

Operazioni dell'algebra relazionale

Classificazione

a linguaggi formali

Algebra relazionale
Calcolo relazionale
Programmazione logica

b linguaggi programmatici

SQL: Structured Query Language
QBE: Query By Example

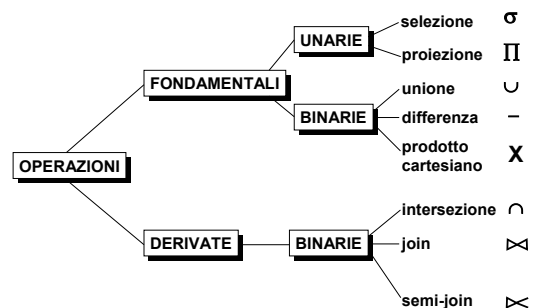
2

Algebra relazionale

- definita da Codd (70)
- molto utile per imparare a formulare query
- insieme minimo di 5 operazioni che danno l'intero potere espressivo del linguaggio

3

Una visione d'insieme



4

Esempio : gestione degli esami universitari

studente

MATR	NOME	CITTA'	INDIR
123	Carlo	Bologna	Inf
415	Paola	Torino	Inf
702	Antonio	Roma	Log

esame

MATR	COD-CORSO	DATA	VOTO
123	1	7-9-03	10
123	2	8-1-03	8
702	2	7-9-03	5

corso

COD-CORSO	TITOLO	DOCENTE
1	matematica	Barozzi
2	informatica	Meo

5

Selezione

$\sigma_{\text{Nome}='Paola'}$ STUDENTE

e' una tabella (priva di nome) con

- schema :
lo stesso schema di STUDENTE
- istanza :
le tuple di STUDENTE che soddisfano il predicato di selezione

Matr	Nome	Città	CDip
415	Paola	Torino	Inf

6

Sintassi del predicato di selezione

espressione booleana di predicati semplici

operazioni booleane :

- AND (P1 AND P2) (\wedge)
- OR (P1 OR P2) (\vee)
- NOT (P1) (\neg)

predicati semplici :

- TRUE, FALSE
- termine
- comparatore
- termine

comparatore :

- =, !=, <, <=, >, >=

termine :

- costante, attributo
- espressione aritmetica di costanti e attributi

7

Esempio di selezione

σ

STUDENTE

(Città='Torino') OR
((Città='Roma')
AND NOT (CDip='Log'))

MATR	NOME	CITTA'	INDIR
123	Carlo	Bologna	Inf
415	Paola	Torino	Inf
702	Antonio	Roma	Log

8

Proiezione

$\Pi_{\text{Nome, CDip}}$ STUDENTE

è una tabella (priva di nome) con

- schema :
gli attributi Nome e CDip
- istanza :
la restrizione delle tuple sugli attributi Nome e CDip

Nome	CDip
Carlo	Inf
Paola	Inf
Antonio	Log

9

Proiezione e duplicati

- nel modello formale la proiezione elimina i duplicati

Π_{CDip} STUDENTE

CDip
Inf
Log

- nel modello informale (e nei sistemi) la eliminazione dei duplicati va richiesta esplicitamente

10

Assegnamento

- serve per dare un nome al risultato di una espressione algebrica
- non fa parte delle operazioni algebriche

INFORMATICI = $\sigma_{\text{CDip}='Inf'}$ STUDENTI

TORINESI = $\sigma_{\text{Città}='Torino'}$ STUDENTI

11

Unione

TABELLA1 \cup TABELLA2

si può fare se TABELLA1 e TABELLA2 sono compatibili

- con lo stesso grado
- oppure (nei sistemi)
- con domini ordinatamente dello stesso tipo

12

Unione

INFORMATICI \cup TORINESI

è una tabella (priva di nome) con

- schema :
lo schema di INFORMATICI
- istanza :
la unione delle tuple di
INFORMATICI e TORINESI

Matr	Nome	Città	CDip
123	Carlo	Bologna	Inf
415	Paola	Torino	Inf

Per quanto riguarda le istanze, è commutativa

13

Differenza

TABELLA1 - TABELLA2

si può fare se TABELLA1 e TABELLA2 sono compatibili

14

Differenza

INFORMATICI - TORINESI

è una tabella (priva di nome) con

- schema :
lo schema di INFORMATICI
- istanza :
la differenza delle tuple di
INFORMATICI e TORINESI

