11.8 Többutas összekapcsolási algoritmus

QUEL (INGRES DBMS) lekérdezés optimalizálója:
Wong-Youssefi algoritmus

Elvek:

0) σ és Π minél hamarabb
1) az összekapcsolások és szorzatok ($\bowtie$, $\times$) heurisztikus rendezése
2) $\bowtie$ átalakítása $\bowtie$($\ltimes$) semi-join után join-ra
3) Több relációs összekapcsolások speciális kezelése

Először 3)-mal foglalkozunk.

$$Q(A,B) \bowtie R(B,C) \bowtie S(C,D)$$

- kommutatív és asszociatív szabály szerint tetszőleges sorrendben kapcsolhatunk össze

RCSSZ: $(Q(A,B) \bowtie S(C,D)) \bowtie R(B,C)$

Egyformán jó: $(Q \bowtie R) \bowtie S$ vagy
$Q \bowtie (R \bowtie S)$

A két összekapcsolást egyszerre elvégezve még jobb eredményt kaphatunk:

FELTEVÉS: $|Q| = |R| = |S| = T_0$ sor
tömör elhelyezés esetén B_0 blokk

További feltevés: $I_0 \leq B_0$ (I_0 minden oszlop képmérete egyenlő)

Klaszterindex: Q.B-re

Ekkor: S.C-re

$Q \bowtie R$ költsége:

$$\left(B_0 + \frac{T_0 B_0}{I_0}\right) + \frac{B_0 T_0 + T_0 B_0}{I_0} = B_0 + \frac{3 B_0 T_0}{I_0}$$

$(Q \bowtie R) \bowtie S$ költsége: $B_{Q \times R}\left(1 + \frac{T_S}{I_0}\right) + \frac{2 T_{Q \bowtie R} \cdot B_S}{I_0}$

$$\frac{2 B_0 T_0}{I_0} + \frac{2 B_0 T_0^2}{I_0^2} + \frac{2 \cdot \frac{T_0^2}{I_0} \cdot B_0}{I_0} =$$

$$\frac{2 B_0 T_0}{I_0} + \frac{4 B_0 T_0^2}{I_0^2}$$

Összköltség

$$\left(B_0 + \frac{3 B_0 T_0}{I_0}\right) + \frac{2 B_0 T_0}{I_0} + \frac{4 B_0 T_0^2}{I_0^2} =$$

$$B_0 + \frac{5 B_0 T_0}{I_0} + \frac{4 B_0 T_0^2}{I_0^2}$$

$Q \bowtie R$-t közben fölöslegesen kiírjuk, majd újra beolvasnuk. Rögtön a memóriában összekapcsoljuk ~~az~~ az abc-t $\sigma_{C=c}(S)$-vel.

Valódi költség

$$B_0 + \frac{B_0 T_0}{I_0} + \frac{4 B_0 T_0^2}{I_0^2}$$

Ki4t- eredményt ad.

for ~~minden~~ az R minden bc sorára do

$$\sigma_{B=b}(Q) \ltimes \{bc\} \rtimes \sigma_{C=c}(S)$$

Feltesszük, hogy a fenti összekapcsolás -hoz szükséges sorok beférnek a memóriába: $\frac{2B_0}{I_0} \leq M$

Költség:

Ciklusmagot T_0-szor hajtjuk végre

$\sigma_{B=b}(Q) \ltimes \{bc\} \rtimes \sigma_{C=c}(S)$ beolvasása:

$$T_0 \frac{B_0}{I_0} + B_0 + \frac{T_0 B_0}{I_0} = B_0 + \frac{2B_0 T_0}{I_0}$$ az összes

beolvasási költség

Output költség: $$\frac{\frac{2B_0T_0}{I_0} \cdot T_0 + \frac{T_0^2}{I_0} B_0}{I_0} = \frac{3B_0T_0^2}{I_0^2}$$

Összköltség:

$$B_0 + \frac{2B_0T_0}{I_0} + \frac{3B_0T_0^2}{I_0^2}$$

$\frac{T_0}{I_0} \to \infty$ esetén az utóbbi költség $\frac{3}{4}$-e az előző-nek.

Általánosítva: (dekompozíciós összekapcsolás)

INPUT: $R, S_1, \ldots, S_n$ $\quad S_i \cap S_j = \emptyset$ (nincs közös attribútum)
(csillag) $\quad R \cap S_i \neq \emptyset$

OUTPUT: $R \bowtie S_1 \bowtie \ldots \bowtie S_n$

Módszer

for az R minden t sorára do begin
  for $i = 1$ to n do
    $T_i := S_i \ltimes \{t\}$,
    /* S_i azon sorai, amelyek $S_i \cap R$-n megegyeznek t-vel */
  Output $\{t\} \bowtie T_1 \bowtie \ldots \bowtie T_n$
end

11.9 Lekérdezések hipergráfos reprezentációja

$\sigma_{F_1 \wedge \ldots \wedge F_n}(R_1 \times \ldots \times R_k)$ lekérdezés reprezentációja

Csúcsok: $R_i.A_j$ attribútumok

$F_1 \wedge \ldots \wedge F_n$ maximális konjunktív lánc

ha F_j: $A = B$, akkor a két csúcsot összevonjuk (tranzitív módon)