

Appendice A

Formulario

Fattore di selettività di una condizione

Condizione (ψ)	Fattore di selettività
$A_i = c$	$f_s(\psi) = \begin{cases} 1/N_{key}(A_i) & \text{se esiste un indice su } A_i \\ 1/10 & \text{altrimenti} \end{cases}$
$A_i = A_j$	$f_s(\psi) = \begin{cases} \frac{1}{\max\{N_{key}(A_i), N_{key}(A_j)\}} & \text{se esistono indici su } A_i, A_j \text{ e} \\ & \text{ } adom(A_i) \subseteq adom(A_j) \text{ o} \\ & \text{ } adom(A_j) \subseteq adom(A_i) \\ 0 & \text{se } adom(A_i) \text{ e } adom(A_j) \text{ sono disgiunti} \\ 1/N_{key}(A_i) & \text{se esiste un indice solo su } A_i \\ 1/10 & \text{altrimenti} \end{cases}$
$A_i > c$	$f_s(\psi) = \begin{cases} \frac{\max(A_i) - c}{\max(A_i) - \min(A_i)} & \text{se } A_i \text{ di tipo numerico e con indice} \\ 1/3 & \text{altrimenti} \end{cases}$
$A_i < c$	$f_s(\psi) = \begin{cases} \frac{c - \min(A_i)}{\max(A_i) - \min(A_i)} & \text{se } A_i \text{ di tipo numerico e con indice} \\ 1/3 & \text{altrimenti} \end{cases}$
$A_i < A_j$	$f_s(\psi) = 1/3$
A_i BETWEEN c_1 AND c_2	$f_s(\psi) = \begin{cases} \frac{c_2 - c_1}{\max(A_i) - \min(A_i)} & \text{se } A_i \text{ di tipo numerico e con indice} \\ 1/4 & \text{altrimenti} \end{cases}$
A_i IN (v_1, \dots, v_n)	$f_s(\psi) = \begin{cases} n \times f_s(A_i = v) & \text{se minore di } 1/2 \\ 1/2 & \text{altrimenti} \end{cases}$
NOT ψ_1	$f_s(\psi) = 1 - f_s(\psi_1)$
ψ_1 AND ψ_2	$f_s(\psi) = f_s(\psi_1) \times f_s(\psi_2)$
ψ_1 OR ψ_2	$f_s(\psi) = f_s(\psi_1) + f_s(\psi_2) - f_s(\psi_1) \times f_s(\psi_2)$

2 Appendice A Formulario

Operatori fisici: descrizione

Operatore logico	Operatore fisico	Descrizione	
Tabella	TableScan (R)	Scansione di R .	
R	IndexScan (R, Idx)	Scansione ordinata di R sugli attributi dell'indice Idx .	
	IndexSequentialScan (R, Idx)	Scansione di R ordinata sulla chiave primaria con l'organizzazione sequenziale con indice.	
	SortScan ($R, \{A_i\}$)	Scansione di R ordinata sugli $\{A_i\}$.	
Proiezione $\pi_{\{A_i\}}^b$	Project ($O, \{A_i\}$)	Proiezione dei record di O senza l'eliminazione dei duplicati.	
$\pi_{\{A_i\}}$	Distinct (O)	Eliminazione dei duplicati dai record di O ordinati sugli $\{A_i\}$.	
Restrizione	Filter (O, ψ)	Restrizione dei record di O .	
σ_ψ	IndexFilter (R, Idx, ψ)	Restrizione di R con l'indice Idx e accessi a R .	
	IndexSequentialFilter (R, Idx, ψ)	Restrizione sulla chiave primaria di R memorizzata con l'organizzazione sequenziale con indice.	
	IndexOnlyFilter ($R, Idx, \{A_i\}, \psi$)	Restrizione di R con l'indice Idx senza accessi a R . L'operatore ritorna i record di R che soddisfano ψ , con attributi gli $\{A_i\}$, un sottoinsieme degli attributi sui quali è definito l'indice.	
Giunzione	NestedLoop (O_E, O_I, ψ_J)	Giunzione con l'algoritmo <i>nested loop</i> .	
	PageNestedLoop (O_E, O_I, ψ_J)	Giunzione con l'algoritmo <i>page nested loop</i> .	
	\bowtie ψ_J	IndexNestedLoop (O_E, O_I, ψ_J)	Giunzione con l'algoritmo <i>index nested loop</i> ; O_I utilizza un indice definito sugli attributi di giunzione.
		MergeJoin (O_E, O_I, ψ_J)	Giunzione con l'algoritmo <i>merge join</i> . I record degli operandi O_E e O_I sono ordinati sugli attributi di giunzione, chiave in O_E e chiave esterna in O_I .
Ordinamento $\tau_{\{A_i\}}$	Sort ($O, \{A_i\}$)	Ordinamento dei record di O sugli $\{A_i\}$.	
Raggruppamento $\{A_i\} \gamma \{f_i\}$	GroupBy ($O, \{A_i\}, \{f_i\}$)	Raggruppamento dei record di O ordinati sugli $\{A_i\}$ usando le funzioni di aggregazione in $\{f_i\}$. L'operatore ritorna record con attributi gli A_i e le funzioni di aggregazione in $\{f_i\}$, ordinati sugli $\{A_i\}$.	
Operatori insiemistici $\cup, -, \cap$	Union, Except, Intersect (O_E, O_I)	Operatori insiemistici con i record degli operandi ordinati e privi di duplicati.	
	UnionAll (O_E, O_I)	Unione con duplicati dei record degli operandi.	