Matr	Nome	Città	CDip
123	Carlo	Bologna	Inf

Attenzione: non è commutativa

15

Prodotto cartesiano

$R \times S$

è una tabella (priva di nome) con

- schema :
gli attributi di R e S
($\text{grado}(R \times S) = \text{grado}(R) + \text{grado}(S)$)
- istanza :
tutte le possibili coppie di tuple di R e S
($\text{card}(R \times S) = \text{card}(R) \times \text{card}(S)$)

16

Esempio

R1(A,B)

A	B
a	1
b	3

R2(C,D)

C	D
c	1
b	3
a	2

$R1 \times R2$ (A,B,C,D)

A	B	C	D
a	1	c	1
a	1	b	3
a	1	a	2
b	3	c	1
b	3	b	3
b	3	a	2

17

Intersezione

TABELLA1 \cap TABELLA2

Come gli altri operatori insiemistici, si può fare se TABELLA1 e TABELLA2 sono compatibili

Derivabile tramite la seguente formula:

$$R \cap S = R - (R - S)$$

18

Intersezione

INFORMATICI \cap TORINESI

è una tabella (priva di nome) con

- schema :
lo schema di INFORMATICI
- istanza :
la intersezione delle tuple di INFORMATICI e TORINESI

Matr	Nome	Città	CDip
415	Paola	Torino	Inf

19

Join

STUDENTE \bowtie STUDENTE.Matr=ESAME.Matr ESAME

è equivalente alla seguente espressione (operatore derivato):

$\sigma_{\text{STUDENTE.Matr=ESAME.Matr}}$ STUDENTE \times ESAME

attributi omonimi sono resi non ambigui usando la notazione “puntata”:
ESAME.Matr, STUDENTE.Matr

20

Join

STUDENTE \bowtie STUDENTE.Matr=ESAME.Matr ESAME

produce una tabella (priva di nome) con

- schema :
la concatenazione degli schemi di STUDENTE e ESAME
- istanza:
le tuple ottenute concatenando quelle tuple di STUDENTE e di ESAME che soddisfano il predicato

STUDENTE. Matr	Nome	Città	CDip	ESAME. Matr	Cod- Corso	Data	Voto
123	Carlo	Bologna	Inf	123	1	7-9-03	10
123	Carlo	Bologna	Inf	123	2	8-1-03	8
702	Antonio	Roma	Log	702	2	7-9-03	5

21

Sintassi del predicato di join

espressione congiuntiva di predicati semplici:

ATTR1 comp ATTR2

ove ATTR1 appartiene a TAB1
ATTR2 appartiene a TAB2
comp: =, !=, <, <=, >, >=

EQUI-JOIN :
soli confronti di uguaglianza

22

Join naturale

equi-join di tutti gli attributi omonimi (si omette il predicato, si elimina la colonna ripetuta)

STUDENTE \bowtie ESAME

Matr	Nome	Città	CDip	Cod- Corso	Data	Voto
123	Carlo	Bologna	Inf	1	7-9-03	10
123	Carlo	Bologna	Inf	2	8-1-03	8
702	Antonio	Roma	Log	2	7-9-03	5

23

Join naturale di tre tabelle

STUDENTE \bowtie ESAME \bowtie CORSO

Matr	Nome	Città	CDip	Cod- Corso	Data	Voto	Titolo	Docente
123	Carlo	Bologna	Inf	1	7-9-03	10	matem	barozzi
123	Carlo	Bologna	Inf	2	8-1-03	8	infor	meo
702	Antonio	Roma	Log	3	7-9-03	5	infor	meo

24

Semi-Join

STUDENTE $\bowtie \leftarrow$ **STUDENTE**.Matr=**ESAME**.Matr **ESAME**

$\Pi_{Attr(Studiante)}$ **STUDENTE**
 $\bowtie \leftarrow$ **STUDENTE**.Matr=**ESAME**.Matr **ESAME**

produce una tabella (priva di nome) con

- schema :
- lo schema di **STUDENTE**
- istanza:

le tuple ottenute proiettando su **STUDENTE** il join di **STUDENTE** e di **ESAME**

Matr	Nome	Città	CDip
123	Carlo	Bologna	Inf
702	Antonio	Roma	Log

25

Semi-Join Naturale

STUDENTE $\bowtie \leftarrow$ **ESAME** =

$\Pi_{Attr(Studiante)}$ **STUDENTE** $\bowtie \leftarrow$ **ESAME**

Proietta sulla tabella **STUDENTE** il join naturale di **STUDENTE** e **ESAME**

Matr	Nome	Città	CDip
123	Carlo	Bologna	Inf
702	Antonio	Roma	Log

26

Equivalenza di espressioni

- quali studenti hanno preso 10 in matematica?

