

# Algebra relazionale e gestione dei valori NULL

Lezioni tenute dalla Prof. Raffalella Gentilini  
Appunti di Robert Parcus      [betoparcus@gmail.com](mailto:betoparcus@gmail.com)

10 ottobre 2012

## 0.1 La divisione

Indicato più intuitivamente controparte della quantificazione universale.

Esempio: Date due relazioni

ISCRIZIONE:	matricola	id_corso	CORSO:	id_corso
	123	BD		BD
	283	BD		INF
	123	INF		INF
	123	MAT		MAT
	283	MAT		
	375	BD		
	283	INF		

$$R \leftarrow ISCRIZIONE \div CORSO$$

(mi darà le matricole iscritte a tutti i corsi appartenenti a CORSO)

R: 

123	283
-----	-----

Vediamo come si può derivare la divisione dalle altre operazioni che conosciamo:

Definizione [DIVISIONE] : Siano  $R1(x), R2(y)$  due schemi di relazione t.c.  $Y$  appartiene a  $X$  e siano  $r1, r2$  due istanze di  $R1R2$ . L'operatore divisione produce una relazione le cui tuple, se estese con una qualunque tupla della seconda relazione producono una tupla di  $R1$ .

$$r1 \div r2 = \{ t \mid \forall t' \in r2, t \cup t' \in r1 \}$$

La divisione  $r1 \div r2$  è dunque un operatore definito sullo schema con attributi  $X - Y$ . È un operatore derivato e può essere definito come:

$$T1 \leftarrow \Pi_{x-y}(R1)$$

$$T2 \leftarrow \Pi_{x-y}(R2 \times T1) - R1$$

Sull'esempio di prima:

$$T1 \leftarrow \Pi_{matricola}(Iscrizione)$$

$$T2 \leftarrow \Pi_{matricola}((T1 \times corso) - R1)$$

$$R \leftarrow T1T2$$

## 0.2 ALGEBRA RELAZIONE E VALORI NULL

La presenza di valori nulli nelle istanze di una base di dati richiede un'estensione della semantica degli operatori dell'algebra relazionale. L'approccio tradizionale (usato anche nei DBMS commerciali ed in SQL) è quello di considerare una logica a 3 valori per la valutazione delle formule proposizionali e quei nodi degli

operatori di Selezione e Join.

### Proiezione, unione, differenza con valori nulli:

Continuano a comportarsi usualmente. L'uguaglianza tra NULL é un livello sintattico e due tuple sono uguali anche se ci sono valori NULL.

Esempio:

IMPIEGATI:	codice	nome	ufficio	Responsabili:	codice	nome	ufficio
	123	Rossi	A12		123	Rossi	A12
	231	Verdi	NULL		NULL	NULL	NULL
	373	Verdi	A27		435	Verdi	A27
	435	Verdi	NULL				

manca roba...

**Selezione e valori nulli:**

Per la selezione, il problema é stabilire se in presenza di valori NULL un predicato é vero o meno.

IMPIEGATI:	codice	nome	ufficio	risultato
	123	Rossi	A12	OK
	231	Verdi	NULL	???
	373	Verdi	A27	OK

$\Pi_{ufficio='A12'}(IMPIEGATI)$

La seconda tupla fa parte del risultato?

*Non si può sapere.*

Vanno definite politiche di gestione del NULL:

Per verificare predicati come quelli dell'esempio precedente, si introduce una logica a 3 valori, dove oltre al valore V ed F si ha un terzo valore ''.

a=a	V
a=b	F
a=NULL	?
NULL=NULL	?
a!=a	F
a!=b	V
a!=null	?
null!=null	?

Queste tabelle di verità possono essere estese ad altri tipi di confronto,. Se  $D$  é un dominio su cui é definito un ordinamento  $< (>, >=, <=)$ , allora per  $x, y \in D$   $x < y$  é appartenente  $<, > =$  il confronto  $x$  e  $y$  é indefinito se  $x$  o  $y$  sono NULL. Le

tabelle di verità nella logica a 3 valori per gli operatori booleani sono NOT:

V	F		V	V F ?		V	V V V
F	V	AND:	F	F F F	OR:	F	V F ?
?	?		?	? F ?		?	V ? ?