Costo e cardinalità del risultato (E_{rec}) degli operatori fisici

Operatori fisici per tabelle

Operatore	Costo	Cardinalità del risultato
TableScan (R)	$N_{\text{pag}}(R)$	$E_{\text{rec}} = N_{\text{rec}}(R)$
IndexScan (R, Idx)		
<i>indice non di ordinamento</i>	$N_{\text{leaf}}(Idx) + N_{\text{rec}}(R)$	
<i>indice di ordinamento</i>	$N_{\text{leaf}}(Idx) + N_{\text{pag}}(R)$	
IndexSequentialScan (R, Idx)	$N_{\text{leaf}}(Idx)$	
SortScan ($R, \{A_i\}$)	$4 \times N_{\text{pag}}(R) + N_{\text{pag}}(R)$	

Per ordinare un insieme di record si suppone di usare un algoritmo che (a) opera sui dati che sono in un file (b) ordina i dati con un solo passo di fusione, con costo $4 \times N_{\text{pag}}(R)$, e (c) lascia i dati ordinati in un file che vanno poi letti con costo $N_{\text{pag}}(R)$.

Operatori fisici per la proiezione

Operatore	Costo	Cardinalità del risultato
Project ($O, \{A_i\}$)	$C(O)$	$E_{\text{rec}} = E_{\text{rec}}(O)$
Distinct (O)	$C(O)$	Sia $\{A_i\}$ l'insieme degli attributi di O .
		$E_{\text{rec}} = \begin{cases} N_{\text{key}}(A_i) & \text{se } \{A_i\} = 1 \\ N_{\text{rec}}(O) & \text{se una chiave di } O \in \{A_i\} \\ \min\{E_{\text{rec}}(O), \prod N_{\text{key}}(A_i)\} & \text{se noti gli } N_{\text{key}}(A_i) \\ E_{\text{rec}}(O) & \text{altrimenti} \end{cases}$

Operatori fisici per la restrizione

Operatore	Costo	Cardinalità del risultato
Filter (O, ψ)	$C(O)$	$E_{\text{rec}} = \lceil f_s(\psi) \times E_{\text{rec}}(O) \rceil$
IndexFilter (R, Idx, ψ)	C_A^I	$E_{\text{rec}} = \lceil f_s(\psi) \times N_{\text{rec}}(R) \rceil$
IndexSequentialFilter (R, Idx, ψ)	$\lceil f_s(\psi) \times N_{\text{leaf}}(Idx) \rceil$	
IndexOnlyFilter ($R, Idx, \{A_i\}, \psi$)	$\lceil f_s(\psi) \times N_{\text{leaf}}(Idx) \rceil$	

Se Idx è di ordinamento,

$$C_A^I = C_I + C_D = \lceil f_s(\psi) \times N_{\text{leaf}}(Idx) \rceil + \lceil f_s(\psi) \times N_{\text{pag}}(R) \rceil$$

altrimenti,

$$C_A^I = \lceil f_s(\psi) \times N_{\text{leaf}}(Idx) \rceil + \lceil f_s(\psi) \times N_{\text{key}}(A_i) \rceil \times \lceil \Phi(N_{\text{rec}}(R)/N_{\text{key}}(A_i), N_{\text{pag}}(R)) \rceil$$

4 Appendice A Formulario

Operatori fisici per la giunzione

Operatore	Costo	Cardinalità del risultato
NestedLoop (O_E, O_I, ψ_J)	$C(O_E) + E_{\text{rec}}(O_E) \times C(O_I)$	$E_{\text{rec}} = \lceil f_s(R.A_i = S.A_j) \times E_{\text{rec}}(O_E) \times E_{\text{rec}}(O_I) \rceil$
PageNestedLoop (O_E, O_I, ψ_J)	$C(O_E) + N_{\text{pag}}(O_E) \times C(O_I)$	
IndexNestedLoop (O_E, O_I, ψ_J)	$C(O_E) + E_{\text{rec}}(O_E) \times C_A^I(O_I)$	
MergeJoin (O_E, O_I, ψ_J)	$C(O_E) + C(O_I)$	

Operatore fisico per l'ordinamento

Operatore	Costo	Cardinalità del risultato
Sort $(O, \{A_i\})$	$C(O) + 6 \times N_{\text{pag}}(O)$	$E_{\text{rec}} = E_{\text{rec}}(O)$

Per ordinare il risultato di O , si suppone di memorizzarlo in un file, con costo $N_{\text{pag}}(O)$, di usare un algoritmo che (a) opera sui dati che sono in un file (b) ordina i dati con un solo passo di fusione, con costo $4 \times N_{\text{pag}}(O)$, e (c) lascia i dati ordinati in un file che vanno poi letti con costo $N_{\text{pag}}(O)$.

Operatore fisico per il raggruppamento

Operatore	Costo	Cardinalità del risultato
GroupBy $(O, \{A_i\}, \{f_i\})$	$C(O)$	$E_{\text{rec}} = \langle \text{come Distinct} \rangle$

Operatori fisici per le operazioni insiemistiche

Operatore	Costo	Cardinalità del risultato
Union (O_E, O_I)	$C(O_E) + C(O_I)$	$E_{\text{rec}} = E_{\text{rec}}(O_E) + E_{\text{rec}}(O_I)$
Except (O_E, O_I)	$C(O_E) + C(O_I)$	$E_{\text{rec}} = E_{\text{rec}}(O_E)$
Intersect (O_E, O_I)	$C(O_E) + C(O_I)$	$E_{\text{rec}} = \langle \text{come } \bowtie \text{ con attributi di giunzione tutti gli attributi} \rangle$
UnionAll (O_E, O_I)	$C(O_E) + C(O_I)$	$E_{\text{rec}} = E_{\text{rec}}(O_E) + E_{\text{rec}}(O_I)$