Assignment 4

Query Optimization

Ilaria Battiston - 03723403 Mareva Zenelaj - 03736071

1 First exercise

1.1 Query Graph

This query graph assumes four relations R_1 , R_2 , R_3 , R_4 , each of them with cardinality 10. We are considering the following query graph:

$$R_1 - R_2$$

$$\mid \qquad \mid$$

$$R_3 - R_4$$

The selectivities of the query graph are:

- $R_1 \bowtie R_2 = 0.4$;
- $R_1 \bowtie R_3 = 0.5$;
- $R_2 \bowtie R_4 = 0.49$;
- $R_3 \bowtie R_4 = 0.6$.

The algorithm GreedyOperatorOrdering will perform the following steps:

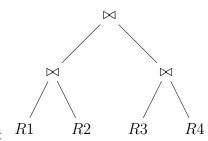
- 1. Join R_1 and R_2 ;
- 2. Join R_3 and R_4 ;
- 3. Join the two results.

$$C_{out}((R_1 \bowtie R_2) \bowtie (R_3 \bowtie R_4)) = 40 + 60 + 588 = 688$$

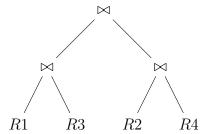
However, there exists another join ordering giving a better result:

$$C_{out}((R_1 \bowtie R_3) \bowtie (R_2 \bowtie R_4)) = 50 + 49 + 588 = 687$$

1.2 Join Tree



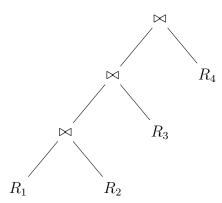
Execution of GreedyOperatorOrdering:



Optimal join tree: R1

1.3 GreedyJoinOrdering-1

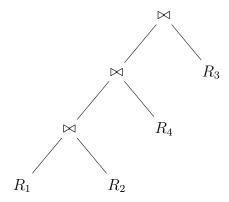
In this example the proposed weight function corresponds to the cardinality. However, all relations have the same cardinality, therefore it is assumed that the chosen ordering is the natural one:



$$C_{out}(((R_1 \bowtie R_2) \bowtie R_3) \bowtie R_4) = 40 + 200 + 588 = 828$$

1.4 GreedyJoinOrdering-2

The algorithm joins according to selectivity:



$$C_{out}(((R_1 \bowtie R_2) \bowtie R_4) \bowtie R_3) = 40 + 196 + 588 = 824$$

1.5 GreedyJoinOrdering-3

In this example we only show the optimal join tree with all the possible C_{out} combinations:

- Starting from R_1 : $C_{out}(((R_1 \bowtie R_2) \bowtie R_4) \bowtie R_3) = 40 + 196 + 588 = 824;$
- Starting from R_2 : $C_{out}(((R_2 \bowtie R_2) \bowtie R_4) \bowtie R_3) = 40 + 196 + 588 = 824;$
- Starting from R_3 : $C_{out}(((R_3 \bowtie R_1) \bowtie R_2) \bowtie R_4) = 50 + 200 + 588 = 838;$
- Starting from R_4 : $C_{out}(((R_4 \bowtie R_2) \bowtie R_1) \bowtie R_3) = 49 + 196 + 588 = 833.$

Clearly, the minimum value is achieved starting from R_1 or R_2 , hence the optimal tree is:

