**School of Computer Science and Engineering**

**FALL SEMESTER – 2024-25**

**Continuous Assessment Test (CAT) – II - Answer key**

**Programme:** B.Tech – CSE **Course Code:** BCSE302L **Course Title:** Database Systems
**Slot:** A1+TA1 **Exam Date:** 13-10-2024 **Exam Time:** 9.30AM to 11.00AM

**1. a. Check R1 and R2 for BCNF, if it is not in BCNF, convert it into BCNF: (6m)**

R1 (elevator_no, building_no, capacity, building_name)
R2 (elevator_no, building_no, capacity)
FD= {elevator_no→building_no, capacity and building_no→ building_name}

**Answer:**

R1 is not in 3NF because of non-key dependencies.
non-key dependency building_no → building_name, that is the non-key attribute building_no uniquely determines the other non-key attribute building_name
To convert R1 into 3NF relations, decompose R1 on non-key dependencies. Thus, we get the following relations;
R1a (elevator_no, building_no, capacity)
***From FD elevator_no → building_no, capacity***
R1b (building_no, building_name)
***From FD building_no → building_name a non-key dependency.***
determinants of the functional dependencies are the primary keys.
Hence the relations R1a and R1b are in BCNF.
R1a = R2, R2 is in BCNF.

**1. b. Identify the MVDs present in the table and normalize it to 4NF. (4m)**

| Name | Areacode | Phone | likes | manufacturer |
|---|---|---|---|---|
| Abi | 632014 | 555-1111 | Strawberry | Tata |
| Abi | 632014 | 555-1111 | chocolate | Britania |
| Abi | 632224 | 555-1199 | Strawberry | Tata |
| Abi | 632224 | 555-1199 | chocolate | Britania |

4NF
R1 → Name ->> Area code, Phone
R2 → Name ->> likes, Manufacturer

---

2. Consider the block size of 256 bytes and block pointer of size 6 byte and search key value of 5bytes and 3bytes for primary index and clustering index on an ordered field respectively. Size of each entry in the data file is 20bytes with 50000 records and record pointer of size 5 bytes. Clustering index has 500 zipcode vslues.

i) Find the blocking factor for the data file, primary and clustering index.

$bfr = \lfloor (B/R) \rfloor = \lfloor (256/20) \rfloor$ = 12 records/block for data file
$bfr_i = \lfloor (B/R_i) \rfloor = \lfloor (256/11) \rfloor$ = 23 records/block for primary index file (5+6 =11)
$bfr_i = \lfloor (B/R_i) \rfloor = \lfloor (256/9) \rfloor$ = 28 records/block for clustering index file (3+6 =9)

ii) Find the number of blocks required for data file, primary and clustering index

$b = \lceil (r/bfr) \rceil = \lceil (50{,}000/12) \rceil = 4167$ blocks for data file
$b_i = \lceil (r_i/bfr_i) \rceil = \lceil (4167/23) \rceil = 181$ blocks for primary index file
$b_i = \lceil (r_i/bfr_i) \rceil = \lceil (500/28) \rceil = 18$ blocks for clustering index file

iii) For the primary and clustering index - calculate the number of accesses needed.

$\lceil(\log_2 b_i)\rceil = \lceil(\log_2 181)\rceil = 7$ block accesses.
$\lceil(\log_2 b_i)\rceil = \lceil(\log_2 18)\rceil = 4$ block accesses.

iv) To calculate the order $p$ of a $B^+$- tree: : $(p * P) + ((p - 1) * V) \leq B$

$(p * 6) + ((p - 1) * 3) \leq 256$ ➔ $p = 29$,

Order $p_{leaf}$ for the leaf nodes is: : $(p_{leaf} * (Pr + V)) + P \leq B$

$(p_{leaf} * (5 + 3)) + 6 \leq 256$ ➔ $p_{leaf} = 32$

On average, each internal node will have 29 * 0.69 or approximately 20 pointers, 19 values.
Each leaf node, on average, will hold 0.69 * $p_{leaf}$ = 0.69 * 32 or 22 data record pointers.

