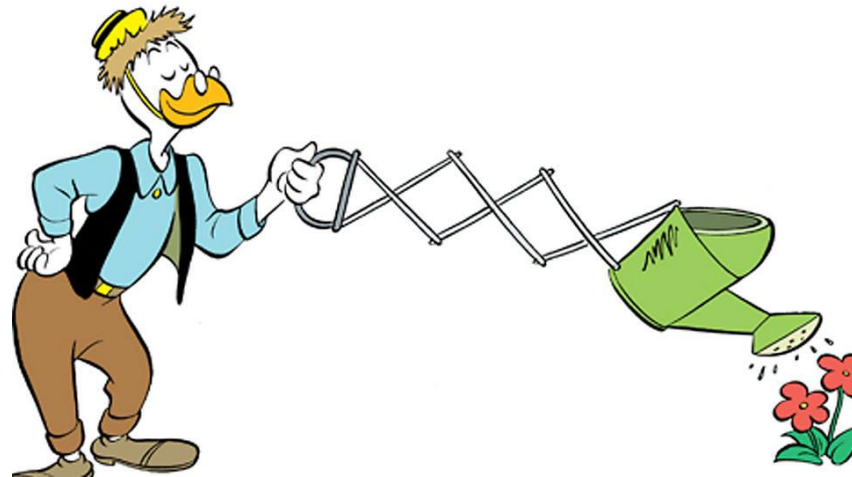


Übung Datenbanksysteme II

Anfrageoptimierung

Maximilian Jenders

Folien basierend auf
Thorsten Papenbrock

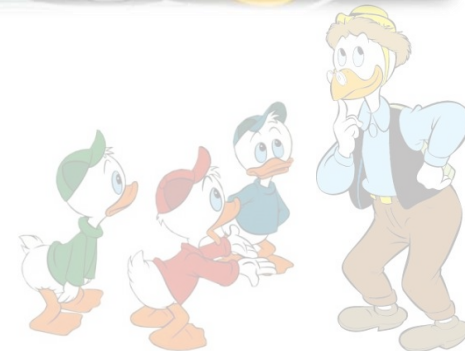


2

- Algebraische Transformation



- Kardinalitätsschätzung



- Join-Reihenfolge



3

- Gegeben: $R(a,b,c)$ und $S(c,d,e)$
- Gesucht: Kostengünstigste Anfragepläne für folgende Anfragen

d.h. möglichst kleine Zwischenergebnisse, also Selektionen und Projektionen so früh wie möglich

a. $\sigma_{b=3 \wedge e=4 \wedge c>10} (R \bowtie S)$

$$\sigma_{c>10 \wedge b=3} (R) \bowtie \sigma_{c>10 \wedge e=4} (S)$$

b. $\pi_{a,d} (R \bowtie S)$

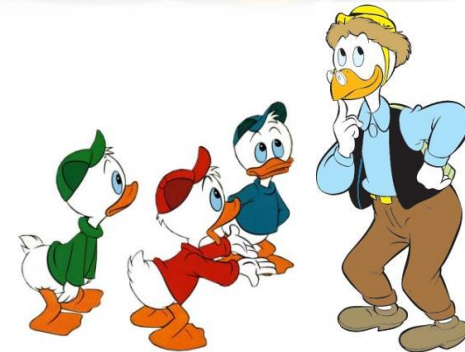
$$\pi_{a,d} (\pi_{a,c} (R) \bowtie \pi_{c,d} (S))$$