Π_{Nome}
 (**STUDENTE** $\bowtie \leftarrow$
 ($\sigma_{Voto=10}$ **ESAME** $\bowtie \leftarrow$
 ($\sigma_{Titolo='matematica'}$ **CORSO**)))

- equivalente a:

$\Pi_{Nome} \sigma_{Voto=10 \wedge Titolo='matematica'}$
 (**STUDENTE** $\bowtie \leftarrow$ **ESAME** $\bowtie \leftarrow$ **CORSO**)

27

Equivalenza di espressioni

- quali professori hanno esaminato Antonio?

$\Pi_{Docente}$
 (**CORSO** $\bowtie \leftarrow$
 (**ESAME** $\bowtie \leftarrow$
 ($\sigma_{Nome='Antonio'}$ **STUDENTE**)))

- equivalente a:

$\Pi_{Docente} \sigma_{Nome='Antonio'}$
 (**STUDENTE** $\bowtie \leftarrow$ **ESAME** $\bowtie \leftarrow$ **CORSO**)

28

Espressioni complesse

- estrarre il nome degli studenti che non hanno mai preso meno di 8

Π_{Nome} **STUDENTE** $\bowtie \leftarrow$
 (Π_{Matr} **ESAME**
 -
 $\Pi_{Matr} \sigma_{Voto < 8}$ **ESAME**)

- spiegazione: prima trovo le matricole di tutti gli studenti meno le matricole di coloro che hanno preso meno di 28, poi trovo i loro nomi.

29

Espressioni complesse

- estrarre il nome degli studenti che hanno sostenuto "informatica" e "matematica" lo stesso giorno

Π_{Nome} **STUDENTE** $\bowtie \leftarrow$
 ((**ESAME** $\bowtie \leftarrow$ $\sigma_{Titolo='informatica'}$ **CORSO**)
 $\bowtie \leftarrow$ $\text{Matr} = \text{Matr} \wedge \text{Data} = \text{Data}$
 (**ESAME** $\bowtie \leftarrow$ $\sigma_{Titolo='matematica'}$ **CORSO**))

- spiegazione: prima trovo le matricole di coloro che hanno dato i due esami nello stesso giorno, poi trovo i loro nomi.

30

Espressioni complesse

- estrarre l'ultimo esame di ciascuno studente

ESAME

-

(ESAME $\triangleright \triangleleft$ | Matr = Matr \wedge Data < Data ESAME)

- spiegazione: prima trovo gli esami che non sono ultimi, cioè che sono seguiti da qualche esame a pari studente in data superiore, poi sottraggo da tutti gli esami questi "esami non ultimi" e trovo gli ultimi esami di ciascuno studente

31

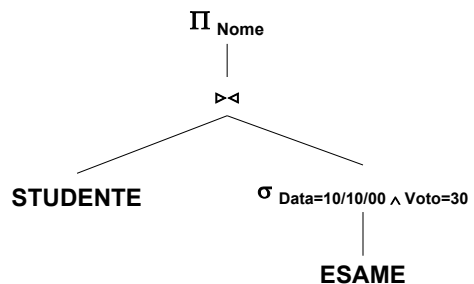
Ottimizzazione delle interrogazioni

Rappresentazione delle espressioni tramite alberi

- Ogni espressione dell'algebra relazionale può essere rappresentata in modo grafico da un albero, che esplicita l'ordine di valutazione degli operatori
- Ogni operatore corrisponde a un nodo
 - operatori unari con un ramo in ingresso e uno in uscita
 - operatori binari con due rami in ingresso e uno in uscita

33

Esempio di albero



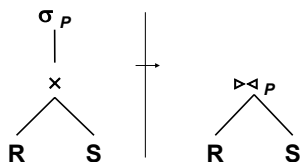
Corrisponde a:

$\Pi_{Nome} \text{ STUDENTE } \triangleright \triangleleft \sigma_{Data=10/10/00 \wedge Voto=30} \text{ ESAME}$

