) Name the strongest normal form that is not violated by the relation containing these attributes.

ii) Decompose it into a collection of BCNF relations if it is not in BCNF

iii) Check if the following decomposition of R into three relations {ABC, ACDE, ADG} with the same set of dependencies F, is

A) dependency-preserving.

B) Lossless-Join.

b) What do you mean by foreign key ? Explain using an example. 2

c) Consider the following database : 8

Person (name text, age int, gender text) ;

Frequents (name text, pizzeria text) ;

Eats (name text, pizza text) ;

Serves (pizzeria text, pizza text, price decimal) ;

Scanned by CamScanner

Answer the following queries in SQL

i) Find the names of all females who eat at least one pizza served by Straw Hat.

ii) Find all pizzerias that serve at least one pizza for less than $10 that both Amy and Fay eat.

iii) Find the names of all people who eat at least one pizza served by Dominos but who do not frequent Dominos.

iv) Find the age of the oldest person (or people) who eat mushroom pizza.

MODULE – III

5. a) Assume the following relations :

Books (DocId, Title, Publisher, Year)

Student (StId, StName, Major, Age)

Author (AName, Address)

borrows (DocId, StId, Date)

has-written (DocId, AName)

Consider the query "find the name and age of students who are 'CS' majors and have borrowed book(s) written by author 'Navathe' ".

Draw the query graph and the initial query tree for the above query and then show (each step) how the query can be optimised. Provide the final optimised query.

b) Consider a relation R with four attributes ABCD. Given the following sets of FDs, assuming those are the only dependencies that hold for R, do the following :

a) Identify the candidate key(s) for R.

b) Identify the best normal form that R satisfies (1NF, 2NF, 3NF, or BCNF).

Scanned by CamScanner

c) If R is not in BCNF, decompose it into a set of BCNF relations that preserve the dependencies.

$$F = \{AB \rightarrow C, AB \rightarrow D, C \rightarrow A, D \rightarrow B\}$$

c) What do you mean by left deep trees ? Why are they preferred by query optimizers ? 4

6. a) Suppose you need to sort a relation of 20 GB, with 4KB block size, using memory size of 20 MB. Suppose the cost of a seek is 5ms, while the disk transfer rate is 40 megabytes per second. Find the cost of sorting the relation, in seconds. Assume bb = 1, where bb is the buffer blocks allocated to each run during merge phase. What is the max/ideal value we can choose for bb ? What is the cost of sorting the relation, in seconds, when we consider bb equal to this ideal value ? 10

b) What do you mean by Multivalued Dependency (MVD) ? Briefly explain MVD's using a example relation and then normalize the relation in 4NF. 7

c) Briefly explain the differences between pipelining and materialization. 3

MODULE – IV

7. a) What is the difference between the 2 phase locking protocol and the tree protocol ? 4

b) Give an example where adding lock/unlock instruction to the transactions (not in 2-phase order) does not give us the serializable schedule but by following 2phase locking protocol we get a serializable schedule devoid of deadlocks.

Scanned by CamScanner

c) What is the difference between strict and basic two phase locking ? Provide a schedule which is allowed under basic 2PL but not under strict 2PL.

d) What are the various deadlock prevention techniques used in databases ?

8. a) Consider a database with objects X and Y and assume that there are two transactions T1 and T2. Transaction T1 reads objects X and Y and then writes object X. Transaction T2 reads objects X and Y and then writes objects X and Y.

i) Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a write-read conflict.

ii) Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a write-write conflict.

iii) Check whether strict 2PL will allow/disallow above two schedules.

b) Consider the following classes of schedules, serializable, conflict-serializable, view-serializable, recoverable and strict. For each of the following schedules, state which of preceding classes it belongs to. If you cannot decide whether a schedule belongs in a certain class based on the listed actions, explain briefly. The actions are listed in the order they are scheduled and prefixed with the transaction name. If a commit or abort is not shown, the schedule is incomplete ; assume that abort or commit must follow all the listed actions. 10

i) T1 : R(X), T1 : R(Y), T1 : W(X), T2 : R(Y), T3 : W(Y), T1 : W(X), T2 : R(Y)

ii) T2 : R(X), T3 : W(X) : T3 : Commit, T1 : W(Y), T1 : Commit, T2 : R(Y), T2 : W(Z), T2 : Commit

Scanned by CamScanner

c) Consider the data given below and determine whether deadlock exists : 4

Transaction	Data items locked by transaction	Data item transaction is waiting for
T1	X2	X1, X3
T2	X3, X10	X7, X8
T3	X8	X4, X5
T4	X7	X1
T5	X1, X5	X3
T6	X4, X9	X6
T7	X6	X5

If it exists, what measures can be taken for handling the deadlock ?

Scanned by CamScanner

T.E. (Comp.) (Semester – V) (RC) (2007 – 08) Examination, Nov./Dec. 2017

DATABASE MANAGEMENT SYSTEMS

Duration : 3 Hours

Total Marks : 100

Instructions : *1) Attempt **any five** questions by selecting at least **one** question from **each** Module.*

*2) Make **necessary** assumptions, **wherever** required and **clearly** specify the assumptions made.*

MODULE – 1

1. a) Design a database with a schema that captures all the information that galleries need to maintain. Galleries keep information about artists, their names (which are unique), birthplaces, age and style of art. For each piece of artwork, the artist, the year it was made, its unique title, its type of art (e.g., painting, lithograph, sculpture, photograph) and its price must be stored. Pieces of artwork are also classified into groups of various kinds, for example, portraits, still lifes, works by Picasso, or works of the 19th century; a given piece may belong to more than one group. Each group is identified by a name (like those above) that describes the group. Finally, galleries keep information about customers. For each customer, galleries keep their unique name, address, total amount of dollars they have spent in the gallery and the artists and groups of art that each customer tends to like. Draw an entity-relationship diagram suitable for representing this information. Convert this ER diagram into tables and represent the constraints on each table.

Scanned by CamScanner

b) What is a query processor ? What are the various components of the query processor ?

c) What are the various levels of schema abstraction in DBMS ? Why is it required ?

2. a) Explain what is attribute inheritance in ER diagram generalization ?

b) An ER diagram can be viewed as a graph. What do the following mean in terms of the structure of an enterprise schema ?

i) The graph is disconnected

ii) The graph has a cycle.

c) List three significant differences between the file processing system and a DBMS.

d) Give an example for the following :

i) Disjoint and complete specialization.

ii) Overlapping and partial specialization.

Give an justification for the example.

e) Give an example of an entity which is existent dependent on another entity but is not a weak entity. Justify.

MODULE – 2

3. a) Consider the following schema :

Faculty (fid, fname, fqualifications, fexperience, salary, deptname)

Course (cid, cname, semester)

Teaches (fid, cid, year)

Answer the following queries in relational algebra :

i) Give the details of the faculty with experience greater than 10 years and teaching a course 'DBMS' in the year 2017.

ii) Identify the faculty who have taught all the courses for semester 5 in the database.

Scanned by CamScanner

iii) List the faculty who have not taught any course in the year 2017.

iv) Find the faculty who has taught maximum courses. (If the same course is taught twice it should be counted as one).

v) Retrieve the names of the faculty who have taught the same course more than once.

b) Consider the following set of functional dependencies : 6

F = {A-> BCE, AB->DE, BI->J}. Compute the canonical cover for F.

c) What do you mean by unsafe expressions in Domain Relational Calculus (DRC) ? Specify the following relational algebra operation in DRC. 4

$\prod_{\langle A,B \rangle}(\sigma_{C>A}(R(A,B,C)))$

4. a) Consider the following relational schema. An employee can work in more than one department; the pct_time attribute of the works relation shows the percentage of time that a given employee works in a given department. 10

Emp (eid, ename, age, salary)

Works (eid, did, pct_time)

Department (did, dname, budget, managerid)

Write the following queries in SQL :

i) Find the details of employees who works for a department with budget greater than 50,000/-

ii) Retrieve the name and salary of the managers.

iii) Find the department with the largest budget also retrieve the name of the manager for the department.

iv) Find the employee with the second highest salary.

v) Find the names of employees who are managing atleast two departments and earning salary over 50,000/-

Scanned by CamScanner

b) User 1 is the owner of relation R.

Provide SQL commands for the following :

i) User 1 wants to grant the select privileges on R to user 2 with a power to pass on this privilege to other users.

ii) User 1 wants to grant the update privilege on the A and D attributes of the relation R to user 3.

iii) User 1 wants to go for a holiday. So he wants user 4 to handle all his tasks once he leaves. What are the privileges he grants to user 4 and what commands to execute once he is back from his holidays.

c) When are two sets of functional dependencies equivalent ? Consider the following two sets of functional dependencies :

$F1 = \{A \rightarrow B, AB \rightarrow C, D \rightarrow AC, D \rightarrow E\}$, $F2 = \{A \rightarrow BC, D \rightarrow AE\}$.

Check whether they are equivalent.

MODULE – 3

5. a) Consider the universal relation Class(course_id, title, dept_name, credits, sec_id, semester, year, building, room_number, capacity, time_slot_id) and the set of functional dependencies.

F = {course_id_>title,dept_name,credits;

building,room_number->capacity;

course_id,sec_id,semester,year-> building,room_number,time_slot_id}

What is the key of relation class ? Decompose Class into 2NF and then 3NF relations. Is the decomposition is in BCNF ? If no, give the corresponding BCNF decomposition. Is the final decomposition in BCNF lossless and dependency preserving decomposition ?

