

# CS 245 Database Systems Principles – Summer 2001

## Assignment 4

- Due Date: Monday July 30<sup>th</sup> 2001, 5 pm.
  - Submission through
    - Box Placed outside Gates 412.
    - Email solution to [masood@cs.stanford.edu](mailto:masood@cs.stanford.edu) (only pdf, ps or plain text files).
    - SITN homework delivery.
  - The deadline is hard, No Late days.
  - Do not forget to write your Leland Ids at the start of your solution.
  - State all assumptions.
  - Email questions to [cs245-staff@lists.stanford.edu](mailto:cs245-staff@lists.stanford.edu)
- 

### Problem 1 (30 points)

Suppose we have two relations: R(a, b, c) and S(c, d, e). Give 6 valid logical plans (either in relational algebra or as logical plan trees) for the following SQL query:

```
SELECT    B, C, D
FROM      R, S
WHERE     R.c = S.c AND R.a = 5
```

### Problem 2 (15 points)

Suppose we have two relations: R(a, b, c) and S(c, d, e). Consider the following relational algebra expression:

$$\Pi_{a,d} [\sigma_{S.e=5 \wedge R.a=10 \wedge R.c=S.c} (R \times S)]$$

Transform this expression into an equivalent expression that:

- Has no Cartesian products
- Performs projections as early as possible
- Performs selections as early as possible

An “early” selection/projection is a selection/projection that occurs as close as possible to the leaves in the logical plan tree form of the expression. Give projections priority over selections.

### Problem 3 (35 points)

Suppose relations R and S are not stored in any particular order. Suppose furthermore that two indexes exist on relation R, one on column a and one on column c. No other indexes are present. Give 7 valid physical plans (as physical plan trees) for the following logical plan:

$$[\sigma_{a=5}(R)] \bowtie_{R.c=S.c} [S]$$

### Problem 4 (20 points)

Here are the statistics about four relations: W, X, Y, and Z:

W(a, b)	X(b, c)	Y(c, d)	Z(d, e)
T(W) = 100	T(X) = 400	T(Y) = 200	T(Z) = 300
V(W, a) = 80	V(X, b) = 200	V(Y, c) = 200	V(Z, d) = 200
V(W, b) = 10	V(X, c) = 1	V(Y, d) = 120	V(Z, e) = 80

Estimate the sizes of the following expressions:

- a)  $\sigma_{a=35}(W)$  (assume that 35 is present in column a of relation W)
- b)  $\sigma_{a=35 \wedge b=5}(W)$  (assume that 35 is present in column a and 5 is present in column b)
- c)  $W \bowtie X$  (natural join)
- d)  $X \bowtie Y$
- e)  $W \bowtie X \bowtie Y \bowtie Z$