

Algebra Relazionale

corso di basi di dati e laboratorio

Prof. Alfio Ferrara

Anno Accademico 2017/2018

Indice

1	Operatori dell'algebra relazionale	1
1.1	Operatori	1
2	Operatori insiemistici	2
2.1	Operatori insiemistici	2
3	Operatori unari	4
3.1	Unari	4
4	Operatori binari	5
4.1	Binari	5
4.2	Join	7
4.3	Divisione	10
5	Equivalenza di espressioni algebriche	14
5.1	Equivalenza	14

1 Operatori dell'algebra relazionale

1.1 Operatori

Algebra relazionale

- Linguaggio procedurale per interrogazione di BD relazionali.

- Il risultato di un'interrogazione è una relazione che può essere manipolata utilizzando gli operatori dell'algebra stessa.

Operatori dell'algebra relazionale

- Gli operatori si dividono in: - **operatori insiemistici** - **operatori su relazioni**
- **Operatori fondamentali:** - Unari (Π, σ) - Binari ($\cup, \times, -$)
- **Operatori derivati:** - Binari ($\bowtie, \bowtie_{\theta}, \cap, \div$)

2 Operatori insiemistici

2.1 Operatori insiemistici

Operatori insiemistici

- Due relazioni si dicono **compatibili all'unione** se sono dello stesso grado e sono definite sugli stessi attributi.
- Date due relazioni r e s compatibili all'unione, è possibile operare su esse con i seguenti operatori insiemistici:

UNIONE: l'unione $r \cup s$ è costituita dalle ennuple t appartenenti a r o a s (o a entrambe).

$$r \cup s = \{t \mid t \in r \vee t \in s\}$$

DIFFERENZA: la differenza $r - s$ è costituita dalle ennuple t appartenenti a r e non a s .

$$r - s = \{t \mid t \in r \wedge \neg(t \in s)\}$$

Operatori insiemistici

INTERSEZIONE: l'intersezione $r \cap s$ è costituita dalle ennuple t sia a r sia a s .

$$r \cap s = \{t \mid t \in r \wedge t \in s\}$$

L'intersezione è un operatore derivato. Come è possibile esprimerlo per mezzo di differenza e unione?

$$r \cap s \equiv (r \cup s) - (r - s) - (s - r)$$

Esempio

Definiamo due relazioni $r(A, B, C)$ e $s(A, B, C)$

r			s		
A	B	C	A	B	C
a1	b1	c1	a1	b1	c1
a1	b2	c1	a2	b2	c2
a2	b1	c2			

Esempio

Operiamo su $r(A, B, C)$ e $s(A, B, C)$ per mezzo degli operatori insiemistici

$r \cup s$			$r - s$			$s - r$			$r \cap s$		
A	B	C	A	B	C	A	B	C	A	B	C
a1	b1	c1	a1	b2	c1	a2	b2	c2	a1	b1	c1
a1	b2	c1	a2	b1	c2						
a2	b1	c2									
a2	b2	c2									

Esempio

$$(r \cup s) - (r - s) - (s - r) = r \cap s$$

$r \cup s$			$r - s$			$s - r$			$r \cap s$		
A	B	C	A	B	C	A	B	C	A	B	C
a1	b1	c1	a1	b2	c1	a2	b2	c2	a1	b1	c1
a1	b2	c1	a2	b1	c2						
a2	b1	c2									
a2	b2	c2									

$$=$$

$r \cap s$		
A	B	C
a1	b1	c1

Compatibilità all'unione

Compatibilità all'unione

$R(A_1, A_2, \dots, A_n)$ e $S(B_1, B_2, \dots, B_n)$ sono **compatibili all'unione** se:

1. R e S hanno lo stesso grado

2. $Dom(A_i) = Dom(B_i)$ per ogni $i = 1 \dots n$

Di conseguenza, la relazione risultato di un'operazione insiemistica ha lo stesso schema della prima relazione R .

3 Operatori unari

3.1 Unari

Selezione (σ)

Selezione (σ)

Restituisce una relazione contenente l'insieme delle sole ennuple della relazione operando che soddisfano particolari condizioni, espresse mediante una **formula proposizionale** (uso costanti, nomi di attributo, operatori di confronto e connettivi logici AND, OR, NOT (\wedge , \vee , \neg)).

Esempio di selezione

r

A	B	C	D
a1	b1	c2	d1
a2	b2	c2	d2

Selezione:

$\sigma_{A=a2}(r)$

A	B	C	D
a2	b2	c2	d2

Proiezione (Π)

Proiezione (Π)

Restituisce come risultato una relazione contenente l'insieme delle ennuple della relazione operando limitate su particolari attributi.

La proiezione può essere considerata ortogonale alla selezione: decomposizione verticale VS decomposizione orizzontale.