Scanned by CamScanner

b) Consider the query to find the names of all the employees who works on a project controlled by 'ADMIN' department for more than 10 hours per week. Draw the initial query tree for the above query and then show how the query can be translated to the optimized query. Show all intermediate steps. 5

When would we prefer a unoptimized query over a optimized one ? 8

Assume the schema as

employee(ename,street,city,dno);

department(dname,dno,mgrname)

Works(ename,pname,hours);

Project(pname,controlling_dnum);

6. a) What are left deep trees ? What are the advantages of the left deep trees during query optimization ? 5

b) A file of 1024 blocks is to be sorted using a buffer size of 5 blocks. How many merge passes are required in the merge phase of the external merge sort algorithm ? Calculate the total cost of sorting the file if the seek time is 0.4 ms, transfer time is 1.1 ms and only one block is allocated to each run during merge phase. 4

c) Briefly explain the differences between pipelining and materialization. 6

d) Let relations r1(A, B, C) and r2(C, D, E) have the following properties : r1 has 50000 tuples, r2 has 25000 tuples, 50 tuples of r1 fit on one block and 25 tuples of r2 fit on one block. Assume that the data files are not sorted and the number of main memory buffer blocks are 50. Attribute C is the primary key for relation r2. 4

6

Scanned by CamScanner

i) Estimate the minimum number of block transfers and seeks required for the r1 join r2 using Block Nested Loop join. What is the minimum number of buffer pages required for this cost to remain unchanged ?

ii) How many tuples does the natural join of r and s produce at most ? Justify.

MODULE – 4

7. a) What is a schedule ? Explain the following types of schedules with examples :

i) Serial schedule

ii) Serializable schedule

iii) Strict schedule

iv) Cascadeless schedule.

Give an example of a serializable schedule that is not strict and vice versa.

b) What is the difference between strict and basic two phase locking ? Provide a schedule which is allowed under basic 2PL but not under strict 2PL. 6

c) What do you mean by cautious waiting ? Does it avoid deadlocks ? If yes, explain how ? 4

8. a) What is the difference between wound-wait and wait-die techniques ? 4

b) Draw a state diagram and discuss the typical states that a transaction goes through during execution. 4

c) When do we say that a schedule is view serializable ? Give an example of the schedule which is conflict serializable but not view serializable. 6

d) Produce a w
whether dea

Transac
T1
T2
T3
T4
T5
T
T

Discuss

1) Tim

2) De

Scanned by CamScanner

d) Produce a wait-for graph for the following transaction scenario and determine whether deadlock exists : 6

Transaction	Data item locked by transaction	Data item transaction is waiting for
T1	X2	X1
T2	X3, X10	X7, X8, X4
T3	X8	X4
T4	X7	X1
T5	X1, X5	X3
T6	X4, X9	X6, X2
T7	X6	X5

Discuss two general techniques for handling deadlock :

1) Time out

2) Deadlock detection and recovery.

Scanned by CamScanner

T.E. (Comp.) (Semester – V) (RC) Examination, Nov./Dec. 2016
DATABASE MANAGEMENT SYSTEMS

Duration : 3 Hours Total Marks : 100

Instructions: *1) Answer **any five** questions by selecting atleast **one** question from **each** Module.*

*2) **Assume** necessary data, **wherever** required.*

MODULE – I

1. a) Assume we have the following application that models soccer teams, the games they play and the players in each team. In the design, we want to capture the following : **10**

We have a set of teams, each team has an ID (unique identifier), name, main stadium and to which city this team belongs.

- Each team has many players and each player belongs to one team. Each player has a number (unique identifier), name, Date of Birth, start year and shirt number that he uses.
- Teams play matches, in each match there is a host team and a guest team. The match takes place in the stadium of the host team.
- For each match we need to keep track of the following :
  - o The date on which the game is played
  - o The final result of the match
  - o The players participated in the match. For each player, how many goals he scored, whether or not he took yellow card and whether or not he took red card.
  - o During the match, one player may substitute another player. We want to capture this substitution and the time at which it took place.
- Each match has exactly three referees. For each referee we have an ID (unique identifier), name, Date of Birth, years of experience. One referee is the main referee and the other two are assistant referee.

Design an ER diagram to capture the above requirements. State any assumptions you have that affects your design. Make sure cardinalities and primary keys are clear.

Scanned by CamScanner

b) What do you mean by data model ? Explain any two data models.

c) What do you mean by consistency constraints ? Explain the various types of constraints.

2. a) Explain the various data abstraction levels. What is physical data independence ?

b) Explain the following constraints on generalization :

i) Completeness constraint

ii) Disjoint v/s overlapping constraint

iii) Condition defined v/s user defined constraint

c) What is a storage manager in DBMS ? Explain the various components of storage manager.

MODULE – II

3. a) What do you mean by a functional dependency ? Do all the tuples in the relation should satisfy these dependencies ? Justify.

Consider a relation R(A, B, C, D) with functional dependency $AB \rightarrow C$ and $CD \rightarrow E$. Suppose there are atmost 4 different values for each of A, B and D. What is the maximum number of different values for E ? Explain. What is the key for relation R.

b) Consider the following relations :

Doctor(SSN,FirstName, LastName, Specialty, YearsOfExperience,PhoneNum)

Patient(SSN, FirstName, LastName, Address, DOB, PrimaryDoctor_SSN)

Medicine(TradeName, UnitPrice, GenericFlag)

Prescription(Id, Date, Doctor_SSN, Patient_SSN)

Prescription_Medicine(Prescription Id, TradeName, NumOfUnits)

Scanned by CamScanner

Answer the following queries in Relational Algebra.

i) List the first and last name of doctors who are not primary doctors to any patient.

ii) For medicines written in more than 24 prescriptions, report the trade name and the total number of units prescribed.

iii) List the SSN of patients who have 'Aspirin' and 'Vitamin' trade names in one prescription.

iv) List the SSN of distinct patients who have 'Aspirin' prescribed to them by doctor named 'James Smith'.

v) List the first and last name of patients who have no prescriptions written by doctors other than their primary doctors.

c) Using the same schema as in 3 b) answer the following query in Tuple relational calculus. 3

List the first and last name of patients who has primary doctor named 'James Smith'.

4. a) Consider the following relations : 10

Employee (emp_no, name, skill, pay_rate)

Position(posting_no, skill)

Duty_allocation (posting_no, emp_no,day,shift)

Write the following queries in SQL.

i) Retrieve employees whose rate of pay is more than or equal to the rate of pay of employee 'XYZ'.

ii) Find the employees with the lowest pay rate.

iii) Find the name of all employees who are assigned to all positions that require a Chef's skill.

iv) Get a list of employees who are not assigned a duty.

v) Get the names of the employees working on at least two dates.

Scanned by CamScanner

b) Let relation R(A,B,C,D) satisfy the following functional dependencies :

$A \rightarrow B, B \rightarrow C, C \rightarrow A$

Call this set S1. A different set S2 of functional dependencies is equivalent to S1 if exactly the same FDs follow from S1 and S2. Which of the following sets of FDs is equivalent to the set above ?

i) $C \rightarrow B, B \rightarrow A, A \rightarrow C$

ii) $A \rightarrow BC, C \rightarrow AB$

iii) $B \rightarrow A, B \rightarrow C, C \rightarrow B$

iv) $A \rightarrow B, B \rightarrow A, C \rightarrow A$.

c) Consider the following set of functional dependencies :

$F = \{A \rightarrow BCE, AB \rightarrow DE, BI \rightarrow J\}$. Compute the canonical cover for F.

MODULE – III

5. a) Find a BCNF decomposition of the relation schema :

SHIPPING (ship,capacity, shipDate,cargo,value) with the following set of functional dependencies :

ship → capacity

shipDate → Cargo

cargo,capacity → Value

Specify the key of the relation. Also mention whether the decomposition is lossless and dependency preserving or not.

b) Consider the relations r1(A, B, C), r2 (C, D, E) and r3 (E, F), with primary keys A, C and E, respectively. Assume that r1 has 1000 tuples, r2 has 1500 tuples and r3 has 750 tuples. Estimate the size of r1 ⋈ r2 ⋈ r3 and give an efficient strategy for computing the join.

c) Explain the fourth normal form.

Scanned by CamScanner

6. a) Consider the following relational schema : **10**

Project(pno, pname, plocation)

Works_on(empno,pno)

Employee (empno, ename, salary)

Write a query to retrieve the name of all employees working on 'research' project at 'Noida' and earning salary over 50,000/-

Draw the query graph and the initial query tree for the above query and then show (each step) how the query can be optimised. Provide the final optimised query.

b) Consider two relations R1(A, B, C) and R2(A, D). A is the primary key for R2 and R1. A is a foreign key for R1. R1 and R2 are stored as an unordered file (blocks are assumed to be 100% full); R1 contains 200000 tuples that are stored in 500 blocks and R2 has 50000 tuples that are stored in 1000 blocks. Moreover, assume that only a small buffer size of 8 blocks is available. Based on these assumptions answer the following questions (indicate, if you assume in your computations that the output relation is written back to disk or not) : **10**

i) How many tuples will the output relation R = R1 ⋈ R2 contain ? Explain.

ii) What is the cost for implementing the natural join using the block-nested loop join ?

iii) Now assume that either a hashed index on A of R1 or a hashed index on A of R2 is available (assume that there are no overflow pages). Compute the cost for using the index nested loops join using the index for R1. A and for using the index nested loops join for R2.A.

iv) Is it possible to apply the hash-join in this case (explain your answer!) ?

v) Which of the 4 proposed implementations (above) is the best ?

