

Aufgabenblatt 5

— Anfrageoptimierung —

Ausgabe am 23.06.2008
Abgabe bis 07.07.2008, 13.00 Uhr

Aufgabe 1: Algebraische Transformationen (6 P)

Zeigen Sie durch Angabe eines Gegenbeispiels, dass folgende Gleichungen *nicht* gelten, wobei R, S Mengen und R', S' Multimengen seien. Mit einem Stern markierte Mengenoperatoren seien deren Erweiterung auf Multimengen.

- (a) $\pi_L(R \cup S) = \pi_L(R) \cup \pi_L(S)$ (1 P)
- (b) $\pi_L(R \cap S) = \pi_L(R) \cap \pi_L(S)$ (1 P)
- (c) $\pi_L(R' \cap^* S') = \pi_L(R') \cap^* \pi_L(S')$ (1 P)
- (d) $\pi_L(R - S) = \pi_L(R) - \pi_L(S)$ (1 P)
- (e) $\pi_L(R' -^* S') = \pi_L(R') -^* \pi_L(S')$ (1 P)
- (f) $\delta(\pi_L(R')) = \pi_L(\delta(R'))$ (1 P)

Aufgabe 2: Kardinalitätsschätzungen (16 P)

Gegeben seien die vier Relationen $W(a, b), X(b, c), Y(c, d)$ und $Z(d, e)$, für die folgende Statistiken vorliegen:

$$\begin{array}{llll} T(W) = 100 & T(X) = 200 & T(Y) = 300 & T(Z) = 400 \\ V(W, a) = 20 & V(X, b) = 50 & V(Y, c) = 50 & V(Z, d) = 40 \\ V(W, b) = 60 & V(X, c) = 100 & V(Y, d) = 50 & V(Z, e) = 100 \end{array}$$

Schätzen Sie die Größe der folgenden Relationen ab, indem Sie die in der Vorlesung vorgestellten Methoden anwenden.

- (a) $W \bowtie X \bowtie Y \bowtie Z$ (2 P)
- (b) $\sigma_{a=10}(W)$ (2 P)
- (c) $\sigma_{c=20}(Y) \bowtie Z$ (2 P)
- (d) $W \times Y$ (2 P)
- (e) $\sigma_{d>10}(Z)$ (2 P)
- (f) $\sigma_{a=1 \wedge b=2}(W)$ (2 P)

- (g) $\sigma_{a=1 \wedge b > 2}(W)$ (2 P)
 (h) $X \bowtie_{X.c < Y.c} Y$ (2 P)

Aufgabe 3: Histogramme (5 P)

Schätzen Sie die Größe der Relation $R(a, b) \bowtie S(b, c)$ unter Verwendung von Histogrammen für $R.b$ und $S.b$ ab. Nehmen Sie an, dass $V(R, b) = 20$, $V(S, b) = 20$ und die Histogramme jeweils die Anzahl der vier am häufigsten auftretenden Attributwerte angeben.

Attributwert	0	1	2	3	4	andere Werte
$R.b$	5	6	4	5		32
$S.b$	10	8	5		7	48

Vergleichen Sie die berechnete Abschätzung mit der einfacheren Abschätzung, die man unter der Annahme, dass alle 20 Attributwerte gleichverteilt sind und $T(R) = 52$ sowie $T(S) = 78$ gilt, erhält.

Aufgabe 4: Optimale Reihenfolge von Join-Operationen (8 P)

Betrachten Sie erneut die in Aufgabe 2 angegebenen Relationen W, X, Y und Z . Berechnen Sie die Kosten für jede denkbare Join-Reihenfolge der Form $((R_1 \bowtie R_2) \bowtie R_3) \bowtie R_4$ und bestimmen Sie daraufhin die optimale Join-Reihenfolge, die dieser Form genügt. Nutzen Sie den in der Vorlesung vorgestellten Bottom-Up Algorithmus und notieren Sie die berechneten Zwischenergebnisse.

Aufgabe 5: Links- und rechts-tiefe Join-Bäume (6 P)

Gegeben seien acht Relationen R_1, \dots, R_8 . Berechnen Sie die Anzahl der

- (a) links-tiefen Join-Bäume (*left-deep join trees*) (2 P)
 (b) rechts-tiefen Join-Bäume (*right-deep join trees*) (2 P)
 (c) Join-Bäume, die weder links- noch rechts-tief sind (*bushy join trees*) (2 P)

für den Join dieser Relationen. Notieren Sie jeweils den Rechenweg.