Esempio di proiezione

r

A	B	C	D
a1	b1	c2	d1
a2	b2	c2	d2

Proiezione:

$\Pi_{A,B}(r)$

A	B
a1	b1
a2	b2

Esempio

r

A	B	C	D
a1	b1	c2	d1
a2	b2	c2	d2

Come posso ottenere una relazione con il solo valore a1?

$\Pi_A(\sigma_{A=a1}r)$

A
a1

4 Operatori binari

4.1 Binari

Prodotto cartesiano (\times)

Prodotto cartesiano (\times)

Date due relazioni $r = R(X)$ e $s = S(Y)$ con $X \cap Y = \emptyset$, il **prodotto cartesiano** $r \times s$ è una relazione definita su XY contenente l'insieme delle ennuple che sono combinazione delle ennuple di r e s .

Esempio

Definiamo due relazioni $r(A, B, C)$ e $s(D, E)$

r			s	
A	B	C	D	E
a1	b1	c1	d1	e1
a1	b2	c1	d1	e2
a2	b1	c2		

Esempio

$r \times s$				
A	B	C	D	E
a1	b1	c1	d1	e1
a1	b2	c1	d1	e1
a2	b1	c2	d1	e1
a1	b1	c1	d1	e2
a1	b2	c1	d1	e2
a2	b1	c2	d1	e2

Esempio

COUNTRY		NAME			
iso3	government	id	short	language	country
GBR	constitutional monarchy	12cf7	France	en-iso	FRA
FRA	republic	5896f	Italy	en-iso	ITA
ITA	republic				

Esempio

COUNTRY \times NAME					
government	iso3	country	id	short	language
constitutional monarchy	GBR	FRA	12cf7	France	en-iso
constitutional monarchy	GBR	ITA	5896f	Italy	en-iso
republic	FRA	FRA	12cf7	France	en-iso
republic	FRA	ITA	5896f	Italy	en-iso
republic	ITA	FRA	12cf7	France	en-iso
republic	ITA	ITA	5896f	Italy	en-iso

Considerazioni

- Poca utilità.
- Correla ennuple nel risultato anche se non sono correlate da un punto di vista semantico.
- In generale è necessario operare una selezione sul risultato del prodotto cartesiano per selezionare un sottoinsieme significativo di tuple.

4.2 Join

Theta-Join (\bowtie_{Θ})

Theta-Join (\bowtie_{Θ})

Date $r = R(X)$ e $s = S(Y)$, con X e Y disgiunti, il Theta-Join $r \bowtie_{\Theta} s$ è una relazione definita su XY composta dalle ennuple risultato della concatenazione di ennuple di r con ennuple di s che soddisfano le condizioni di confronto $A\Theta B$ (con $A \in X$ e $B \in Y$).

- Θ è un'espressione composta per mezzo di operatori di confronto ($=, \neq, >, <, \geq, \leq$).
- Se il confronto è di uguaglianza, l'operazione è detta **EQUIJOIN**.
- Si noti che il join $R \bowtie_{A=B} S$ è del tutto equivalente all'operazione $\sigma_{A=B}(R \times S)$

Esempio

COUNTRY \times NAME

government	iso3	country	id	short	language
constitutional monarchy	GBR	FRA	12cf7	France	en-iso
constitutional monarchy	GBR	ITA	5896f	Italy	en-iso
republic	FRA	FRA	12cf7	France	en-iso
republic	FRA	ITA	5896f	Italy	en-iso
republic	ITA	FRA	12cf7	France	en-iso
republic	ITA	ITA	5896f	Italy	en-iso

COUNTRY $\bowtie_{iso3=country}$ NAME

government	iso3	country	id	short	language
republic	FRA	FRA	12cf7	France	en-iso
republic	ITA	ITA	5896f	Italy	en-iso

Si noti che il join $COUNTRY \bowtie_{iso3=country} NAME$ è del tutto equivalente all'operazione $\sigma_{iso3=country}(COUNTRY \times NAME)$

Join naturale (\bowtie)**Join naturale (\bowtie)**

Date due relazioni $r(YX)$ e $s(XZ)$, il **join naturale** di r e s è una relazione definita su YXZ composta dalle ennuple risultato della concatenazione di ennuple di r con ennuple di s con valori identici sugli attributi X .

- Due ennuple sono dette *joinabili* se contribuiscono al join.
- Una ennupla che non contribuisce al join è detta *dangling*.
- Un join è completo se nessuna delle relazioni contiene ennuple dangling.
- Due attributi si possono rendere compatibili al join naturale attraverso una ridenominazione (ρ).