4

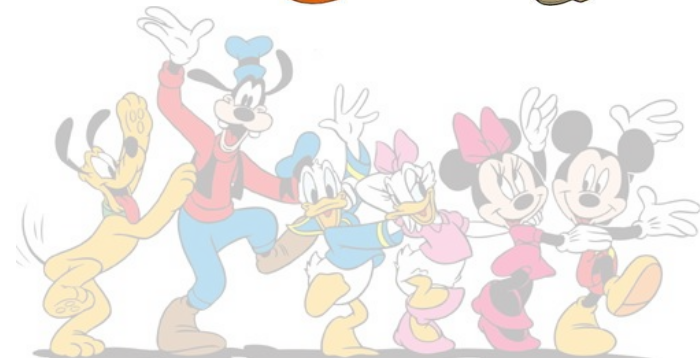
- Algebraische Transformation



- Kardinalitätsschätzung



- Join-Reihenfolge



5

- Gegeben: $R(a,b,c,d)$ und $S(d,e)$
 $T(R)=100$; $V(R,a)=100$; $V(R,b)=10$; $V(R,c)=1$; $V(R,d)=50$
 $T(S)=500$; $V(S,d)=30$; $V(S,e)=100$

- Gesucht: Geschätzte Ergebniskardinalität für folgende Anfragen

a. $\sigma_{b=25}(R)$

$$T(R)/V(R,b) = 10$$

b. $\sigma_{c=30}(R)$

$$T(R)/V(R,c) = 100$$

c. $\sigma_{b=25 \wedge c=30}(R)$

$$T(R)/(V(R,b) \cdot V(R,c)) = 10$$

d. $\sigma_{b>25}(R)$

$$T(R)/3 \approx 33$$

e. $\sigma_{a>30 \wedge b=10}(R)$

$$T(R)/(3 \cdot V(R,b)) \approx 3$$

f. $\sigma_{b>25 \wedge b=11}(R)$

0 (widersprüchliche Selektion)

g. $\sigma_{b=25 \vee d=13}(R)$

$$T(R)/V(R,b) + T(R)/V(R,d) - T(R)/(V(R,b) \cdot V(R,d)) = 11,8 \approx 12$$

h. $R \bowtie S$

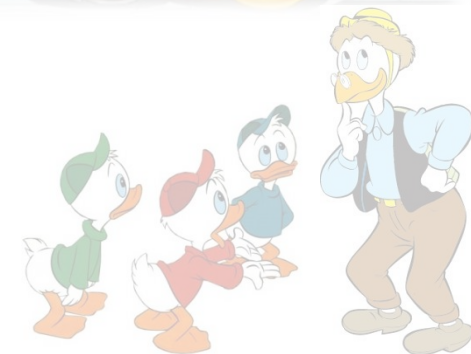
$$T(R) \cdot T(S) / \max[V(R,d), V(S,d)] = 1000$$

6

- Algebraische Transformation



- Kardinalitätsschätzung



- Join-Reihenfolge



Aufgabe 3: Join-Reihenfolge

7

- Gegeben: Folgende Relationen und deren Statistiken

$W(A, B)$	$X(B, C)$	$Y(C, D)$	$Z(D, A)$
$T(W) = 100$	$T(X) = 200$	$T(Y) = 300$	$T(Z) = 400$
$V(W, A) = 20$			$V(Z, A) = 100$
$V(W, B) = 60$	$V(X, B) = 50$		
	$V(X, C) = 100$	$V(Y, C) = 50$	
		$V(Y, D) = 50$	$V(Z, D) = 40$

- Gesucht: Optimale Join-Reihenfolge für $W \bowtie X \bowtie Y \bowtie Z$
- Bestimme die Join-Reihenfolge als Left Deep Tree. Nutze dazu Dynamische Programmierung und gib die Tabellen aller Zwischenschritte an.