Scanned by CamScanner

MODULE – IV

7. a) Consider the following classes of schedules: conflict-serializable, view-serializable, recoverable, avoids-cascading-aborts and strict. Explain your answer. For the following schedule, state which of the preceding classes it belongs to. If you cannot decide whether the schedule belongs in a certain class based on the listed actions, explain briefly.

The actions are listed in the order they are scheduled and prefixed with the transaction name.

T1:R(X), T1:R(Y), T2:R(Z), T1:W(X) T2:R(X) T2:W(Y) T2: commit, T3:R(A), T1:R(Z), T1: Commit, T3:R(Z) T3:W(A), T3:Commit

b) Consider the following actions taken by transaction T1 on database objects A and B:

R(A), W(A), R(B), W(B)

i) Give an example of another transaction T2 that, if run concurrently to transaction T1 without some form of concurrency control, could interface with T1.

ii) Explain how the use of Strict 2PL would prevent interference between the two transactions.

iii) Strict 2PL is used in many database systems. Give two reasons for its popularity.

c) Bring out the differences between the basic two-phase locking and strict two phase locking using a appropriate example.

8. a) For each of the following locking protocols, assuming that every transaction follows that locking protocol, state which of these desirable properties are ensured : serializability conflict-serializability, recoverability, avoidance of cascading aborts. Justify your answer.

i) Always obtain an exclusive lock before writing; hold exclusive locks until end-of-transaction. No shared locks are ever obtained.

ii) In addition to (i), obtain a shared lock before reading; shared locks can be released at any time.

iii) As in (ii) and in addition, locking is two-phase.

iv) As in (ii) and in addition, all locks held until end-of-transaction.

Scanned by CamScanner

b) Suppose that a DBMS recognizes increment, which increments an integer valued object by 1 and decrement as actions, in addition to reads and writes. A transaction that increments an object need not know the value of the object; increment and decrement are versions of blind writes. In addition to shared and exclusive locks, two special locks are supported: An object must be locked in I mode before incrementing it and locked in D mode before decrementing it. An I lock is compatible with another I or D lock on the same object, but not with S and X locks. Illustrate how the use of I and D locks can increase concurrency. (Show a schedule allowed by Strict 2PL that only uses S and X locks. Explain how the use of I and D locks can allow more actions to be interleaved, while continuing to follow Strict 2PL.) 6

c) Show that there are schedules that are possible under the two phase locking protocol, but are not possible under the timestamp protocol and vice versa. 6

Scanned by CamScanner

T.E. (Comp.) (Semester – V) Examination, May/June 2017
DATABASE MANAGEMENT SYSTEMS

Duration : 3 Hours Total Marks : 100

Instructions : *1)* ***All*** *questions are* ***compulsory****.*
2) ***Assume*** *necessary data,* ***wherever*** *required.*

MODULE – I

1. a) Consider a database in which researchers submit their research papers for consideration. Reviews by reviewers are recorded for use in the paper selection process. The database system caters primarily to reviewers who record answers to evaluation questions for each paper they review and make recommendations regarding whether to accept or reject the paper. Authors of papers are uniquely identified by e-mail id. First and last names are also recorded. Each paper is assigned a unique identifier by the system and is described by a title, abstract and the name of the electronic file containing the paper. A paper may have multiple authors, but one of the authors is designated as the contact author. Reviewers of papers are uniquely identified by e-mail address. Each reviewer's first name, last name, phone number, affiliation and topics of interest are also recorded. Each paper is assigned between two and four reviewers. A reviewer rates each paper assigned to him or her on a scale of 1 to 10 in four categories: technical merit, readability, originality and relevance to the conference. Finally, each reviewer provides an overall recommendation regarding each paper. Each review contains two types of written comments: one to be seen by the review committee only and the other as feedback to the author(s). Design an Entity-Relationship diagram for the above database, convert the ER diagram to relational schema. **10**

b) What are the various types of database users ? **4**

c) What is a storage manager ? What are the various data structures implemented by the storage manager as part of the physical system implementation ? **6**

Scanned by CamScanner

2. a) What is aggregation ? Explain using an example and convert the aggregation example to relational tables (schema).

b) An ER diagram can be viewed as a graph. What do the following mean in terms of the structure of an enterprise schema ?

i) The graph is disconnected

ii) The graph has a cycle.

c) An ER enterprise schema may define certain constraints to which the contents of a database must conform. Briefly explain these constraints.

MODULE – II

3. a) What are views ? Can views be updated ? Justify using examples.

b) Given 2 sets F1 and F2 of functional dependencies for a relation R, check whether they are equivalent.

$F1 = \{A \rightarrow B, AB \rightarrow C, D \rightarrow AC, D \rightarrow E\}$

$F2 = \{A \rightarrow BC, D \rightarrow AE\}$.

c) Consider the following relations for a database of business trip of sales persons in a sales office :

Sales – person (ssn, name, start-year, dept_no)

Trip (ssn,from_city, to_city, departure-date, return_date, trip_id)

Expense (Trip_id, account_no, amount)

Specify the following queries in relational algebra:

i) For every salesperson, find the total amount spent on the trips.

ii) Find the most expensive trips.

iii) For every department, find the person spending maximum amount on the trip and also display the amount spent.

iv) Find the person going maximum number of times on trips.

v) Find the person with maximum number of trips in the last 5 years.

4. a) Consider the following relations for a database of business trip of sales persons in a sales office :

Person (driverID, name, address)

Car (license, model, year)

Accident (report no, data, location)

Owns(driverID, license)

participated (driverID, license, reportno, damage_amt)

Scanned by CamScanner

Specify the following queries in SQL :

i) Find the total number of accidents in which the cars belonging to 'Ram' were involved.

ii) Find the total number of people who owned cars that were involved in accidents in 2000.

iii) Find the car involved in an accident with the maximum damage amount.

iv) For each car involved in more than five accidents, find the total damage amount.

v) For each person in the database, retrieve the name, address, model and year of each car owned by that person.

b) Consider the following schema :

Faculty (fid, fname, fqualifications, fexperience)

Course (cid, cname, semester)

Teaches (fid, cid, year)

Answer the following queries using TRC and QBE.

Retrieve the names of the faculties who have taught 'DBMS' atleast once. **4**

c) Consider the relation Took(sID, name, courseNum, title). Students have a unique student ID and (possibly not unique) name; courses have a unique course number and (possibly not unique) title. Each tuple in the relation encodes the fact that the given student took the given course. What are all of the functional dependencies for relation Took ? **6**

MODULE – III

5. a) Consider the join $R \bowtie_{R.a = S.b} S$, given the follwoing information about the relations to be joined. The cost metric is the number of page I/Os unless otherwise noted and the cost of writing out the result should be uniformly ignored. **10**

Relation R contains 10,000 tuples and has 10 tuples per page.

Relation S contains 2000 tuples and also has 10 tuples per page.

Attribute b of relation S is the primary key for S.

Both relations are stored as simple heap files.

Neither relation has any indexes built on it.

52 buffer pages are available.

i) What is the cost of joining R and S using a simple nested loops join ? What is the minimum number of buffer pages required for this cost to remain unchanged ?

Scanned by CamScanner

ii) What is the cost of joining R and S using a hash join ? What is the minimum number of buffer pages required for this cost to remain unchanged ?

iii) What would be the lowest possible I/O cost for joining R and S using any join algorithm and how much buffer space would be needed to achieve this cost ? Explain briefly.

iv) How many tuples does the join of R and S produce, at most and how many pages are required to store the result of the join back on disk ?

v) With 52 buffer pages, if clustered B+ indexes existed on R.a and S.b, would either provide a cheaper alternative for performing the join (using the index nested loops algorithm) than a block nested loops join ? Explain. Would your answer change if only five buffer pages were available ?

b) The following relation schema can be used to register information on the repayments on micro loans.

Repayment(borrower_id, name, address,loanamount, requestdate, repayment_date, request_amount)

A borrower is identified with an unique borrower_id, and has only one address. Borrowers can have multiple simultaneous loans, but they always have different request dates. The borrower can make multiple repayments on the same day, but not more than one repayment per loan per day.

i) State a key (candidate key) for repayment.

ii) State the functional dependencies for repayment.

iii) Normalize Repayment to BCNF. State for every step in the normalization, which functional dependency that causes it.

6. a) What is join dependency ? Explain 5NF.

b) Consider the attribute set R = {ABCDEG} and the FD set

F = {AB → C, AC → B, AD → E, B → D, BC → A, E → G}.

Check whether the decomposition of R into 4 sub-relations {AB, BC, ABDE, EG} is

a) dependency-preserving ?

b) lossless-join ?

Scanned by CamScanner

c) Consider the bank database given below. 10

Customer (cid,cname,ccity) ; Deposit (cid,accno) ;

Account (accno,branch_city, balance)

Consider the following query :

Find the names of the customers who live in the same city in which they have opened the account with a balance grater than 50,000/-. Draw the initial query tree and query graph for the above query. Show (Each step) how the query can be optimised. Provide the final optimised query.

MODULE – IV

7. a) Consider the following classes of schedules: conflict-serializable, view-serializable, recoverable, avoids-cascading-aborts, and strict. Explain your answer. For the following schedule, state which of the preceding classes it belongs to. If you cannot decide whether the schedule belongs in a certain class based on the listed actions, explain briefly.

The actions are listed in the order they are scheduled and prefixed with the transaction name.