Esempio

r			s		
A	B	C	C	D	E
a1	b1	c1	c1	d1	e1
a2	b2	c2	c2	d2	e2
a3	b3	c3	c1	d3	e3

$r \bowtie s$				
A	B	C	D	E
a1	b1	c1	d1	e1
a1	b1	c1	d3	e3
a2	b2	c2	d2	e2

Proprietà del join naturale

- Il join naturale è **commutativo** e **associativo**.
- Può essere espresso in termini di: - Prodotto cartesiano. - Selezione ennuple con valori uguali per attributi in comune. - Proiezione su attributi di r e s eliminando duplicazioni.
- Se $r(X), s(Y)$ con $X \cap Y = \emptyset$,

$$r \bowtie s = r \times s$$

- Se $r(X), s(Y)$ con $X = Y$,

$$r \bowtie s = r \cap s$$

Esempio

WEB			WEBSITE	
url	country	official	url	title
governo.it	ITA	False	governo.it	Governo Italiano
igmi.org	ITA	False	igmi.org	NULL
stat.it	ITA	False	stat.it	Statistics
poste.it	ITA	False	poste.it	Poste Italiane

WEB ⋈ WEBSITE			
url	country	official	title
governo.it	ITA	False	Governo Italiano
igmi.org	ITA	False	NULL
stat.it	ITA	False	Statistics
poste.it	ITA	False	Poste Italiane

Outer join

- Gli operatori THETA JOIN e JOIN NATURALE combinano ennuple di relazioni che hanno valori 'in corrispondenza' su particolari attributi: ne consegue che non tutte le ennuple entrano a far parte del risultato.
- L'operazione di **outer join** appende alle ennuple risultato del join anche quelle ennuple di una o entrambe le relazioni che non soddisfano i criteri di join, estendendole con opportuni 'NULL'.
- Abbiamo: - LEFT OUTER JOIN: considera la sola relazione di sinistra. - RIGHT OUTER JOIN: considera la sola relazione di destra. - FULL OUTER JOIN: considera entrambe le relazioni.

Esempio

WEB			WEBSITE	
url	country	official	url	title
governo.it	ITA	False	governo.it	Governo Italiano
igmi.org	ITA	False	igmi.org	NULL
stat.it	ITA	False	stat.it	Statistics
poste.it	ITA	False	poste.it	Poste Italiane
			google.com	Google

WEB RIGHT OUTER JOIN WEBSITE

url	country	official	title
governo.it	ITA	False	Governo Italiano
igmi.org	ITA	False	NULL
stat.it	ITA	False	Statistics
poste.it	ITA	False	Poste Italiane
google.com	NULL	NULL	Google

4.3 Divisione

Divisione

Divisione

La divisione si applica **solo** a relazioni $r(Z)$, $s(X)$ in cui $X \subseteq Z$. Dato l'insieme di attributi $Y = Z - X$, la divisione $r \div s$ è una relazione $T(Y)$ contenente tutte le tuple t tali che: - vi siano tuple t_r in r con $t_r[Y] = t$, e - $t_r[X] = t_s$ per ogni tupla t_s di s .

Intuitivamente, la divisione verifica una condizione o una corrispondenza fra una o più tuple di r e **tutte** le tuple di s .

Esempio

r							
A	B			s			
x	a			B		$r \div s$	
y	a			a		A	
x	b	\div		b	=	x	
z	b						

Come ottenere la divisione da altre operazioni (I)

Proietto gli attributi non in comune alle due relazioni

r							
A	B			s		$\Pi_A(r)$	
x	a			B		A	
y	a			a		x	
x	b	\div		b	\rightarrow	y	
z	b					z	

Come ottenere la divisione da altre operazioni (II)

Definisco tutte le possibili combinazioni fra gli elementi in A e in s

$\Pi_A(r)$		s		$\Pi_A(r) \times s$	
A		B		A	B
x	\times	a	\rightarrow	x	a
y		b		x	b
z				y	a
				y	b
				z	a
				z	b

Come ottenere la divisione da altre operazioni (III)

Tolgo alle possibili combinazioni le combinazioni effettive in r

$\Pi_A(r) \times s$					
A	B		r		$(\Pi_A(r) \times s) - r$
x	a	-	A	B	
x	b		x	a	
y	a		y	a	
y	b		x	b	
z	a		z	b	
z	b				

A	B
y	b
z	b

$\Pi_A((\Pi_A(r) \times s) - r)$
A
y
z

Come ottenere la divisione da altre operazioni (IV)

Otengo la divisione

$\Pi_A(r)$		$\Pi_A((\Pi_A(r) \times s) - r)$		$(\Pi_A(r)) - ((\Pi_A(r) \times s) - r)$
A		A		A
x	$-$	y	\rightarrow	x
y		z		
z				

$r \div s$

A
x

Esempio

Trovare le misure che sono disponibili per tutte le nazioni.

