



Übung zur Vorlesung *Grundlagen: Datenbanken* im WS23/24

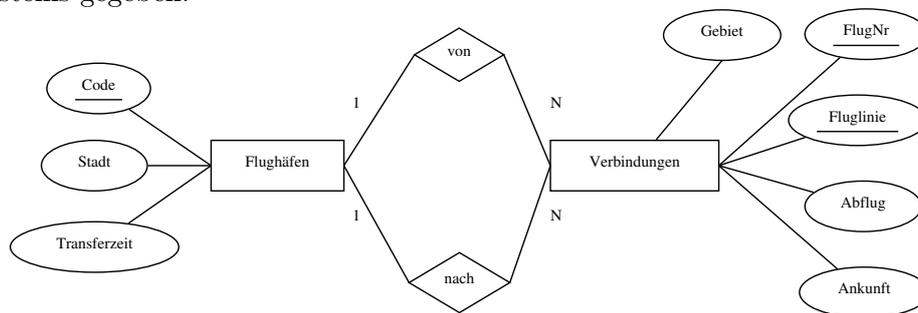
Christoph Anneser, Michael Jungmair, Stefan Lehner, Moritz Sichert, Lukas Vogel
 (gdb@in.tum.de)

<https://db.in.tum.de/teaching/ws2324/grundlagen/>

Blatt Nr. 12

Hausaufgabe 1

Betrachten Sie die Anfrage „Finde alle Flüge von New-York nach Sydney mit einmaligem Umsteigen“. Dazu sei das nachfolgende (vereinfachte) ER-Diagramm eines Fluginformationssystems gegeben:



- Geben Sie eine SQL-Query für die oben genannte Anfrage an.
- Führen Sie die kanonische Übersetzung des SQL-Statements in die relationale Algebra durch.
- Schätzen Sie die Relationsgrößen sinnvoll ab (z.B. so wie in den Beispielen der Vorlesung) und transformieren Sie den kanonischen Operatorbaum aus Teilaufgabe b) zur optimalen Form. Wie haben sich die Kosten dabei geändert? (Kosten = Anzahl der Zwischenergebnistupel)

Hausaufgabe 2

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

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

Sei $p_{i,j}$ das Joinprädikat 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 3

Gegeben sei die Anfrage:

```
select *
  from R, S, T
 where R.A = S.A and S.B = T.B and T.C = R.A
```

Des Weiteren soll gelten:

- S.A und T.C seien Fremdschlüssel auf R
 - S.B sei Fremdschlüssel auf T
 - R.A, T.B seien Primärschlüssel von R respektive T
 - Ihre Query-Engine unterstützt nur Nested-Loop-Joins
 - Kardinalitäten: $|R| = 100, |S| = 1000, |T| = 10$
 - Es gibt keine Indexe
- a) Für Equijoins zwischen R_i mit Fremdschlüssel auf den Primärschlüssel in R_j gilt die Abschätzung:

$$f_{i,j} = \frac{1}{|R_j|}$$

Warum?

- b) Bestimmen Sie, wie in der Vorlesung gezeigt, den optimalen Ausführungsplan als Baum mit Kosten-/Kardinalitätsabschätzungen mit Hilfe von Dynamischem Programmieren. Verwenden Sie die Kostenfunktion C_{out} .

Hausaufgabe 4

Geben Sie die kanonische Übersetzung der folgenden Anfrage in die relationale Algebra an. Verwenden Sie zur Darstellung des relationalen Algebraausdrucks die Baumdarstellung.

```
select v.VorlNr, v.Titel, p.Name, count(h.MatrNr) as hoerer
from
  Vorlesungen v left outer join
  hoeren h on (v.VorlNr = h.VorlNr),
  Professoren p
where
  v.gelesenVon = p.PersNr
group by v.VorlNr, v.Titel, p.Name
having count(h.MatrNr) > 3
```