T2:R(X), T2:W(X), T1:R(X), T2:R(Z), T1:W(Y), T1:commit, T4:R(X), T4:W(X), T4:W(Y), T2: commit, T3:R(Y), T4:Commit, T3:W(Z), T3:Commit. 10

b) What is the difference between wound-wait and wait-die techniques ? 4

c) Prove that the basic two-phase locking protocol guarantees conflict serializability of schedules. 6

8. a) What is Serializable schedule ? Why is serializability desirable ? 4

b) What is the phantom problem ? Can it occur in a database where the set of database objects is fixed and only the values of objects can be changed ? 5

c) Show that, if two schedules are conflict equivalent, then they are also view equivalent. 5

d) Show by example that there are schedules possible under the tree protocol that are not possible under the 2PL protocol and vice versa. 6

Scanned by CamScanner

COMP 5 – 5 (RC)

T.E. (Comp.) (Semester – V) (RC) Examination, November/December 2015

DATABASE MANAGEMENT SYSTEMS

Duration : 3 Hours Total Marks : 100

Instructions : 1) *Answer* ***five*** *questions in* ***all****, selecting at least* ***one*** *question from* ***each*** *Module.*

2) *Assume necessary data,* ***wherever*** *required.*

MODULE – I

1. a) The organisers of the EXAM 2015 international multi-conference need to keep track of a large collection of workshops associated with the event. Initial requirements analysis brings out the following information about what needs to be recorded. 10

- Each workshop has a name, and happens on a particular date – or dates, as some workshops last more than one day.
- There are several participants, each of which may sign up to one or more workshops.
- For each participant, it is important to record their name, email address, and the workshops which they wish to attend.
- There are a number of meeting rooms at the conference venue, each of a fixed capacity. Meetings rooms are identified by a floor and room number.
- Every workshop needs an allocated meeting room; where a workshop lasts for two days, it will use the same room on both days.

A) Draw an entity-relationship diagram suitable for representing this information, in particular the connections between participants, workshops, rooms and dates.

P.T.O.

Scanned by CamScanner

B) Further analysis reveals additional requirements. However, not all of these can be captured easily in an ER diagram.

- Each workshop must have an identified organiser among the conference participants.
- No participant may register for two workshops on the same day.
- Every participant may register for at least one workshop.

i) Identify two of these which can be captured in an ER diagram.

ii) For those two, show the additions required to your diagram.

b) What are the various levels of schema abstraction in DBMS ? Why is it required ?

c) Explain the responsibilities of the following subsystems of database (any 2) :

i) Query processor

ii) Storage manager

iii) Transaction manager.

d) List any two advantages of DBMS.

2. a) Explain the distinction between condition defined and user defined constraints. Which of these constraints can the system check automatically ? Explain your answer.

b) For each of the following concepts give a brief description of what it means, how it is represented in the ER diagram and provide an example.

i) Composite key

ii) Total participation

iii) Key constraint.

c) What do you mean by identifying relationship ?

d) Explain the following terms with the help of an example. Also discuss why are they required ?

I) Aggregation

II) Specialization.

Scanned by CamScanner

MODULE – II

3. a) Differentiate between a relation schema and a relation instance. Define the term relation degree. What are domain constraints ? 3

b) Let the following relational schema be given : 10

EMPLOYEE (SSN, Name, Age, salary, Dno)

WORK_ON (Project#, SSN)

PROJECT (Project#, pname, plocation)

For each of the following queries, give an expression in relational algebra :

i) Display the names of employees working on project at "Delhi".

ii) Retrieve the name of the working on maximum projects.

iii) For each project retrieve the number of employees working on that project.

iv) Find the names of the employees who work on all the projects that "Mr. ABC" works on.

v) For each department that has more than 5 employees, retrieve the department number and the number of its employees who are making more than 50,000/-.

c) What do you mean by safety of expressions in DRC ? 2

d) Check whether the following set of functional dependencies are equivalent or not. 5

F1 = {A->C, AC->D, E->AD, E->H} F2 = {A->CD, E->AH}

4. a) Consider the following relations : 10

Student (snum: integer, sname: string, major: string, level: string, age: integer)

Class (name: string, meets at: time, room: string, fid: integer)

Enrolled (snum: integer, cname: string)

Faculty (fid: integer, fname: string, deptid: integer)

Scanned by CamScanner

The meaning of these relations is straightforward; for example, Enrolled has one record per student-class pair such that the student is enrolled in the class. Write the following queries in SQL. No duplicates should be printed in any of the answers.

i) Find the names of all Juniors (Level = JR) who are enrolled in a class taught by Prof. 'AB'.

ii) Find the age of the oldest student who is either a History major or is enrolled in a course taught by Prof. BB.

iii) Find the names of all classes that either meet in room R128 or have five or more students enrolled.

iv) Find the names of all students who are enrolled in two classes that meet at the same time.

v) Find the names of students who are enrolled in the maximum number of classes.

b) Define a view. How is it different from a table ? Why are they used ? Provide an example of a view which cannot be updated and explain why it cannot be updated ? 5

c) Consider the following set of functional dependencies :

F = {A->BCE, AB->DE, BI->J}. Compute the canonical cover for F. 5

MODULE – III

5. a) How does hybrid hash join improve on the basic hash join algorithm ? How many buffer blocks are required to implement hash join and hybrid hash join ? Explain. 7

b) Define the term functional dependency. Why are some functional dependencies called trivial ? 6

Scanned by CamScanner

c) Consider a relation R with five attributes ABCDE. You are given the following dependencies : 7

$A \rightarrow B$, $BC \rightarrow E$ and $ED \rightarrow A$

i) List all keys for R.

ii) Is R in 3NF ?

iii) Is R in BCNF ?

6. a) Consider the following relational schema : 10

Student(<u>rollno</u>, name, sem, dno) dept(<u>dno</u>, dname)

Subject(<u>subid</u>, sname, sem) marks_obt(<u>rollno, subid</u>, marks)

Write a query to retrieve the names of Computer Engineering students who obtained marks greater than 50 in the subject 'DBMS'.

Draw the query graph and the initial query tree for the above query and then show (each step) how the query can be optimised. Provide the final optimised query.

b) Suppose you have a file with 20,000 blocks and you have five buffer blocks. Answer the following questions, assuming that general external sorting algorithm is used : 10

i) How many runs will you produce in the first pass ?

ii) How many passes will it take to sort the file completely ?

iii) What is the total I/O cost of sorting the file ?

iv) How many buffer pages do you need to sort the file completely in just two passes ?

Scanned by CamScanner

MODULE – IV

7. a) Explain the following terms using an example : 10

i) Partial schedule

ii) Recoverable schedule

iii) Cascadeless schedule

iv) Strict schedule

v) Consider the following schedule.

r1(x);w1(x);r2(x);w2(x);r2(y);w2(y);r1(z);w1(z);c1;c2;

Check whether the given schedule is cascadeless. If not, write down the minimum changes required to convert it into cascadeless schedule.

b) Consider the following actions taken by transaction T1 on database objects X and Y :

R(X), W(X), R(Y), W(Y) 4

i) Give an example of another transaction T2 that, if run concurrently to transaction T without some form of concurrency control, could interfere with T.

ii) Explain how the use of Strict 2PL would prevent interference between the two transactions.

c) Bring out the differences between the basic two-phase locking, strict two phase locking and rigorous two-phase locking using a appropriate example. 6

Scanned by CamScanner

8. a) What is the difference between conflict serializability and view serializability ? Which one of them is a stricter form ? Under what conditions conflict and view serializability are similar. Give example of the schedule which follows conflict serializability but not view serializability or vice versa. 10

b) Briefly explain why concurrency control is required. Give an example where adding lock/unlock instruction to the transactions (not in 2-phase order) does not give us the serializable schedule but by following 2 phase locking protocol we get a serializable schedule devoid of deadlocks. 6

c) What is the difference between wound wait and wait-die techniques ? 4

Scanned by CamScanner

REFERENCE SECTION LIBRARY PANDA G.E.C.

COMP 5 – 5 (RC)

T.E. (Comp.) (Semester – V) (RC) Examination, May/June 2015

DATA BASE MANAGEMENT SYSTEMS

Duration : 3 Hours Total Marks : 100

Instructions : *Attempt **any five** questions by selecting at least **one** question from **each** module.*

MODULE 1

1. a) Construct an E-R diagram for a car-insurance company whose customers own one or more cars each. Each car has associated with it zero to any number of recorded accidents. It is not necessary that a person involved in a accident is the owner of the car. Clearly specify cardinality and the participation constraints. Construct tables for the resulting ER diagram. Specify constraints. 6

   b) What is an identifying relationship ? Explain with the help of an example. Can the cardinality ratio of an identifying relationship be many to one ? 4

   c) Existence dependency need not always result in a weak entity set. Is this true ? Provide example to justify your answer. 3

   d) Explain the concept of physical data independence and its importance in database systems. 3

   e) List the various users of the database system. 4

2. a) What do you mean by recursive relationship ? Why is it required ? Provide example. 3

   b) Provide examples for the following : 4

      i) Total and disjoint specialization

      ii) Partial and overlapping specialization.

   c) Define the concept of aggregation. Give two examples of where this concept is useful. 5

   d) Explain the distinction between condition defined and user defined constraints. Which of these constraints can the system check automatically ? Explain your answer. 5

   e) What is a composite attribute ? How is it stored in a relation for a given entity ? Provide example. 3

P.T.O.

