

FORMULARIO SU OTTIMIZZAZIONE DI INTERROGAZIONI

Selezione - Stima del fattore di selettività

- $F(A=v) = 1/V(A,R)$
[se per qualche motivo $V(A,R)$ non è noto, si assume $F=1/10$]
- $F(A>v) = (\text{Max}(A,R) - v)/(\text{Max}(A,R)-\text{Min}(A,R))$
- $F(A<v) = (v-\text{Min}(A,R))/(\text{Max}(A,R)-\text{Min}(A,R))$
[se non si conoscono i valori di max e min dell'attributo o l'attributo non è numerico, si stima $F=1/3$]
- $F(A \text{ IN } (v_1, v_2, \dots, v_N)) = N * F(A=v)$
- $F(A \text{ BETWEEN } (v_1, v_2)) = (v_2 - v_1)/(\text{Max}(A,R)-\text{Min}(A,R))$
[se non si conoscono i valori di max e min dell'attributo o l'attributo non è numerico, si stima $F=1/4$]
- $F(A1 = A2) = 1/ \text{MAX} (V(A1,R), V(A2,R))$
[se i valori non sono noti si assume $1/10$]
- $F(C_1 \text{ AND } C_2) = F(C_1) * F(C_2)$
- $F(C_1 \text{ OR } C_2) = F(C_1) + F(C_2) - F(C_1) * F(C_2)$
- $F(\text{NOT } C) = 1 - F(C)$

Stima della dimensione del risultato

Proiezione (senza eliminazione dei duplicati):

$$T(R) * (S(A1,R) + \dots + S(An,R))$$

Selezione:

$$T(R) * F(P) * S(R)$$

Prodotto Cartesiano:

$$T(R) * T(S) * (S(R) + S(S))$$

Join naturale:

cardinalità nel caso $R1 \cap R2$ non vuoto e non chiave né di $R1$ né di $R2$

- se $R1$ e $R2$ hanno un unico attributo comune $A: T(r1) * T(r2) / \max(V(A,r1), V(A,r2))$
- se $R1$ e $R2$ hanno due attributi comuni A, B :
 - $T(r1) * T(r2) / \max(V(A,r1), V(A,r2)) * \max(V(B,r1), V(B,r2))$

Eliminazione dei duplicati:

$$\text{Min}(T(R)/2, \prod_i V(A_i(R)) * S(R))$$

Interrogazioni su singola relazione – costo globale di accesso

- $B(R)$ se si sceglie la scansione sequenziale
- $F(C_A) * [L(I) + B(R)]$
se si sceglie un indice I clusterizzato
- $F(C_A) * [L(I) + V(A,R) * \Phi(\text{NTK}(A), B(R), T(R))]$
con $\text{NTK}(A) = T(R) / V(A;R)$
se si sceglie un indice I non clusterizzato
[N.B. Φ è sicuramente minore di tutti i suoi argomenti]