MEASURE

id	value	label	country
952	66028467.0	Population	FRA
957	2.7390000001e+13	GDP	FRA
1345	59831093.0	Population	ITA
1350	2.0680000001e+13	GDP	ITA
2848	64097085.0	Population	GBR

COUNTRY

iso3	government
GBR	constitutional monarchy
FRA	republic
ITA	republic

$$(\rho_{country \rightarrow iso3}(\Pi_{label, country}(MEASURE)) \div \Pi_{iso3}(COUNTRY))$$

label
Population

Esempio

$$\rho_{country \rightarrow iso3}(\Pi_{label, country}(MEASURE))$$

label	iso3
Population	FRA
GDP	FRA
Population	ITA
GDP	ITA
Population	GBR

$$\Pi_{iso3}(COUNTRY)$$

iso3
GBR
FRA
ITA

Esempio

Per leggibilità, rinomino:

$$\rho_{country \rightarrow iso3}(\Pi_{label, country}(MEASURE)) \rightarrow R$$

e

$$\Pi_{iso3}(COUNTRY) \rightarrow S$$

Definisco tutte le possibili combinazioni fra gli elementi in $label$ e in S

$\Pi_{label}(R)$		S		$\Pi_{label}(R) \times S$	
label		iso3		label	iso3
Population		GBR		Population	GBR
GDP		FRA		Population	FRA
		ITA		Population	ITA
				GDP	GBR
				GDP	FRA
				GDP	ITA

Esempio

$\Pi_{label}(R) \times S$		R		$(\Pi_{label}(R) \times S) - R$	
label	iso3	label	iso3	label	iso3
Population	GBR	Population	GBR	GDP	GBR
Population	FRA	Population	FRA		
Population	ITA	Population	ITA		
GDP	GBR	GDP	FRA		
GDP	FRA	GDP	ITA		
GDP	ITA				

Esempio

$\Pi_{label}(R)$		$\Pi_{label}((\Pi_{label}(R) \times S) - R)$	
label		label	
Population		GDP	
GDP			

$(\Pi_{label}(R)) - (\Pi_{label}((\Pi_{label}(R) \times S) - R))$	
label	
Population	

5 Equivalenza di espressioni algebriche

5.1 Equivalenza

Caratteristiche generali

- L'algebra permette di formulare espressioni fra loro equivalenti (che producono lo stesso risultato).
- Le interrogazioni SQL sono tradotte in algebra e si valuta il costo in termini di dimensioni dei risultati intermedi; la scelta fra varie alternative equivalenti è per quella a costo minore.
- In questo contesto vengono utilizzate delle trasformazioni di equivalenza, cioè operazioni che sostituiscono un'espressione con un'altra equivalente.
- Sono di interesse le trasformazioni che riducono le dimensioni dei risultati intermedi e quelle che preparatorie a ciò.

Regole di trasformazione

- Atomizzazione delle selezioni:

$$\sigma_{C1 \wedge C2}(R) = \sigma_{C1}(\sigma_{C2}(R)) = \sigma_{C2}(\sigma_{C1}(R))$$

Una selezione congiuntiva può essere sostituita da una serie di selezioni atomiche (preliminare ad altre).

- Commutatività di σ :

$$\sigma_{C1}(\sigma_{C2}(R)) = \sigma_{C2}(\sigma_{C1}(R))$$

- Idempotenza delle proiezioni:

$$\Pi_X(R) = \Pi_X(\Pi_{XY}(R))$$

cascata di proiezioni che eliminano i vari attributi in fasi diverse (preliminare ad altre).

Regole di trasformazione

- Anticipazione della selezione rispetto al join:

$$\sigma_C(R \bowtie S) = (\sigma_C(R)) \bowtie S$$

la condizione C coinvolge solo attributi di R .

- Conversione di una sequenza (σ, \times) in un join:

$$\sigma_C(R \times S) = R \bowtie_C S$$

se la condizione C corrisponde a una condizione di join.

Regole di trasformazione

Anticipazione della proiezione rispetto al join: R e S definite su X_1 e X_2 rispettivamente; J_1 e J_2 denotano rispettivamente gli attributi di R e S coinvolti nella condizione di join C ; $Y_1 = (X_1 \cap Y)$ e $Y_2 = (X_2 \cap Y)$. Se C coinvolge solo attributi inclusi in Y ($J_1 \cup J_2 \subseteq Y$):

$$\Pi_Y(R \bowtie_C S) = (\Pi_{Y_1}(R)) \bowtie_C (\Pi_{Y_2}(S))$$

Se C coinvolge ulteriori attributi non inclusi in Y , allora:

$$\Pi_Y(R \bowtie_C S) = \Pi_Y((\Pi_{Y_1 J_1}(R)) \bowtie_C (\Pi_{Y_2 J_2}(S)))$$

In sintesi, si possono eliminare subito gli attributi di ciascuna relazione che non sono coinvolti nel join e non sono utilizzati nella lista degli attributi della proiezione.