Scanned by CamScanner

MODULE 2

3. a) Consider the following DreamHome rental database

Branch (BranchNo, Street, City, PostCode)
Staff (StaffNo, Name, Position, Gender, DOB, Salary, BranchNo)
Property_for_rent (PropertyNo, Street, City, PostCode, Type, Rooms, Rent, OwnerNo, StaffNo, BranchNo)
Client (ClientNo, Name, TelNo, Preftype, MaxRent)
PrivateOwner (OwnerNo, Name, Address, Telno)
Viewing (ClientNo, PropertyNo, ViewDate, Comment)
Registration (ClientNo, BranchNo, StaffNo, DateJoined)

I) Answer the given queries in relational algebra : 10

i) List all cities where there is a branch office but no properties for rent.
ii) Identify the clients who have viewed all properties with 3 rooms.
iii) Identify the staff members that have registered maximum clients.
iv) List the property with the maximum rent and display its owner information.
v) List all cities where there is either a branch office or a property for rent.

II) Retrieve the names and comments of all clients who have viewed a property using Tuple Relational Calculus. 2

III) Identify all the staff members earning salary greater than Fifty Thousand using Domain relational calculus. 2

IV) Write the following query in QBE : Retrieve the name and telephone numbers of all the owners from Delhi. 2

b) What do you mean by entity integrity ? Why is it required ? 2

c) What is the difference between schema and instance of a relation ? 2

4. a) Consider the following database

Employee (empno,name,skill,salary)
Position (posting_no,skill)
Duty_allocation (posting_no,emp_no,day,shift)

Find SQL queries for the following : 10

i) Retrieve the name of the employee earning second highest salary.
ii) Find the names of all employees who are assigned to all positions that require a chef's skill.
iii) Get a list of all employees who are not assigned a duty.
iv) Get the employee numbers of all employees working on at least 2 days.
v) Retrieve the names of employees whose salary is greater than the salary of employee 'XYZ'.

Scanned by CamScanner

b) Giver a relation R = (A, B, C, D) and the corresponding set of functional dependencies, F = {A → BC, CD → E, B → D, E → A}. Find the canonical cover for F. What is the key of R ? 6

c) What do you mean by a functional dependency ? Consider the following relation student_guide(Name, Department, Guide). Identify all the functional dependencies in this relation. 4

Name	Department	Guide
Seema	CSE	Prof. X
Jina	Chemical	Prof. Y
Purva	Physics	Prof. Z
Nita	CSE	Prof. X
Anjali	Maths	Prof. T
Meena	Chemical	Prof. N

MODULE 3

5. a) Consider relation R = {A,B,C,D,E,F,G,H,I} and a set of functional dependencies F={AB→C, A→DE, B→F, F→GH, D→IJ}. Consider the following decomposition D={R1,R2,R3,R4,R5}, R1={A,B,C}, R2={A,D,E}, R3 ={B,F}, R4={F,G,H} and R5={D,I,J}.
Determine whether each decomposition has the following properties :
i) Dependency preservation and
ii) Lossless Join. 6

b) For a given relation R = {A,B,C,D,E,F,G,H} with the following set of functional dependencies
F = {A→BCD, AE→F, E→G, D→H}.
i) Find the key
ii) Relation R is in which normal form ?
iii) Apply Normalization on R until it cannot be decomposed further. Also give reasons for decomposition. 8

c) What are multivalued dependencies ? Why are they required ? How are they different from functional dependencies ? Provide examples to support your answer.

d) What is domain key normal form ?

Scanned by CamScanner

6. a) Suppose you need to sort a relation of 40 gigabytes, with 4 kilobyte blocks, using a memory size of 40 megabytes. Suppose the cost of a seek is 5 milliseconds, while the disk transfer rate is 40 megabytes per second. 8

i) Find the cost of sorting the relation, in seconds, with

A) number of buffer blocks allocated to each run bb=1 and

B) number of buffer blocks allocated to each run bb=100.

ii) In each case how many merge passes are required ?

b) Consider the following relations 8

employee(id,name,dept,salary)

works_on(eid,pno,hours)

project (pno,name,plocation)

Consider the query :

"Retrieve the name of the employees belonging to admin department and working on 'Project1' project for more than 5 hours per week."

i) Draw the query graph and the initial canonical query tree for the above query.

ii) Optimize the query using heuristic rules and show the resultant optimized query.

c) What is the different between pipelining and materialization ? 2

d) What do you mean by semantic query optimization ? 2

MODULE 4

7. a) Consider a transaction T that transfers Rs. 500/- from account A to account B as defined below : 6

Read(a)

a=a-500

Write(a)

Read(b)

b=b+500

Write(b)

Assume that prior to the execution of transaction T, the amount in account A is Rs. 2,000/- and in B is Rs. 3,000/-, Furthermore, assume that a power failure occurs. What will happen ? Will our database be in a consistent state ? Apply ACID test to check whether the four properties hold or not ?

b) During its execution, a transaction passes through several states. With the help of a state diagram, explain the various states. 6

Scanned by CamScanner

c) Consider the following schedules :

s1 : r1(x);r2(z);r1(z);r3(x);r3(y);w1(x);c1;w3(y);c3;r2(y);w2(z);w2(y);c2;

s2 : r1(x);r2(z);r1(z);r3(x);r3(y);w1(x);w3(y);r2(y);w2(z);w2(y);c1;c2;c3;

Determine whether each schedule is strict, cascadeless, recoverable, or nonrecoverable. Are the schedules conflict serializable ? For each serializable schedule, determine the equivalent serial schedule. **8**

8. a) Differentiate between wait-die and wound-wait schemes ? **4**

b) Produce a wait-for graph for the following transaction scenario and determine whether deadlock exists :

Transaction	Data item Locked by transaction	Data Item Transaction is waiting for
T1	X2	X1, X3
T2	X3, X10	X7, X8
T3	X8	X4, X5
T4	X7	X1
T5	X1, X5	X3
T6	X4, X9	X6
T7	X6	X5

Discuss two general techniques for handling deadlock

1) time out
2) deadlock detection and recovery.

c) What benefit does rigorous – 2PL provide ? How does it compare with other forms of 2PL ?

d) Discuss how serializability is used to enforce concurrency control in a database system. Why serializability is sometimes considered too restrictive as a measure of correctness for schedules ?

e) What is meant by catastrophic failure ?

Scanned by CamScanner

T.E. (Comp.) (Semester – V) (RC) Examination, Nov./Dec. 2014
DATABASE MANAGEMENT SYSTEMS

Duration : 3 Hours Total Marks : 100

Instruction : *Attempt **any five** questions by selecting at least **one** question from **each** module.*

MODULE – 1

1. a) A database is to be designed for a college. The college provides a number of module, each being characterised by its code, title, credit value, module leader, teaching staff and the department they come from. A module is co-ordinated by a module leader who shares teaching duties with one or more lectures. A lecturer may teach (and be a module leader for) more than one module. Students are free to choose any module they wish but the following rules must be observed : some modules require pre-requisites modules and some degree programmes have compulsory modules. The database is also to contain some information about students including their numbers, names, addresses, degrees they read for, and their past performance (i.e. modules taken and examination results). Draw an ER diagram to represent this information and convert the resulting ER into tables. 10

b) How is the physical data independence different from the logical data independence ? 3

c) Define the concept of aggregation. Give an example where this concept is useful and show the resulting tables. 5

d) What do you mean by role in a ER diagram ? When is it mandatory to specify a role in a ER diagram ? 2

2. a) Give an example for the following : 4

i) Disjoint and complete specialization.

ii) Overlapping and complete specialization.

P.T.O

Scanned by CamScanner

b) State any four advantages of the DBMS system over file processing system. 4

c) Give an example of an entity which is existent dependent on another entity but is not a weak entity. 2

d) What is a data model ? Explain in brief the relational data model using appropriate example. 4

e) Draw an E-R Diagram for the following scenarios. 6

i) An Employee may be assigned a company CAR, but some Employees do not qualify for a company CAR. A CAR will be assigned to one and only one employee.

ii) To be an athlete, you must play one of the following sports : Basketball, Tennis or Soccer. Athletes can play more than one sport.

iii) In a Manufacturing company, a part has at least one alternative part (or substitute part). Some parts do not serve as an alternative (substitute) for any parts. Also, a part may serve as an alternative for many different parts.

MODULE – 2

3. a) What is the difference between schema and the instance of a relation ? 2

b) Specify the following relational algebra operations in DRC. 6

i) $\sigma_{A=C}(R(A,B,C))$

ii) $\Pi_{<A,B>}(\sigma_{B=C}(R(A,B,C)))$

iii) $R(A,B,C) \cap S(C,D,E)$

c) Given 2 sets F1 and F2 of functional dependencies for a relation R, Check whether they are equivalent. 4

$F1 = \{A \rightarrow B, AB \rightarrow C, D \rightarrow AC, D \rightarrow E\}$.

$F2 = \{A \rightarrow BC, D \rightarrow AE\}$

Scanned by CamScanner

d) What is a functional dependency ? Consider the relation given below. Identify all the functional dependencies satisfied by this relation. Explain whether your answer will reflect the semantics of your data or not ? What are trivial functional dependencies ? 6

A	B	C
a1	B1	c1
a1	b1	c2
a2	b1	c1
a2	b1	c3

e) What do you mean by the degree of a relation ? 2

4. a) What are the restrictions that are put on the update operations using a view ? Explain using examples. 4

b) Consider the following database schema

Student (Name, studNo, Class, Major)

Course (Title, CourseNo, Credit Hours, Department)

Section (SectionIdentifier, CourseNo, Semester, Year, Instructor)

Grade_report (StudNo, SectionId, Grade)

Prerequisite (Course No, PrerequisiteNo).

I) Answer the following queries in SQL.

i) Retrieve the titles of courses taught by Professor Sinha in 2001 and 2005.

ii) For each section taught by professor Gupta, retrieve the course No, semester, year and the number of students who took the section.

iii) Retrieve the names and Major department of those students who have a grade A in all their courses.

Scanned by CamScanner

iv) Identify the students who have taken all the courses offered by computer science department in fall 2010.

v) Identify the course in fall 2012 with maximum A credits.

II) Answer the following queries in relational algebra. 6

i) Retrieve the names and Major department of those students who do not have a grade A in any of their courses.

ii) Identify the student with maximum A credits.

iii) Find the courses having more than two prerequisites.

MODULE – 3

5. a) Consider the universal relation R = {A,B,C,D,E,F,G,H,I} and the set of functional dependencies F = {AB->C, BD->EF, AD->GH, A->I, H->J}. 8

i) Find the key.

ii) Relation R is in which normal form ?

iii) Apply Normalization on R until it cannot be decomposed further. Also give reasons for decomposition.

b) What do you mean by double buffering ? 2

c) Define join dependencies and 5 NF. 4

d) Let relations r1 (A,B,C) and r2 (C,D,E) have the following properties : r1 has 30000 tuples, r2 has 50000 tuples, 30 tuples of r1 fit on one block and 50 tuples of r2 fit on one block. A primary index is available on attribute A of relation r1, with the number of levels is equal to 4. Estimate the minimum number of block transfers and seeks required for each of the below mentioned operations :

i) r1 join r2 using block nested loop join

ii) $\sigma_{A=100}(r1)$. 6

6. a) Extend the sort merge join algorithm to implement the outer join algorithm. 5

b) Consider the following relation 10

emp (id, name, salary, city)

Workson (eid, pno, hours)

project (pno, name, plocation)

Scanned by CamScanner

Consider the following query :

Retrieve the project no, project name and the number of employees from 'Mapusa' who work on 'Research' project for than 20 hours per week.

i) Draw the initial query tree and the query graph for the above query.

ii) Optimize the given query, and provide relational algebra expression for the final optimized query.

iii) Provide SQL expression for the optimized query.

c) With the help of a neat diagram, explain the various steps in query processing. 5

MODULE – 4

7. a) Consider the following two transactions : 8

T0	T1
Read (A)	Read (A)
A = A-50	T = A*0.5
Write (A)	A = A-T
Read (B)	Write (A)
B = B + 50	Read (B)
Write (B)	B = B + T
	Write (B)

i) Show, any parallel execution of T0 and T1 which produces a serializable schedule.

ii) Show, any parallel execution of T0 and T1 which produces a non-serializable schedule.

iii) For concurrency control in the above case, which lock can be used and how will that lock solve the problem ?

Scanned by CamScanner

b) What is a recoverable schedule ? Why is recoverability of schedules desirable ? Are there any circumstances under which it would be desirable to allow non-recoverable schedules ? Explain your answer. 6

c) Consider the following schedule :

T1	T2	T3
Read (A) A = f1 (A) Read (C) Write (A) A = F2 (C)		
	Read (A) Read (B)	
Write (C)		
		Read (C)
	B = f3 (B) Write (B)	
		C = f4 (C) Read (B) Write (C)
	A = f5 (A) Write (A)	
		B = f6 (B) Write (B)

Draw the precedence graph. Is the schedule conflict serializable ? If yes then state the corresponding serial schedule. Does the schedule also satisfy view serializability ? 6

Scanned by CamScanner

[Total No. of Questions : 8]

T.E. (COMP) (Semester - V) (RC) Examination, Nov. / Dec. - 2011

DATA BASE MANAGEMENT SYSTEM

8. a) Draw a state diagram and discuss the typical states that a transaction goes through during execution. 4

b) Does 2PL protocol ensures conflict serializability ? Justify with the help of an example. 6

c) Describe strict 2PL. What benefit does strict 2PL provide ? What are its disadvantages ? 6

d) What are deadlocks ? How cautious waiting does avoids deadlock ? 4

Scanned by CamScanner

[Total No. of Questions : 8]

T.E. (COMP) (Semester - V) (RC) Examination, Nov. / Dec. - 2011

DATA BASE MANAGEMENT SYSTEM

Duration : 3 Hours **Total Marks : 100**

Instructions : *1) Answer any five questions, at least one from each Module.*
2) Assume suitable data, if necessary.

REFERENCE SECTION LIBRARY PONDA G.E.C.

MODULE - I

Q1) a) With the help of a neat diagram explain the main phases of database design. [7]

b) An existence dependent entity need not be a weak entity. Justify with an example. [4]

c) What is meant by a recursive relationship type? Give an example. [4]

d) What is the difference between a specialization heirarchy and a specialization lattice? [5]

Q2) a) List the various types of attributes used in ER models. [2]

b) Construct an ER diagram for a Hospital with a set of patients and a set of Docters. Associate with each patient a log of the various tests and examinations conducted and treatment given. [5]

c) Discuss the two main types of constraints on specialization. [5]

d) Explain the difference between external, internal and conceptual schemas. How are these different schema layers related to the concepts of logical and physical data independence. [8]

MODULE - II

op ***Q3)*** a) Discuss the entity integrity and referential integrity constraints why is each considered important. [6]

b) Differentiate between : [4

i) Outer join and inner join.

ii) Relational Algebra and Relational Calculus.

c) Consider the following schema :

Hotel (<u>Hotel No</u>, Name, City)

Room (<u>Room No, Hotel No</u>, Type, Price)

Booking (<u>Hotel No, Guest No, Date From</u>, Date To, Room No)

Guest (<u>Guest No</u>, Name, Address)

P.

Scanned by CamScanner

B - 869 (Pages : 03) COMP 5-5 (RC)

[Total No. of Questions : 8]

T.E. (COMP) (Semester - V) (RC) Examination, Nov. / Dec. - 2011

DATA BASE MANAGEMENT SYSTEM

Duration : 3 Hours Total Marks : 100

Instructions : *1) Answer any five questions, at least one from each Module.*

2) Assume suitable data, if necessary.

REFERENCE SECTION LIBRARY PONDA G.E.C.

MODULE - I

Q1) a) With the help of a neat diagram explain the main phases of database design. [7]

b) An existence dependent entity need not be a weak entity. Justify with an example. [4]

c) What is meant by a recursive relationship type? Give an example. [4]

d) What is the difference between a specialization heirarchy and a specialization lattice? [5]

Q2) a) List the various types of attributes used in ER models. [2]

b) Construct an ER diagram for a Hospital with a set of patients and a set of Docters. Associate with each patient a log of the various tests and examinations conducted and treatment given. [5]

c) Discuss the two main types of constraints on specialization. [5]

d) Explain the difference between external, internal and conceptual schemas. How are these different schema layers related to the concepts of logical and physical data independence. [8]

MODULE - II

Q3) a) Discuss the entity integrity and referential integrity constraints why is each considered important. [6]

b) Differentiate between : [4]

i) Outer join and inner join.

ii) Relational Algebra and Relational Calculus.

c) Consider the following schema :

Hotel (Hotel No, Name, City)

Room (Room No, Hotel No, Type, Price)

Booking (Hotel No, Guest No, Date From, Date To, Room No)

Guest (Guest No, Name, Address)

Scanned by CamScanner

Write SQL statements for the following :

i) How many hotels are there? [1]

ii) List the price and type of all rooms at the Grosvenor Hotel. [2]

iii) List the number of rooms in each hotel in London. [2]

iv) How many different guests have made bookings for August. [2]

v) List the names of guests currently staying at Grosvenor Hotel. [3]

Q4) a) Define the division operation in relational algebra. What kind of queries is it used for? Illustrate with an example. [5]

b) Discuss various types of join operations. Why is theta join required? [7]

c) What is a correlated nested query? Explain with an example. [4]

d) What is meant by closure of a set of functional dependencies. [4]

MODULE - III

Q5) a) Why should need values in a relation be avoided as far as possible? [6]

b) Analyse the sample data in the following table carefully :

PET ID	PET NAME	PET TYPE	PET AGE	OWNER	VISIT DATE	PROCEDURE
246	ROVER	DOG	12	SAM COOK	JAN 13/2002	01 - RABIES VACCINATION
					MAR 27/2002	10 - EXAMINE and TREAT WOUND
					APR 02/2002	05 - HEART WORM TEST
298	SPOT	DOG	2	TERRY KIM	JAN 21/2002	08 - TETANUS VACCINATION
					MAR 10/2002	05 - HEART WORM TEST
341	MORRIS	CAT	4	SAM COOK	JAN 23/2001	01 - RABIES VACCINATION
					JAN 13/2002	01 - RABIES VACCINATION
519	TWEEDY	BIRD	2	TERRY KIM	APR 30/2002	20 - ANNUAL CHECK UP
					APR 30/2002	12 - EYE WASH

Identify the functional dependencies and show all the intermediate steps in normalizing the above relation upto third normal form. [8]

c) Explain multi valued dependencies and fourth normal form. [6]

Scanned by CamScanner

Q6) a) Discuss query tree and query graph with respect to query optimization techniques. [6]

b) Suppose we have an available buffer space of 64 blocks and a file of 4096 blocks is to be sorted. How many posses will be needed in the merge phase of the external sort-merge algorithm. [4]

c) What is meant by the term heuristic optimisation? Discuss the main heuristics that are applied during query optimization. [10]

REFERENCE SECTION LIBRARY PONDA G.E.C.