Aufgabe 3: Join-Reihenfolge

$W(A,B)$	$X(B,C)$	$Y(C,D)$	$Z(D,A)$
$T(W) = 100$	$T(X) = 200$	$T(Y) = 300$	$T(Z) = 400$
$V(W,A) = 20$			$V(Z,A) = 100$
$V(W,B) = 60$	$V(X,B) = 50$	$V(Y,C) = 50$	
	$V(X,C) = 100$	$V(Y,D) = 50$	$V(Z,D) = 40$

8

	{W}	{X}	{Y}	{Z}
Kardinalität	100	200	300	400
Kosten	0	0	0	0
Opt. Plan	W	X	Y	Z

Kleinste Relation links

	{W,X}	{W,Y}	{W,Z}	{X,Y}	{X,Z}	{Y,Z}
Kard.	$100 \cdot 200$ $/60 =$ 33333	$100 \cdot 300$ $=$ 30000	$100 \cdot 400$ $/100 =$ 400	$200 \cdot 300$ $/100 =$ 600	$200 \cdot 400$ $=$ 80000	$300 \cdot 400$ $/50 =$ 2400
Kosten	0	0	0	0	0	0
Opt. Plan	$W \bowtie X$	$W \bowtie Y$	$W \bowtie Z$	$X \bowtie Y$	$X \bowtie Z$	$Y \bowtie Z$

Aufgabe 3: Join-Reihenfolge

$W(A,B)$	$X(B,C)$	$Y(C,D)$	$Z(D,A)$
$T(W) = 100$	$T(X) = 200$	$T(Y) = 300$	$T(Z) = 400$
$V(W,A) = 20$			$V(Z,A) = 100$
$V(W,B) = 60$	$V(X,B) = 50$	$V(Y,C) = 50$	
	$V(X,C) = 100$	$V(Y,D) = 50$	$V(Z,D) = 40$

9

	$\{W,X\}$	$\{W,Y\}$	$\{W,Z\}$	$\{X,Y\}$	$\{X,Z\}$	$\{Y,Z\}$
Kard.	$100 \cdot 200$ $/60 =$ $333,33$	$100 \cdot 300$ $=$ 300	$100 \cdot 400$ $/100 =$ 400	$200 \cdot 300$ $/100 =$ 600	$200 \cdot 400$ $=$ 800	$300 \cdot 400$ $/50 =$ 2400
Kosten	0	0	0	0	0	0
Opt. Plan	$W \bowtie X$	$W \bowtie Y$	$W \bowtie Z$	$X \bowtie Y$	$X \bowtie Z$	$Y \bowtie Z$

Kardinalität ist für alle
Joinreihenfolgen gleich!

Kosten: (Kardinalität + Kosten)
des Zwischenergebnisses

	$\{W,X,Y\}$	$\{W,X,Z\}$	$\{W,Y,Z\}$	$\{X,Y,Z\}$
Kardinalität	$333,33 \cdot 300$ $/100 =$ 1000	$333,33 \cdot 400$ $/100 =$ $1333,33$	$400 \cdot 300$ $/50 =$ 2400	$600 \cdot 400$ $/50 =$ 4800
Kosten	333,33	333,33	400	600
Opt. Plan	$(W \bowtie X) \bowtie Y$	$(W \bowtie X) \bowtie Z$	$(W \bowtie Z) \bowtie Y$	$(X \bowtie Y) \bowtie Z$

Aufgabe 3: Join-Reihenfolge

$W(A,B)$	$X(B,C)$	$Y(C,D)$	$Z(D,A)$
$T(W) = 100$	$T(X) = 200$	$T(Y) = 300$	$T(Z) = 400$
$V(W,A) = 20$			$V(Z,A) = 100$
$V(W,B) = 60$	$V(X,B) = 50$	$V(Y,C) = 50$	
	$V(X,C) = 100$	$V(Y,D) = 50$	$V(Z,D) = 40$

10

	$\{W,X\}$	$\{W,Y\}$	$\{W,Z\}$	$\{X,Y\}$	$\{X,Z\}$	$\{Y,Z\}$
Kard.	$100 \cdot 200 / 60 = 333,33$	$100 \cdot 300 = 30000$	$100 \cdot 400 / 100 = 400$	$200 \cdot 300 / 100 = 600$	$200 \cdot 400 = 80000$	$300 \cdot 400 / 50 = 2400$
Kosten	0	0	0	0	0	0
Opt. Plan	$W \bowtie X$	$W \bowtie Y$	$W \bowtie Z$	$X \bowtie Y$	$X \bowtie Z$	$Y \bowtie Z$

