NATIONAL INSTITUTE OF TECHNOLOGY, KURUKSHETRA

BTech(IT) 3rd Sem. End Term Examination Dec 2022

Theory Examination

Course no.:ITPC 25..... Course Name: Database Systems

Date of Exam: 21 Dec 2022..... Time allowed: 03 Hours.....

No. of Questions to be attempted: 05.. Maximum Marks: 50.....

No. of Page used: 03..... Roll No.: ....121.1.3.0.76.....

Note: Answer all the questions, an internal choice given for Qno. 4(c). You may answer the questions in any order. However, all parts of the same question must be answered together. Clearly state any reasonable assumption you make.

Q.1 Answer the following (*briefly*): **(1*10)**

(a) Why database is important?

(b) How data independence plays important role in database architecture design?

(c) What is the difference between DBMS and RDBMS?

(d) Why are NULL values needed in the relational model?

(e) How following concepts are related to each other: (i) *Primary key & Foreign key* (ii) *Candidate key & Super key*?

(f) Between the properties of *dependency preservation* and *loss-lessness join*, which one must definitely be satisfied? Why?

(g) What undesirable dependencies are avoided when a DB relation is in 2NF & 3NF?

(h) Which Normal Form (NF) is considered optimal (best possible) for a DB relation?

(i) How do you describe a DB Transaction in '*Partially Committed*' state?

(j) What is *Lost-Update* problem? Dose a Shared/Exclusive 2PL protocol will be able to solve it?

Q.2 (a) Consider that in an ER Model, there is an entity set ***People*** with attribute ***id, name*** and ***designation***, where ***id*** is the key attribute. There are two specializations of people: ***Employee*** *and* ***Student***. This generalization/specialization between ***People, Employee*** *and* ***Student*** is overlapping and partial. Employee has a attribute ***date_of_joining***. There are two specializations of students: ***UG*** *and* ***PG***. This generalization/specialization between **Student, UG** and **PG** is disjoint and total. **UG** has the multi-valued attribute **hobby** and **PG** has the attribute **previous_degree**. Develop a suitable EER schema with clearly identified primary key and appropriate participation information's. **(3,3)**

(b) Design an ER schema for an IT training group database that will meet the information needs for its training program. Clearly indicate the entities, relationship and the key constraints. The description of the environment is as follows: '***The company has 12 instructors and can handle up to 100 trainees for each training session. The company offers 5 Advanced technology courses, each of which taught by a team of 2 or more instructors. Each instructor is assigned to a maximum of two teaching teams or may be assigned to do research. Each trainee undertakes one Advanced technology course per training session***'.

NATIONAL INSTITUTE OF TECHNOLOGY, KURUKSHETRA

BTech(IT) 3rd Sem. End Term Examination Dec 2022

Theory Examination

Q.3 (a) SQL allows a *foreign key* dependency to refer to the same relation, as in the following example: **(3,6)**

- **create table** manager (employee_name **varchar** (20) **primary key**, manager_name **varachar** (20) **not null**,
  foreign key (manager_name) **references** manager (employee_name) **on delete cascade set null on update cascade**);

Here, *employee_name* is a key to the table, meaning that each employee has at most one manager. The foreign key clause requires that every manager also be an employee. Explain exactly what will happen when a tuple in the relation *manager* is deleted or updated.

(b) Consider the ***Company*** relational DB, where the primary key(s) are underlined.

employee (employee-name, salary, city, country, gender)
company (company-name, city)
works (employee-name, company-name)
manages (employee-name, manager-name)

Give an equivalent **SQL & Relational Algebra expression** for each of the following queries:

(i) Find the names of all female employees in this database who work for company named as '*TIBCO*' but do not live in the same ~~country~~ city.

(ii) Retrieve the employees list, who lives in the same city as company located at, but their manager do not lives in.

(iii) Assume the companies may be located in several cities. Find all companies located in every city in which '*Small Bank Corporation*' is located.

(iv) For each company that has more than twenty five thousand employees and located in at least 05 cities, retrieve the company name and number of its employees, who are making more than $99,000.

Q.4 (a) Consider ~~two~~ one FD sets for a relation R(A,B,C,G,H,I): F={A->B, A->C, CG->H, CG->I, B->H}. Evaluate and discuss whether a FD: AG->I is logically implied or not? **(3,3,6)**

(b) Consider a relation R (A, B, C, D, E) and FD set F= {A->BC, CD->E, B->D, E->A, AE->BD}. This relation is decomposed into two sub relations r_1(ABC) and r_2(ADE). Check if the decomposition is dependency preserving or not. Also check whether it is Lossless or lossy decomposition. Discuss Briefly.

(c) Consider the relation **TRIP (Trip_id, Start_date, Cities_visited, Cards_used)**. *This relation refers to business trips made by company salespeople.* Suppose that a *TRIP* has a single *Start_date*, but involves many *Cities* and *Salespeople* may use multiple credit cards on the *Trip.*'

Answer the following:

(i) Discuss what FDs exist in this relation.

(ii) Identify the candidate keys for derived FD set.

(iii) Show how you will go about normalizing it to highest NF.

OR

NATIONAL INSTITUTE OF TECHNOLOGY, KURUKSHETRA

BTech(IT) 3rd Sem. End Term Examination Dec 2022

Theory Examination

Consider the relation **CAR_SALE(Car#, Date_sold, Salesperson#, Commission%, Discount_amt)**. Assume that a car may be sold by multiple *salespeople*, and hence {Car#, Salesperson#} is the primary key. Additional dependencies are: *Date_sold→Discount_amt* (ii) Salesperson#→Commission%.

(i) Discuss what functional dependencies (FDs) exist in this relation.
(ii) Identify the candidate keys for FD set.
(iii) Show how you will go about normalizing it to highest NF.

Q.5 (a) Consider the three DB *transactions* T_1, T_2, and T_3 and a *schedule* S as given below. The actions are listed in the order they are scheduled and prefixed with the transaction name. Evaluate and discuss whether *Schedule (s)* is serializable or not? **(3,4,6)**

S: $r_1(X)$, $r_2(Z)$, $r_1(Z)$, $r_3(X)$, $r_3(Y)$, $w_1(X)$, $w_3(Y)$, $r_2(Y)$, $w_2(Z)$,$w_2(Y)$;

(b) Consider the following two DB transactions:

T_1: read(A);
read(B);
if A = 0 then B := B + 1;
write(B).

T_2: read(B);
read(A);
if B = 0 then A := A + 1;
write(A).

Add *lock* and *unlock* instructions to *transactions* T_1 and T_2, so that they observe the *two-phase locking (2PL) protocol*. Can the execution of these transactions result in a deadlock? Generalize your view.

(c) Consider the below mentioned ***Schedule*** and ***Data Items*** (with $X=20$, $Y=30$ as initial values):

T_1	T_2
read_lock(Y);	read_lock(X);
read_item (Y);	read_item(X);
unlock (Y);	unlock (X);
write_lock(X);	write_lock (Y);
read_item(X);	read_item(Y);
X=X+Y;	Y=X+Y;
write_item (X);	write_item (Y);
unlock (X);	unlock (Y);

Answer the following(s):

(i) Derive the possible values of data items X & Y, for both orders **T_1-> T_2** & **T_2-> T_1**.
(ii) Verify, whether above mention schedule follows 2PL protocol? If yes explain the order and data item values for the sequence, otherwise justify your answer.