MODULE - IV

Q7) a) What is a schedule (history)? Define the concepts of recoverable, cascadeless and strict schedules. [8]

b) Discuss the wait-die and wound-wait protocols for deadlock prevention. [6]

c) Explain multiversion two phase locking using certified locks. [6]

Q8) a) With the help of an example explain ACID properties of a transaction. [8]

b) What is concurrency control? What are its objectives? [4]

c) With respect to two phase locking discuss : [2 x 4 = 8]

i) Basic two phase locking.

ii) Conservative two phase locking.

iii) Strict two phase locking.

iv) Rigorous two phase locking.

✦✦✦✦

Scanned by CamScanner

REFERENCE SECTION LIBRARY PONDA G.E.C.

COMP 5-5 (RC)

T.E. (Comp.) (Semester – V) (RC) Examination, Nov./Dec. 2010
DATABASE MANAGEMENT SYSTEMS

Duration : 3 Hours Total Marks : 100

Instructions :1) Answer any five questions, atleast one from each Module.
2) Assume suitable data, if necessary.

MODULE – I

1. a) Discuss in detail the advantages of DBMS over conventional file systems. 8

b) Draw an ER diagram which stores information about interviews of job applicants in various companies. Store the name and address of the company, job applicants number, name, address and phone number. Some of the interviews conducted result in a job offered. For every interview conducted store the name of the person conducting the interview and the phone number. Store the type of job offered for job applicants who have successfully answered the interview. 6

c) How is a weak entity different from a strong entity ? How is it represented in an ER diagram. 4

d) List the various types of attributes used in ER models. 2

2. a) A rail transport company wishes to hold details of trains services, stations and operators. 10

Use the following business rules and construct an ER diagram

Trains : A train has a unique id code, a collection of days it runs (Monday, Sunday, etc...) on and the number of carriages it has. Each train must have a number of drivers who can drive it and a driver may drive a number of trains. The train driver has a unique pay roll number and a name. A train is designated as either an intercity or district train. Each train must have one operator who has a name and phone number for complaints. The operator can operate many trains (or none at all!)

P.T.O.

Scanned by CamScanner

Stations : A station has a name and the number of platforms. All stations are either main or district.

Timetable : A train has a number of stops at a station during its journey. It never visits the same station twice during the same journey. At each stop the arrival and departure time is recorded. A station has many trains passing through it, often arriving at the same time and a train visits many stations.

b) Differentiate between specialization and generalization with an example. 6

c) Discuss the functions of a database administrator. 4

MODULE – II

3. a) What is a foreign key ? Explain the concept of referential integrity with an example. 6

b) What do you understand by canonical cover of F ? For R = {A, B, C, D, E, F, G, H} compute the canonical cover for set F

$F = \{A \rightarrow B, ABCD \rightarrow E, ACDF \rightarrow EG, EF \rightarrow GH\}$. 8

c) Consider the following relational schema :

sailors (<u>sid</u>, sname, rating, age)

boats (<u>bid</u>, bname, colour)

reserves (<u>sid</u>, <u>bid</u>, day)

Answer the following w.r.t. the schema above

i) Write the DDL for all the tables. 3

ii) Give SQL formations for the following : 3

- insert a record into the reserves table
- change the rating of all the sailors

Whose age is greater than 30 and set the value to 10.

Scanned by CamScanner

REFERENCE SECTION LIBRARY PONDA G.E.C.

4. a) Consider the following relational schema :

employee (empno, name, office, age)

book (isbn, tittle, authors, publisher)

loan (empno, isbn, date).

Write relational algebra statements for the following :

i) Find the names of employees who have borrowed one or more books published by "TMH". 2

ii) Find the employee number of employees who have never borrowed a book written by "Galvin". 3

iii) Find the average age of employees who have borrowed books written by "PS Gill". 3

b) List and explain the various pitfalls in relational database design. 6

c) Explain the concept of views in SQL. Illustrate with an example. Discuss the problem that may arise when one attempts to update a view. 6

MODULE – III

5. a) Consider the following relation for published books :

book (book_title, AuthorName, book_type, ListPrice, author_affiliation, publisher)

book_title and AuthorName are the keys

Suppose the following dependencies exist :

book_title → publisher, book-type

book_type → ListPrice

AuthorName → author_affiliation

Decompose the relation into 2NF and 3NF.

b) Explain with a diagram the processing of a high level query.

c) Why do we need to normalize tables ? Explain.

Scanned by CamScanner

6. a) Define multivalued dependencies. Give a suitable example. 4

b) Explain the concept of heuristic optimization of query trees. Give an example. 8

c) What is external sorting ? Explain with an example. 8

MODULE – IV

7. a) In context of concurrent access in a multiuser database system describe the following terms : 6

i) transaction

ii) commit

iii) rollback

b) Differentiate between exclusive lock and shared lock. 6

c) Discuss the basic time stamp algorithm. Explain with procedures how are write and read operations implemented. 8

8. a) What is concurrency control ? What are its objectives ? 4

b) Explain multiversion 2PL using certified locks. 8

c) With the help of an example explain ACID properties of a transaction. 8

Scanned by CamScanner

T.E. (Comp.) (Semester – V) Examination, Nov./Dec. 2009
DATABASE MANAGEMENT SYSTEM (RC)

Duration : 3 Hours

17

COMP 5 – 5 (RC)

T.E. (COMP) (Sem. – V) (RC) Examination, May 2010
DATABASE MANAGEMENT SYSTEM

Duration: 3 Hours Total Marks: 100

MODULE – 1

1. a) With the help of neat diagram explain the main phases of database design. 5

b) Define the following terms with appropriate examples : 10

1) Entity 2) Attributes
3) Relationship 4) Instance
5) Metadata.

c) Explain the concepts of Generalization and Aggregation used in ER with the help of examples. 5

2. a) Consider a database used to record the marks that students get in different exams of different course offerings. Construct an E-R diagram that models exams as entities and uses a ternary relationship for the above database. 8

b) Explain the difference between a weak and a strong entity set. 4

c) Construct an ER diagram in which the same entity set appears several times. Why is allowing this redundancy a bad practice that one should avoid whenever possible ? 8

MODULE – 2

3. a) Consider the following tables. The primary keys are underlined.

Singers : SingerId, Name, Language
Songs : SongId, Title, Mood
Relationship : SongId, SingerId, Year

Answer the following using Relational Algebra : 10

1) Find names of all singers who have sung songs with blue mood.
2) Find names of all singers who have sung blue and cheerful songs.
3) Find names of all singers who have sung blue or cheerful songs.
4) Find names of all singers who have sung only blue songs.
5) Find names of all singers who have sung songs of every mood.

P.T.O.

Scanned by CamScanner

b) Define foreign key. What is this concept used for ? How does it play role in join operations ? 5

c) How is view defined in SQL ? Give examples. 5

4. a) Specify the following relational algebra operations in both tuple and domain relational calculus : 12

a) $\sigma_A = C(R(A, B, C))$ b) $\pi_{<B, A>}(R(A, B, C))$

c) $R(A, B, C) * S(C, D, E)$ d) $R(A, B) \div S(A)$

e) $R(A, B, C) - S(A, B, C)$ f) $R(A, B, C) \times S(D, E, F)$

b) Explain the pitfalls of bad relational database design. 5

c) What is COBE and how does it differ from SQL ? 3

MODULE – 3

5. a) What is a query plan ? 5

b) What are the elements in estimating the cost of a query plan ? 5

c) Explain the use of heuristics in query optimisation. 10

6. a) Explain binary sorting for executing query operations. 5

b) What do you mean by BCNF ? How it is different from 4th normal form ? 5

c) Explain 3NF with an example. 5

d) Explain the concept of join dependencies and fifth normal form. 5

MODULE – 4

7. a) Explain why concurrency control is needed ? And state different techniques used. 10

b) List and explain the desirable properties of transactions. 5

c) Illustrate "serializability with examples". 5

8. a) What is 2PL protocol ? How does it guarantee serializability ? 5

b) Discuss the wait-die and wound-wait protocols for deadlock prevention. 5

c) Discuss the problems of deadlock and starvation and the different approaches used in dealing with this problem. 10

Scanned by CamScanner

11

COMP – 5-5 (RC)

T.E. (Comp.) (Semester – V) Examination, Nov./Dec. 2009
DATABASE MANAGEMENT SYSTEM (RC)

Duration : 3 Hours Total Marks : 100

Instruction : Assume any data if required.

Module – 1

1. a) State and explain the various functions of a Database Administrator. 5

   b) Explain the concepts of total participation and partial participation w.r.t. relationship in dbms. 5

   c) State and explain the various characteristics of the database approach that distinguishes it from traditional approach of programming with files. 10

2. a) A database is being constructed to keep track of the teams and games of a sports league. A team has a number of players, not all of whom participate in each game. It is desired to keep track of the players participating in each game for each team and the result of the game. Design an E-R diagram considering your favourite sport. 8

   b) Differentiate between aggregation and generalization with an example. 4

   c) Construct an E-R diagram for a car-insurance company whose customers own one or more cars each. Each car has associated with it zero to any number of recorded accidents. 8

Module – 2

3. a) Discuss the various reasons that lead to the occurance of null valves in relations. 5

   b) Explain the following terms w.r.t. relation schema. 5

   i) Super key ii) Candidate key iii) Primary key

   c) Consider the following relations shown below.

   Works (PName, CName, Salary)

   Lives (PName, Street, City)

   Located In (CName, City)

   Managers (PName, MgrName)

   PName refers to the person name, CName refers to company name and MgrName refers to manager name and it is a foreign key that references PName of works. 10

P.T.O.

Scanned by CamScanner

Specify the following queries in relational algebra.

a) Find the names of the person who work for the company 'IBM'

b) List the names of the person working for 'HP' along with the cities they live in.

c) Find the names of the person who live and work in the same city.

d) Find the names of person who do not work for the company 'IBM'

e) Find the names of the companies that are located in every city where the company 'HP' is located in.

4. a) What is view in SOL and how is it defined ? Discuss the problems that may arise when one attempt to update a view. 7

b) How does SOL allow implementation of the entity and referential integrity constraint ? Illustrate with suitable example. 7

c) What do you mean by safe expression in relational calculus ? Explain. 6

Module – 3

5. Write query processing briefly discuss the following algorithms.**(any 4)** 5×4=20

1) Join 2) Sorting 3) Merge-Join

4) Pipelining 5) Nested loop Join

6. a) Illustrate the following normal forms with example :

1) First Normal Form

2) Second Normal Form

b) R = {A, B, C, D, E, F, G, H, I, J} 5

FD = { AB → C

A → DE

B → F

F → GH

D → IJ }

1) Decompose R into 2 NF and 3 NF

2) Why do you need to normalise the tables ? 5

3) What do you mean by BCNF ? Explain with an example. 5

5

Scanned by CamScanner

Module – 4

7. a) What is a schedule (history) ? Define the concepts of recoverable, cascadeless and strict schedules and compare them in terms on their recoverability. 10

b) Differentiate between exclusive lock and shared lock. 5

c) What is concurrency control ? What is its objectives ? 5

8. a) Explain multiversion two phase locking using time stamp ordering. 8

b) Write 2 phase locking (2PL) discuss 12

1) Basis 2PL

2) Conservative 2PL

3) Strict 2PL

4) Rigorous 2PL.

Scanned by CamScanner

T.E. (Comp.) (Semester – V) (RC) (2007-08) Examination, May/June 2018
DATABASE MANAGEMENT SYSTEMS

Duration : 3 Hours Total Marks : 100

Instructions : *1) Attempt **any five** questions by selecting at least **one** question from **each** Module.*
*2) Make necessary assumptions, **wherever** required and **clearly** specify the assumptions made.*

MODULE – 1

1. a) A general hospital consists of a number of specialized wards. Each ward hosts a number of patients. On admission, the personal details of every patient are recorded. A separate register is to be held to store the information of the tests undertaken and the results of a prescribed treatment. A number of tests may be conducted for each patient. Each patient is assigned to one leading consultant but may be examined by another doctor, if required. Doctors are specialists in some branch of medicine and may be leading consultants for a number of patients, not necessarily from the same ward. Draw an entity-relationship diagram suitable for representing this information. Convert this ER diagram into tables and represent the constraints on each table. **10**

b) What is the importance of the storage manager in the DBMS ? What are the various components of the storage manager ? **5**

c) What is a query language ? What is the difference between procedural and non procedural query language ? Is SQL a procedural or non-procedural language ? **5**

2. a) What is the role of DBA in the DBMS system. **4**

b) What is a participation role in an ER diagram ? When is it necessary to use role names in the description of entity types ? Provide example. **5**

c) What is aggregation ? Why is it required ? Justify using an example. **4**

d) What do you mean by a weak entity ? Provide example. **3**

e) What is a data model ? Explain any one data model. **4**

Scanned by CamScanner

MODULE – 2

3. a) Consider the following relations : 10

Student (snum : integer, sname : string, major : string, level : string, age : integer)

Class (name : string, meets at : time, room : string, fid : integer)

Enrolled (snum : integer, name : string)

Faculty (fid : integer, fname : string, deptid : integer)

The meaning of these relations is straightforward; for example, Enrolled has one record per student-class pair such that the student is enrolled in the class.

Answer the following queries in SQL :

1) Write the SQL statements required to create the enrolled and class relations, including appropriate versions of all primary and foreign key integrity constraints.
2) For each student find the number of enrolled classes.
3) Find the name of the faculty taking maximum number of class.
4) Find the name of the student who has not enrolled for any class.
5) Find the room in which more than two classes are held.

b) Given a relation, R = {A, B, C, D, E, F, G, H, I} and the corresponding set of FD's, F = {ABCD → E, AB → G, B → F, C → J, CJ → I, G → H}. Find the canonical cover for F. 6

c) What do you mean by unsafe expressions in Tuple Relational Calculus (TRC) ? Specify the following relational algebra operation in TRC. 4

$\Pi_{\alpha,\beta}(\sigma_{\gamma\rho z}(R(X, Y, Z)))$

4. a) Consider the following relations : 10

Sailors (sid, sname, rating, age)

boats (bid, bname, colour)

reservations (sid, bid, date)

Answer the following queries in relational algebra

i) Find sailors who've reserved a red and a green boat.
ii) Find the names of sailors who've reserved all red boats.
iii) Find the names of sailors who has reserved maximum boats.
iv) For each sailor find the number of boats reserved.
v) Retrieve the date on which maximum reservations were made ?

Scanned by CamScanner

b) Suppose that we have the following three tuples in a legal instance of a relation schema S with three attributes ABC (listed in order) : (1, 2, 3), (4, 2, 3) and (5, 3, 3).

i) Which of the following dependencies can you infer does not hold over schema S ?

a) $A \rightarrow B$, (b) $BC \rightarrow A$, (c) $B \rightarrow C$.

Justify your answer.

ii) Can you identify any other dependencies that hold over S ? 4

c) What are views ? Can views be updated ? Justify using examples. 4

d) What is referential integrity ? 2

MODULE – 3

5. a) Consider the relation for the published books : 7

BOOK (book_title, authorname, book_type, listprice, author_affil, publisher)

where author_affil refers to the affiliation of author.

Suppose the following dependencies exist :

book_title → publisher, book_type

book_type → listprice

authorname → author_affil

What normal form is the relation in ? Explain your answer.

Apply normalization until you cannot decompose the relations further. State the reasons behind each decomposition.

b) Consider the employee database given below : 10

employee (<u>ename</u>, street, city, age, gender)

works (<u>ename</u>, cname, salary)

company(<u>cname</u>, city)

Consider the query to retrieve the names of the female employees who works on a company located in 'Mumbai' city and earns salary over 50,000. Draw the initial query tree for the above query and then show how the query can be translated to the optimized query. Show all intermediate steps. When would we prefer a un-optimized query over a optimized one ?

c) What is multivated dependency ? Define fourth normal form. 3

6. a) What do you mean by left deep trees ? Why are they preferred by query optimizers ? 4

b) Consider the join $R \bowtie_{R.a = S.b} S$, given the following information about the relations to be joined. The cost metric is the number of block transfers and number of seeks and the cost of writing out the final result should be uniformly ignored. 16

Scanned by CamScanner

Relation R contains 25000 tuples and has 25 tuples per block
Relation S contains 50000 tuples and has 10 tuples per block
Attribute b of relation S is the primary key for S
There is a B+ tree primary index available on attribute R.a with h = 20;
52 buffer blocks are available.
Answer the following questions and make necessary assumptions and specify the assumptions made.

i) What is the cost of joining R and S using block nested loops join ? Which one would you choose to be the inner relation and why ? What is the minimum number of buffer pages required for this cost to remain unchanged ?

ii) What is the cost of sorting S using a external sort-merge algorithm ? What is the minimum number of buffer pages required for this cost to remain unchanged ?

iii) What will be the minimum cost (Number of transfers and seeks) for following selections : (specify the selection algorithm used and justify why the selected algorithm is the best)

A) $\sigma_{R.a = 7}(R)$

B) $\sigma_{R.a <= 125}(R)$

C) $\sigma_{R.a = 3 \wedge R.C > 20}(R)$

MODULE – 4

7. a) How do we determine whether a given schedule is conflict serializable or not using a precedence graph. Determine whether the following schedule is conflict serializable, recoverable and cascadeless schedule. 8

Schedule :

T1		R(A)	W(A)			C		
T2	R(A)	R(B)			W(A)		C	
T3				R(B)				C

C represents commit of transaction ; R(x)-represents read of data item x; W(x) represents write on data item x.

b) Explain tree based protocol in concurrency control. Also give an example of serializable schedule under the tree protocol. 6

Scanned by CamScanner

c) Explain the distinction between serial schedule and serializable schedule. Consider. 6

T_0	T_1
Read (A)	Read (A)
A=A–50	T = A * 0.5
Write (A)	A = A – T
Read (B)	Write (A)
B = B + 50	Read (B)
Write (B)	B = B + T
	Write (B)

i) Show, any parallel execution of T_0 and T_1 which produces a serializable schedule.

ii) Show, any parallel execution of T_0 and T_1 which produces a non-serializable schedule.

8. a) Show by example that there are schedules possible under the tree protocol that are not possible under the 2PL protocol and vice versa. 5

b) Briefly explain why the serialization order matches the commit order using rigorous two-phase locking. 6

c) What is a cascadeless schedule ? Why is cascadelessness of schedules desirable ? Are there any circumstances under which it would be desirable to allow non cascadeless schedules ? Explain your answer. 5

d) If deadlock is avoided by deadlock avoidance, is starvation still possible ? Explain your answer. 4

Scanned by CamScanner