TU München, Fakultät für Informatik
Lehrstuhl III: Datenbanksysteme
Prof. Dr. Thomas Neumann

Query Optimization

2. Exercise
Due November 6, 2017, 9 AM
submit via email (radke@in.tum.de)

Exercise 1

For relations R_1, R_2, predicates p_1, p_2, $\mathcal{F}(p_1) \subseteq \mathcal{A}(R_1)$ and assuming set semantics.

- Prove the following equivalence:
  $\sigma_{p_1}(R_1 \bowtie_{p_2} R_2) = \sigma_{p_1}(R_1) \bowtie_{p_2} R_2$
- Does the equivalence also hold for outer joins? Justify your answer.
  $\sigma_{p_1}(R_1 ⟗_{p_2} R_2) \stackrel{?}{=} \sigma_{p_1}(R_1) ⟗_{p_2} R_2$

Exercise 2

Given $|R_1|$, $|R_2|$, the domain of $R_1.x$ and $R_2.y$, and the information if $R_1.x$ and/or $R_2.y$ are keys of R_1 and R_2.

1. How can we estimate the selectivity of $\sigma_{R_1.x=c}$, where c is a constant?
2. How can we estimate the selectivity of $\bowtie_{R_1.x=R_2.y}$?

Note that we don't know the output size of $\sigma_{R_1.x=c}$ ($\bowtie_{R_1.x=R_2.y}$, respectively), so we can't simply use the definition of selectivity.

Exercise 3

Given are two relations R and S, with sizes 1,000 and 100,000 pages respectively. Each page has 50 tuples. The relations are stored on a disk, the average access time for the disk is 10 ms and the transfer speed is 10,000 pages/sec. How long does it take to perform the Nested Loops Join of R and S? How long does it take to perform the Block Nested Loops Join with a block size of 100 pages? Assume that CPU costs are negligible and ignore I/O costs for the join output.