CS F212 – DATABASE SYSTEMS
II SEMESTER 2017-2018
Comprehensive examination

Date: 07th May, 2018   **Time: 3 Hours**   **Weightage: 40%**

Part A – 5 Short Answer Questions (closed book) [10% - 5*4 = 20 marks]
[Max. Time - 45 mins.]

1. Suppose we store a relation R(a,b,c,d) in a partitioned hash table with 4096 buckets (i.e., 12-bit bucket address). Each query on R specify exactly one of the attributes, and the attributes have probability of 0.4, 0.3, 0.2. & 0.1 respectively, to be specified. The hash function produces 5 bits based only for a, 3 bits based only on b, 2 bits based only on c, and 2 bits based only on d.
    (a) Find the bucket number of the bucket containing record (127,31,15,7).
    (b) What is the average number of buckets that need to be searched to answer a query?
    (show some calculations)

2. Let F be a set of FDs on a relation R that has been decomposed into n relations, R_i ($i=1,2,…n$). Write a sufficient condition for dependency preservation. Also give a necessary condition using the concept of restrictions of F to R_i.
    Sufficient condition:

    Necessary Condition:

3. In the absence of any information on 6 relations, r_i ($i=1,2,3,4,5,6$), give a heuristic that will avoid enumerating all possible join orders and justify your answer in 1 line. Simply give the RA expression tree. How many join orders will be there?

4. Suppose a materialized view, **vold** = **r (full outer join) s.** Describe how we can incrementally maintain this view on insertions & deletions on **r.** It is given that **ir** is the set of tuples inserted in **r** and **dr** is the set of tuples deleted from **r. vnew** is the view after the inserts or deletes on **r.**

   For inserts **ir in r:**
   vnew =

   For deletes **dr from r:**
   vnew =

5. Graphically depict the hash join algorithm between two relations R and S. The relation R occupies 10 blocks, whereas S occupies 5 blocks. The hash function used is modulo 5. Annotate the figure properly so that the idea behind hash join algorithm is highlighted.

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI
Department of Computer Science and Information Systems

II SEMESTER 2017-2018

CS F212 – Database Systems

07th May, 2018   Comprehensive Exam   Weightage: 40%

PARTS – B & C (30%-60 marks; Open book - text book, hand-written class notes, printed lecture slides, & lab notes provided)

PART B

1. Consider the schema:

    Registration_Performance_Current_Semester
    (IDNO,course_no,semester,Academic_year,registered_category,Grade)
    Registration_Performance_Past
    (IDNO,course_no,semester,Academic_year, registered_category,Grade)

    The first relation contains grades of students in the current semester and the second relation contains grades of students in previous semesters. Records of only continuing students are kept in these two relations. Due to some reason the registered_category (Open_EL, Disc_EL, Repeat,…) of some students is not known at the time of populating the first relation. For calculating CGPA, it is important to know if a student is repeating a course. Write a relational algebra expression to create a temporary relation to find out students who are repeating courses along with the courses they are repeating. The schema should be Repeat (IDNO, course_no). What type of database object this temporary relation should be if this exercise needs to be carried out in every semester?

    [10]

2. Consider the interleaved schedule of two transactions, T1 & T2. Transaction T1 is transferring amount 100 from account X to account Z, whereas transaction T2 is summing up the balances of accounts X and Z. The balance in accounts X and Z before the transactions begin is 500 and 250, respectively.

Time	T1	T2
t1		Begin Tx
t2	Begin Tx	Sum=0
t3	Read(balX)	
t4	balx=balx-100	Read(balX)
t5	Write(balx)	Sum+=balx
t6	Read(balZ)	
t7	balz=balz+100	
t8	Write(balz)	
t9	Commit	Read(balz)
t10		Sum+=balz
t11		Write(sum)
t12		Commit

(a) Assuming that accounts X & Z are on different blocks, BX and BZ, respectively. Issue appropriate instructions so that the transactions can start working.

(b) Construct and populate a table with schema (time, bal_X of T1, bal_X of T2, bal_X in memory, bal_Z of T1, bal_Z of T2, bal_Z in memory, SUM in memory)

(c) What will be the value of SUM that will eventually be displayed to user? Did the transactions produce the expected result? Justify your answer.

[10]

3. In extensible hashing, the capacity of buckets is 4. If we start with 4 buckets which are completely full, what is the maximum/minimum number of buckets we will have after 8 inserts. Also give the respective size of the directory, the global depth, minimum and maximum values for local depth. Justify your answer. Assume that redistribution will lead to almost equal distribution of records among the bucket and its split image.

[10]

4. Suppose we store a relation $R(x,y)$ in a grid file. Both attributes have a range from 0 to 300. The partitions of the grid file are uniformly spaced. For x, partitions every 20 units, and for y, every 50 units. In order to perform the Nearest-Neighbor query for the point P(110,205), we start by searching the bucket with lower left corner at (100,200) and upper right corner at (120,250). The closest point in this bucket is Q(115,220). What other buckets must be searched to check whether this point is the nearest neighbor or not? How many I/Os will be required for answering the query, assuming that there are no overflow chains.

[10]

PART C
(use a separate answer sheet)

Consider the following schema related to hotel reservation:
City (Zip, Name, State)
Hotel (Hno, Name, Pin no., Address, max_capacity)
Room (Roomno, Hno, Type, Price)
Reservation (Rid, Gid, Roomno, Hno, Type, Arrival-Date, Departure-Date)
Guest (Gid, Firstname, Lastname, Pin no., Address)
Employee (Hno, Eno, Firstname, Lastname, Manager_eno.)
Booking_Type (Type, Description)

Q1. Create a view containing the Firstname and Lastname for each Guest in the XYZ hotel who will arrive exactly after 3 days.

[5]

Q2. Create a view containing the hotel name starting with "S" and ends with "T" along with the names of the guest staying in that particular hotel. Also, write a separate query on the created view to obtain the total no. of Guest staying in each hotel.

[5]

Q3. Write a trigger to generate a message if a hotel is fully booked. Also, write a supporting Sql query to check the correctness of your trigger.

[5]

Q4. Write a TSQL query to print the ID of all the reservations whose reservation Type is MAP and Hno is GH7X2.

[5]

***