34

Trasformazioni di equivalenza per espressioni algebriche

1) Eliminazione dei prodotti cartesiani

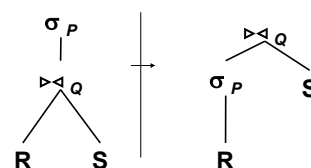


Vale se p è una congiunzione di predicati del tipo ATTR comp ATTR

35

Trasformazioni di equivalenza per espressioni algebriche

2) Push della selezione rispetto al join

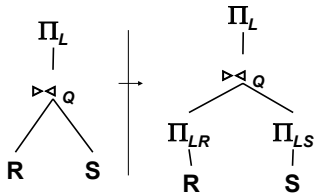


Vale se p è un predicato che si applica ai soli attributi di R

36

Trasformazioni di equivalenza per espressioni algebriche

3) Push della proiezione rispetto al join

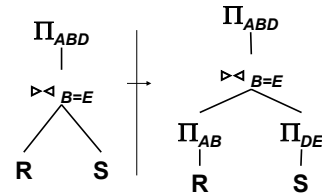


JR e JS sono gli attributi di R e S necessari a valutare Q
 $LR = L - \text{schema}(S) + JR$
 $LS = L - \text{schema}(R) + JS$

37

Esempio

3) Push della proiezione rispetto al join



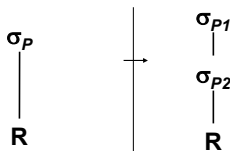
$R(A,B,C)$
 $S(D,E,F,G)$

$L=ABD$
 $JR=B$
 $JS=E$
 $LR=AB$
 $LS=DE$

38

Trasformazioni di equivalenza per espressioni algebriche

4) Idempotenza della selezione

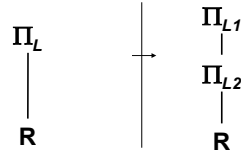


$P = P1 \wedge P2$

39

Trasformazioni di equivalenza per espressioni algebriche

5) Idempotenza della proiezione

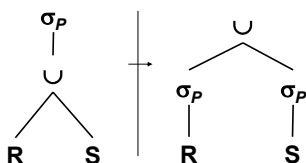


$L1 = L$
 $L2 \supseteq L$

40

Trasformazioni di equivalenza per espressioni algebriche

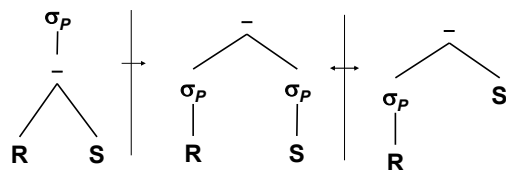
6) Push della selezione rispetto all'unione



41

Trasformazioni di equivalenza per espressioni algebriche

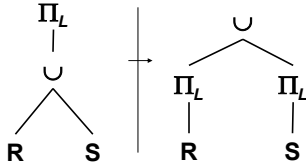
7) Push della selezione rispetto alla differenza



42

Trasformazioni di equivalenza per espressioni algebriche

8) Push della proiezione rispetto all'unione

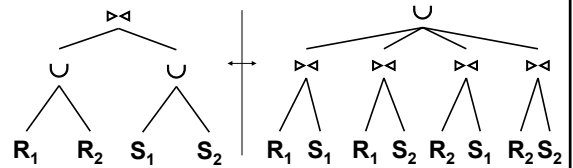


Attenzione: non vale il push della proiezione rispetto alla differenza e all'intersezione

43

Trasformazioni di equivalenza per espressioni algebriche

9) Commutazione di join e unione



44

Formule utili

$$R \bowtie R = R$$

$$R \cup R = R$$

$$R - R = \emptyset$$

$$R \bowtie \sigma_P R = \sigma_P R$$

$$R \cup \sigma_P R = R$$

$$R - \sigma_P R = \sigma_{\neg P} R$$

$$\sigma_{P1} R \bowtie \sigma_{P2} R = \sigma_{P1 \wedge P2} R$$

$$\sigma_{P1} R \cup \sigma_{P2} R = \sigma_{P1 \vee P2} R$$

$$\sigma_{P1} R - \sigma_{P2} R = \sigma_{P1 \wedge \neg P2} R$$

$$\sigma_P \emptyset = \emptyset$$

$$\Pi_L \emptyset = \emptyset$$

$$R \cup \emptyset = R$$

$$R - \emptyset = R$$

$$\emptyset - R = \emptyset$$

$$R \cap \emptyset = \emptyset$$

$$R \times \emptyset = \emptyset$$

$$R \bowtie \emptyset = \emptyset$$

45

Ottimizzazione algebrica

- Tra tutte le rappresentazioni equivalenti, conviene scegliere quella meno costosa da eseguire

- Criterio informale: minimizzare la dimensione dei risultati intermedi

- Tecnica:

- utilizzare dove possibile le trasformazioni di push (2, 3, 6, 7, 8);
- usare le trasformazioni di idempotenza (4, 5) per generare nuove selezioni e proiezioni

46

Esempio di ottimizzazione

Si ha lo schema:

R(A,B,C)

S(C,D,E)

