

Übung zur Vorlesung *Datenbanksysteme II* im SS 20

Gabriel Haas (gabriel.haas@uni-jena.de)

dbis1.github.io/courses/ss20/db2/

Blatt Nr. 07

Abgabe auf moodle.uni-jena.de. Die Bedingungen für den Notenbonus finden Sie auf der Vorlesungswebsite.

Hausaufgabe 1

Für einen Join-Baum T sei folgende Kostenfunktion gegeben

$$C_{out}(T) = \begin{cases} 0 & \text{falls } T \text{ eine Basisrelation } R_i \text{ ist} \\ |T| + C_{out}(T_1) + C_{out}(T_2) & \text{falls } T = T_1 \bowtie T_2 \end{cases}$$

Die Kardinalität sei dabei

$$|T| = \begin{cases} |R_i| & \text{falls } T \text{ eine Basisrelation } R_i \text{ ist} \\ \left(\prod_{R_i \in T_1, R_j \in T_2} f_{i,j} \right) |T_1| |T_2| & \text{falls } T = T_1 \bowtie T_2 \end{cases}$$

Sei $p_{i,j}$ das Join Prädikat zwischen R_i und R_j , dann sei

$$f_{i,j} = \frac{|R_i \bowtie_{p_{i,j}} R_j|}{|R_i \times R_j|}$$

und die Kardinalität eines Join-Resultats ist $|R_i \bowtie_{p_{i,j}} R_j| = f_{i,j} |R_i| |R_j|$.

Gegeben sei eine Anfrage über die Relationen R_1, R_2, R_3 und R_4 mit $|R_1| = 10, |R_2| = 20, |R_3| = 20, |R_4| = 10$. Die Selektivitäten der Joins seien $f_{1,2} = 0.01, f_{2,3} = 0.5, f_{3,4} = 0.01$, alle nicht gegebenen Selektivitäten sind offensichtlich 1 (Warum?). Berechnen Sie den optimalen (niedrigste Kosten) Join-Tree. Als Vereinfachung reicht es, wenn Sie nur Joins mit Prädikat und keine Kreuzprodukte betrachten.

Hausaufgabe 2

Zu welcher Joinreihenfolge würde man mit dem Greedy Operator Ordering Algorithmus kommen? Verwenden Sie die Zwischenergebnisse aus Aufgabe 1.