	$\{W,X,Y\}$	$\{W,X,Z\}$	$\{W,Y,Z\}$	$\{X,Y,Z\}$
Kardinalität	$333,33 \cdot 300 / 100 = 1000$	$333,33 \cdot 400 / 100 = 1333,33$	$400 \cdot 300 / 50 = 2400$	$600 \cdot 400 / 50 = 4800$
Kosten	333,33	333,33	400	600
Opt. Plan	$(W \bowtie X) \bowtie Y$	$(W \bowtie X) \bowtie Z$	$(W \bowtie Z) \bowtie Y$	$(X \bowtie Y) \bowtie Z$

Aufgabe 3: Join-Reihenfolge

$W(A,B)$	$X(B,C)$	$Y(C,D)$	$Z(D,A)$
$T(W) = 100$	$T(X) = 200$	$T(Y) = 300$	$T(Z) = 400$
$V(W,A) = 20$			$V(Z,A) = 100$
$V(W,B) = 60$	$V(X,B) = 50$	$V(Y,C) = 50$	
	$V(X,C) = 100$	$V(Y,D) = 50$	$V(Z,D) = 40$

11

	$\{W,X,Y\}$	$\{W,X,Z\}$	$\{W,Y,Z\}$	$\{X,Y,Z\}$
Kardinalität	$333,33 \cdot 300$ /100 = 1000	$333,33 \cdot 400$ /100 = 1333,33	$400 \cdot 300$ /50 = 2400	$600 \cdot 400$ /50 = 4800
Kosten	333,33	333,33	400	600
Opt. Plan	$(W \bowtie X) \bowtie Y$	$(W \bowtie X) \bowtie Z$	$(W \bowtie Z) \bowtie Y$	$(X \bowtie Y) \bowtie Z$

	$\{W,X,Y,Z\}$	$\{W,X,Y,Z\}$	$\{W,X,Y,Z\}$	$\{W,X,Y,Z\}$
Kard.	$1000 \cdot 400$ /(100·50) = 80	$1333,33 \cdot 300$ /(100·50) = 80	$2400 \cdot 200$ /(100·60) = 80	$4800 \cdot 100$ /(100·60) = 80
Kosten	1333,33	1666,66	2800	5400
Plan	$((W \bowtie X) \bowtie Y) \bowtie Z$	$((W \bowtie X) \bowtie Z) \bowtie Y$	$((W \bowtie Z) \bowtie Y) \bowtie X$	$((X \bowtie Y) \bowtie Z) \bowtie W$



12

X(a,b)	Y(a,b)	Z(a,b)
T(X)=100	T(Y)=200	T(Z)=300
V(X,a)=10	V(Y,a)=20	V(Z,a)=30
V(X,b)=3	V(Y,b)=2	V(Z,b)=1

$$X \bowtie Y = 100 \cdot 200 / (20 \cdot 3) = 333,33$$

$$(X \bowtie Y) \bowtie Z = 333,33 \cdot 300 / (30 \cdot \mathbf{2}) = 1666,66$$

- *Preservation of Value Sets* gilt nicht, wenn das Attribut ein Join-Attribut war!
- Wegen *Containment of Value Sets* wissen wir aber, dass nach dem Join $V(X \bowtie Y, a) = 10$ und $V(X \bowtie Y, b) = 2$ sein müssen.

Allgemein:

$$X \bowtie Y \bowtie Z = T(X) \cdot T(Y) \cdot T(Z) / [V(Y,a) \cdot V(Z,a)] \cdot [V(X,b) \cdot V(Y,b)]$$

Für jedes Join-Attribut teile durch alle Kardinalitäten außer der Kleinsten!