| | | | |
|---|---|---|---|
| Root node → | 1 node | 19 key entries | 20 pointers |
| Internal node 1 → | 20nodes | 380 entries | 400pointers |
| Leaf node → | 400 nodes | 8800 data record pointers | |

==========

3. Derive equivalent relational Algebra expression/SQL for the following:

i) $\pi_{\text{movieTitle}}\sigma_{\substack{\text{starName=name}\\ \wedge\text{birthdate=1960}}}(\text{MovieStar})$

ii) $\pi_{\text{movieTitle}}\sigma_{\text{starName=M.name}}(\text{StarsIn} \times \rho_M\pi_{\text{name}}\sigma_{\text{birthdate=1960}}(\text{MovieStar}))$

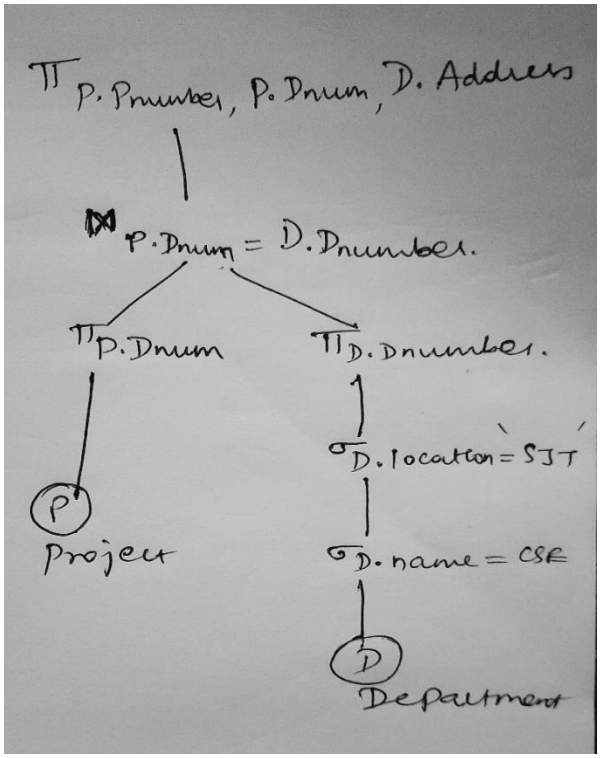
iii) $\pi_{\text{fname,lname}}$ ( $\sigma_{\text{dno = 4 } \wedge \text{ salary > 50000}}$ (employee) )

iv) Select Id,Name From Person Where Hobby ='stamps' OR Hobby ='coins'

v) Select StudId, CrsCode From Transcript Where Grade ! = 'F'

==========

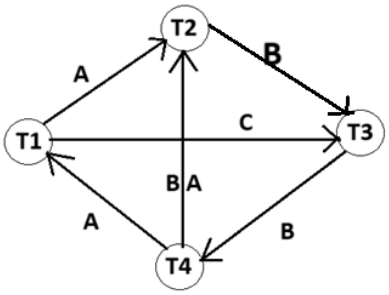
4. Find the optimal query tree for the SQL query and explain the steps in detail.
Select P.Pnumber, P.Dnum, D.Address
From Project as P, Department as D
Where P.Dnum=D.Dnumber And D.location='SJT' And D.name=CSE




==========

5. Draw the precedence graph for the schedule of 4 transactions: W3(B), R1(C), W4(B), W4(A), R1(A), W3(C), W2(A), R2(D), W2(B), R3(D). Identify whether it is conflict serializable or not? Present the equivalent serial schedules.

precedence graph and explanation – 6m




conflict serializable or not – 4m

This schedule isn't conflict serializable because there is two cycles (T1, T3, T4) and (T1, T2, T3, T4) in the graph. Moreover, this schedule is not conflict equivalent (every pair of conflicting operations is ordered in the same way) to any serial schedule of transaction execution.

equivalent serial schedules - In general, as there are 4 transactions, there is 4! no of Serial Schedules possible = 24. But, as there exists cycle in the graph, this cannot be converted to any serial schedule.