Indian Statistical Institute

M.Tech. (CS), Mid-Semester of Second Semester Examination, 2025-26

Database Management Systems

Answer Keys

1. (a) There are $t/(1+5) = t/6$ matches from R_1 to R_2. There are $t/(1+3) = t/4$ matches from R_2 to R_1. Then every matching tuple of R_1 joins with every matching tuple of R_2. This gives us the maximum possible number of tuples as $(t/6)*(t/4) = t/24$. Note that the distinct 10% of $t = t/10$ elements in the common attribute does not restrict the maximum.

   (b) For two non-identical relations R_1 and R_2, there might either be a tuple t (i) in R_1 but not in R_2 or (ii) in R_2 but not in R_1. For (i), t will get selected in $R_1 - R_2$ but not in $R_2 - R_1$ and for (ii), t will get selected in $R_2 - R_1$ but not in $R_1 - R_2$. Hence, proved.

2. (i) $\pi_{T_1.team\_owner}\left(\sigma_{(T_1.team\_owner=T_2.team\_owner)\wedge(T_1.team\_id\neq T_2.team\_id)}\left(\rho_{T_1}(\text{TEAMS}) \times \rho_{T_2}(\text{TEAMS})\right)\right)$.

   (ii) $\pi_{team\_name}\left(\text{TEAMS} - (\text{TEAMS} \rhd\lhd_{team\_id=team\_winner} \text{RESULT})\right)$.

3. (a) (i) Efficient using a B^+ tree but not hash table because it will require range search.

   (ii) Efficient using a hash table but not B^+ tree because it will require point search.

   (b) The size of the part $\sigma_{(X \leq 150) \vee (Z > 0)}(R_1)$ is

   $$\frac{(120)^2 - \left(120 - 120 * \frac{150 - 100 + 1}{200 - 100 + 1}\right)\left(120 - 120 * \frac{0 - (-9) + 1}{10 - (-9) + 1}\right)}{120} \approx 90,$$ and the size of the part $\sigma_{(Y\ !=\ 10)}(R_2)$ is $(20 - 1) = 19$. As the common attribute between R_1 and R_2 is a key for R_2, hence the estimated size is bounded by the first part of the semijoin, i.e., 90.

4. It has to decomposed into 3NF subrelations for satisfying all the constraints with guarantee. The candidate key is VZ. Violations happen for $V \rightarrow W$, $W \rightarrow X$, $X \rightarrow Y$, $Y \rightarrow W$. Hence, the decomposed relations are $<V, W>$, $<V, Z>$ and $<W, X, Y>$.

Moreover, it can be decomposed into some BCNF subrelations, which might satisfy the required constraints. For satisfying BCNF, violations happen for $V \rightarrow W$, $W \rightarrow X$, $X \rightarrow Y$, $Y \rightarrow W$. Hence, the decomposed relations are $<V, W>$, $<V, Z>$ and $<W, X, Y>$.