T(D,F,G)

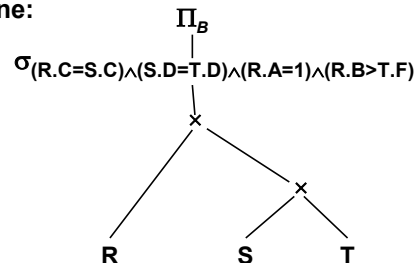
Si deve ottimizzare l'espressione algebrica:

$$\Pi_B \sigma_{(R.C=S.C) \wedge (S.D=T.D) \wedge (R.A=1) \wedge (R.B>T.F)} R \times S \times T$$

47

Esempio di ottimizzazione

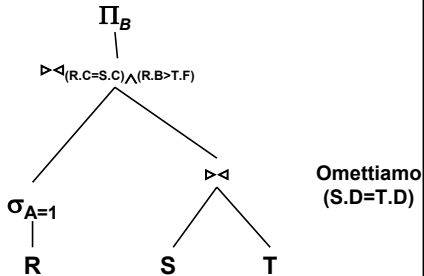
- Con la rappresentazione ad albero, si ottiene:



48

Esempio di ottimizzazione

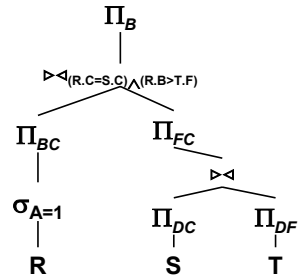
- Applicando le trasformazioni, si ottiene:



49

Esempio di ottimizzazione

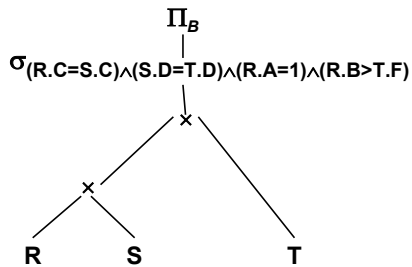
- Aggiungendo le proiezioni, si ottiene:



50

Esercizio

- Ripetere con:



51

Esercizi di ottimizzazione

- 1) Si ha lo schema: $R(A,B)$
 $S(A,B)$

Ottimizzare l'espressione algebrica:

$$\sigma_{(S.A=R.A) \wedge (R.A>2) \wedge (S.A=1)} R \times S$$

Il risultato è una relazione vuota, perchè il predicato è una CONTRADDIZIONE!!!!

52

Esercizi di ottimizzazione

- 2) Si ha lo schema: $R(A,B,C)$
 $S(C,D,E)$
 $T(C,D,E)$

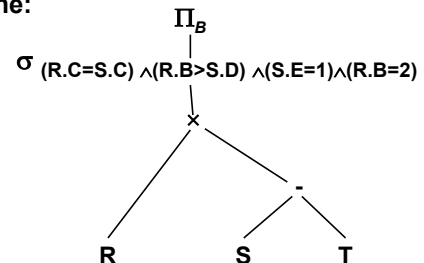
Ottimizzare l'espressione algebrica:

$$\Pi_{AD} \sigma_{(R.C=S.C) \wedge (S.E=1) \wedge (R.B=2) \wedge (R.B>S.D)} (R \times (S - T))$$

53

Esempio di ottimizzazione

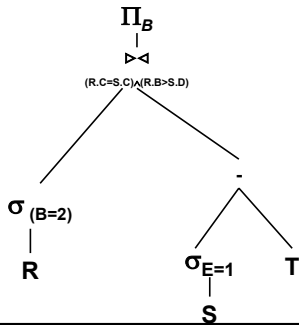
- Con la rappresentazione ad albero, si ottiene:



54

Esempio di ottimizzazione

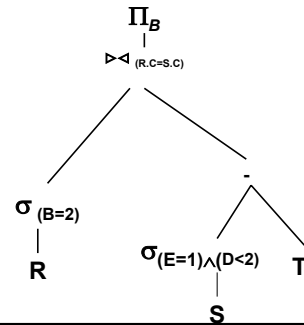
- Applicando le trasformazioni, si ottiene:



55

Esempio di ottimizzazione

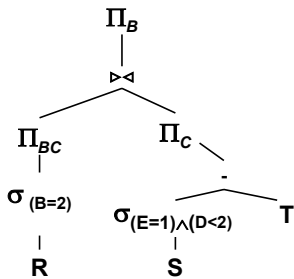
- Ragionando sui predicati, si ottiene:



56

Esempio di ottimizzazione

- Aggiungendo le proiezioni, si ottiene:



57