Una selezione  $G(f)$  produce le sole tuple per cui la condizione di selezione risulti true.

Esempio:  $\sigma_{ufficio='A12' \text{ OR } ufficio \neq 'A12'}(IMPIEGATI)$  Non restituisce tutte le tuple!!! Per lavorare esplicitamente con i valori nulli si introducono le condizioni atomiche - a IS NULL - a IS NOT NULL Selezione in ufficio= 'A12' || ufficio != 'A12' || ufficio IS NULL (IMPIEGATI)

JOIN con valori NULL:

Valgono le stesse regole di valutazione per F utilizzate nella selezione. Il JOIN NATURALE non combina due tuple se queste hanno entrambe valori nulli su un attributo in comune.

Esempio: Impiegati: codice 123 231 272 435 nome Rossi Verdi " " ufficio A12 NULL A27 NULL Responsabili: ufficio A12 A27 NULL codice 123 NULL 231

### 0.3 esercizi 18/10

Consideriamo la seguente porzione di BD:

FORNITORE(CodForn, nome, città)  
 PRODOTTI(CodProd, nome, marca, modello)  
 CATALOGO(CodForn,CodProd, Costo)

2) Trovare i nomi dei fornitori che distribuiscono prodotti di marca IBM:

$$T_1 \leftarrow \sigma_{marca=IBM} (PRODOTTI \bowtie CATALOGO)$$

3) Trovare i codici di tutti i prodotti che sono forniti da almeno 2 fornitori:  
Idea: Concateniamo due copie di catalogo su valori uguali dell'attributo codice prodotto e troviamo le tuple su cui i fornitori sono diversi.

$$\begin{aligned} COPIACATALOGO(codforn1, codforn2) &\leftarrow \Pi_{codforn, prod}(CATALOGO) \\ T_1 &\leftarrow COPIACATALOGO \bowtie CATALOGO \\ RIS &\leftarrow \Pi(\sigma_{codforn1, codforn2}(T_1)) \end{aligned}$$

4) Nomi dei fornitori che distribuiscono tutti i prodotti:

$$\begin{aligned} T1 &\leftarrow (\Pi_{codFor, codProd}(CATALOGO)) \div (\Pi_{codProd}(PRODOTTI)) \\ RIS &\leftarrow \Pi(T_1 \bowtie FORNITORE) \end{aligned}$$

Oppure, senza usare la divisione:

// Tutte le coppie fornitore e prodotto

$$\begin{aligned}
T_1 &\leftarrow \Pi_{\text{codForn}}(\text{CATALOGO}) \times \P_{\text{codProd}}(\text{PRODOTTO}) \\
&\quad // \text{CodForn} \notin T_2 \text{secodForn forniscetuttiiprodotti.} \\
T_2 &\leftarrow \Pi(T_1 - \Pi_{\text{codForn}, \text{codProd}}(\text{CATALOGO})) \\
RIS &\leftarrow \Pi_{\text{codForn}}(\text{CATALOGO}) - T_2
\end{aligned}$$

STADIO(nome, città, capienza) PARTITA(stadio, dato, ora, squadra1, squadra2) SQUADRA(nazione, continente, livello)

1) Determinare gli stadi in cui non gioca alcuna squadra africana:

$$\begin{aligned}
AFRICA &\leftarrow \Pi_{(\text{nazione})}(\sigma_{(\text{continente} \neq \text{africa})}(\text{SQUADRA})) \quad BAD \leftarrow \\
&\Pi_{(\text{stadio})}(\text{PARTITA} \bowtie_{(\text{nazione} = \text{squadra1} \text{ AND } \text{nazione} = \text{squadra2})} AFRICA) \\
RIS &\leftarrow \Pi_{(\text{nome})}(\text{STADIO}) - BAD
\end{aligned}$$

