Indian Statistical Institute

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

Database Management Systems

Full Marks: 30 Date: 25-02-2026 Time: 2 Hours

Answer any *three* of the following questions

$3 \times 10 = 30$

1. (a) Suppose the pair relations R_1 and R_2, both having t number of tuples, share a common attribute A. Further assume that the ratio between the number of tuples returned by the operations $(R_1 \ltimes R_2)$ and $(R_1 \triangleright R_2)$ is 1:5 and $(R_2 \ltimes R_1)$ and $(R_2 \triangleright R_1)$ is 1:3. If the number of unique elements in A independently for R_1 and R_2 are exactly 10% of the total elements, then what will be the maximum number of tuples returned by the operation $R_1 \bowtie R_2$? Recall that $R_1 \ltimes R_2$ and $R_1 \triangleright R_2$ denote the semijoin and antijoin operations, respectively.

   (b) Prove or disprove the following statement:
   "For two non-identical relations R_1 and R_2, $(R_1 - R_2) \neq (R_2 - R_1)$ will always hold."

6+4

2. Consider a Cricket League RRL that follows a single Round-Robin structure, in which every team plays every other team once. The final Champion and Runner-up are decided based on the maximum and second maximum number of wins, respectively. The following schema represents the details associated with the entire league. The attributes *team1*, *team2* and *team_winner* all take *team_id* values as entries. Note that the primary keys in the respective tables are underlined therein.

   TEAMS = <*team_id*: integer, *team_name*: string, *team_owner*: string>
   MATCH = <*match_id*: integer, *team1*: integer, *team2*: integer>
   RESULT = <*match_id*: integer, *team_winner*: integer>

   Write relational algebra expressions to return the following things.
   (i) Name(s) of the owner(s) who own multiple teams.
   (ii) Name(s) of the team(s) who never won a match in the RRL.

5+5

3. (a) Suppose a relation $R = <X, Y, Z>$ is to be indexed. Compare the performance of using a B^+-tree and a hash table to execute the following operations in an effective way. Note that the underlined attribute is the primary key in the relation.
   (i) $\sigma_{(Y > 0)}(R)$
   (ii) $\sigma_{(X = 0) \vee (Z != 0)}(R)$

   (b) Consider the relations R_1 = <X, Y, Z> and R_2 = <Y, Z> having 120 and 20 tuples, respectively. Notably the primary keys in the respective relations are underlined therein. The number of distinct values of Y in R_1 and R_2 are the same. The range of values of X

and Z in R_1 are [100, 200] and [-9, 10], respectively. Estimate the size of the following query considering the attribute values are uniformly distributed.

$$\sigma_{(X \leq 150) \vee (Z > 0)} (R_1) \ltimes \sigma_{(Y != 10)} (R_2)$$

(2+2)+6

4. Let there be a relation $R = <V, W, X, Y, Z>$ having all the attributes as atomic and holding the functional dependencies FD = $\{V \rightarrow W, W \rightarrow X, X \rightarrow Y, Y \rightarrow W, VZ \rightarrow X\}$. Decompose R into subrelations (at least two) such that it satisfies all the following requirements.
   (R1) The redundancy is eliminated from R as much as possible.
   (R2) The decomposition is lossless join.
   (R3) The decomposition is dependency preserving.

   Show with explanations how the aforementioned requirements are satisfied in your proposed decomposition.

6+4