2) Determinare le squadre che incontrano soltanto squadre dello stesso livello:

$$\begin{aligned}
&\text{SQUADRA1\_LIVELLO}(S1, S2, L1) \\
&\leftarrow \Pi_{(\text{squadra1}, \text{squadra2}, \text{livello})}(\text{PARTITA} \bowtie_{(\text{squadra1} = \text{nazione})} \text{SQUADRA} \\
&\quad \text{SQUADRA2\_LIVELLO}(S1, S2, L1, l2) \\
&\leftarrow \Pi_{(S1, S2, L1, \text{livello})}(\text{SQUADRA1\_LIVELLO} \bowtie_{(S2 = \text{nazione})} \text{SQUADRA} \\
&\quad \text{BAD1} \leftarrow \Pi_{(S1)}(\sigma_{L1 \neq L2}(\text{SQUADRE\_LIVELLO})) \\
&\quad \text{BAD2} \leftarrow \Pi_{(S2)}(\sigma_{L1 \neq L2}(\text{SQUADRE\_LIVELLO})) \\
&\quad RIS \leftarrow \Pi_{(\text{nazione})}(\text{squadra}) - (\text{BAD1} \cup \text{BAD2})
\end{aligned}$$

Sia dato il seguente schema relazione che descrive gli esami (obbligatori e non) in un anno ( I, II, III) di un indirizzo di laurea triennale:

CORSO(cod\_c, nome, CFU, SSD)  
INDIRIZZO(COD\_1, titolo)  
COMPOSIZIONE(COD\_I, COD\_C, tipo, anno)

1) Determinare i titoli dei corsi che in almeno un indirizzo possono essere collocati indifferentemente in ciascuno dei tre anni:

$$\begin{aligned}
ANNO1(\text{codc1}, \text{codI1}) &\leftarrow \Pi_{(\text{codc}, \text{codI})} \sigma_{(\text{anno}=1)}(\text{COMPOSIZIONE}) \\
ANNO2(\text{codc2}, \text{codI2}) &\leftarrow \Pi_{(\text{codc}, \text{codI})} \sigma_{(\text{anno}=2)}(\text{COMPOSIZIONE}) \\
ANNO3(\text{codc3}, \text{codI3}) &\leftarrow \Pi_{(\text{codc}, \text{codI})} \sigma_{(\text{anno}=3)}(\text{COMPOSIZIONE}) \\
GOOD &\leftarrow (ANNO1 \bowtie_{(\text{codc1} = \text{codc2} \text{ AND } \text{codI1} = \text{codI2})} \\
&\quad ANNO2) \bowtie_{(\text{codc2} = \text{codc3} \text{ AND } \text{codI2} = \text{codI3})} ANNO3) \\
RISULTATO &\leftarrow \Pi_{(TITOLO)}(GOOD \bowtie_{(\text{codc1} = \text{codc})} CORSO)
\end{aligned}$$

2) Determinare i titoli degli indirizzi che prevedono lo stesso insieme di esami obbligatori dell'indirizzo 'Sistemi': //

// Corsi obbligatori dell'indirizzo sistemi

$$\begin{aligned}
& OBBSIST \leftarrow \Pi_{(codc)}(\sigma_{(TITOLO=sistemi \wedge Dtipo=obb)}(INDIRIZZO \bowtie \\
& \quad COMPOSIZIONE)) \\
& // \text{Tutti i corsi obbligatori } OBB \leftarrow \Pi_{(codI, codc)}(\sigma_{(tipo=obb)}(COMPOSIZIONE)) \\
& // \text{indirizzi che non voglio nel risultato } \Pi_{incodi}((\pi_{incodi}(indirizzi) \times OBBSIST) - \\
& OBB) \cup \Pi_{codI}(OBB - (\Pi_{codI}(INDIRIZZI) \times OBBSIST))
\end{aligned}$$