



PEGASO
Università Telematica



Indice

1. INTRODUZIONE.....	3
2. JOIN	4
3. JOIN NATURALE	6
4. JOIN COMPLETO	10
5. JOIN INCOMPLETO	11
6. JOIN VUOTO	12
BIBLIOGRAFIA	13

1. Introduzione

In questa unità didattica viene introdotto l'operatore dell'algebra relazionale più importante: l'operatore di JOIN. Vengono illustrati il Join naturale, i Join completi e Join incompleti, accompagnati anche da esempi.

L'operatore di Join consente di relazionare informazioni distribuite su più tabelle attraverso semplici considerazioni soprattutto di tipo insiemistico.

2. Join

Partiamo dai due operatori dell'algebra relazionale illustrati precedentemente, ovvero l'operatore Selezione e l'operatore di Proiezione. Tali operatori, opportunamente combinati insieme, permettono di estrarre informazioni da **una** sola relazione. Non si possono però correlare informazioni presenti in **relazioni diverse**, né informazioni in ennuple diverse di una stessa relazione. Per tale esigenze, ci viene in aiuto il Join, l'operatore più interessante dell'algebra relazionale attraverso il quale possiamo correlare dati in relazioni diverse.

Per comprendere tale operatore, partiamo dal seguente esempio che consiste in un concorso pubblico dove, al fine di garantire la correttezza della prova, la commissione decide di implementare la seguente procedura:

Prove scritte in un concorso pubblico

- I compiti sono anonimi e ad ognuno è associata una busta chiusa con il nome del candidato.
- Ciascun compito e la relativa busta vengono contrassegnati con uno stesso numero.

Come si può notare osservando la Figura 1, si vengono a definire le due tabelle della parte superiore della figura. La tabella sottostante raffigura l'abbinamento voto-persona, necessario alla fine per valutare la graduatoria del concorso.

1	25	1	Mario Rossi
2	13	2	Nicola Russo
3	27	3	Mario Bianchi
4	28	4	Remo Neri

Mario Rossi	25
Nicola Russo	13
Mario Bianchi	27
Remo Neri	28

Figura 1: la suddivisione delle informazioni relative ad un concorso pubblico: in alto le due correlazioni che garantiscono la correttezza della procedura. In basso l'abbinamento necessario ai fini della valutazione dei candidati.

Ci poniamo la seguente domanda: come effettuare la correlazione tra le due tabelle ID-Voto e ID-Persona al fine di ottenere come relazione risultato la tabella VOTO-CANDIDATO?

Un join è un'operazione SQL eseguita per stabilire una connessione tra due o più tabelle del database in base alle colonne corrispondenti, creando così una relazione tra le tabelle. Le query più complesse in un sistema di gestione di database SQL implicano comandi di join. Esistono diversi tipi di join.

Il tipo di join utilizzato da un programmatore determina quali record vengono selezionati dalla query.

Nel prossimo paragrafo vedremo come risolvere tale correlazione di dati.

3. Join Naturale

Dato l'esempio precedente, portato in termini di database relazionale, abbiamo le relazioni, con le opportune istanze illustrate in Figura 2. La relazione $R(\text{Numero}, \text{Candidato}, \text{Voto})$ rappresenta la relazione risultante. Come si può notare, a partire dalle due tabelle $R_1(\text{Numero}, \text{Voto})$ e $R_2(\text{Numero}, \text{Candidato})$, si è riusciti in qualche modo a formare una terza tabella fondamentale per il nostro esempio.

Numero	Voto	Numero	Candidato
1	25	1	Mario Rossi
2	13	2	Nicola Russo
3	27	3	Mario Bianchi
4	28	4	Remo Neri

Numero	Candidato	Voto
1	Mario Rossi	25
2	Nicola Russo	13
3	Mario Bianchi	27
4	Remo Neri	28

Figura 2: le informazioni relative al concorso messe in un database relazionale. Due tabelle in input ed una tabella in output.

Pertanto possiamo introdurre l'operatore di Join naturale come quell'operatore binario dell'algebra relazionale il quale, avendo in input due relazioni (generalizzabile):

- **Produce** un risultato:
 - sull'unione degli attributi degli operandi;
 - con ennuple costruite ciascuna a partire da una ennupla di ognuno degli operandi.
- **Combina** le ennuple di due relazioni sulla base dell'uguaglianza dei valori degli attributi comuni alle due relazioni.

In termini formali, date due relazioni $R_1(X_1)$, $R_2(X_2)$, definiamo l'operatore di **Join naturale** nel modo seguente:

- $R_1 \bowtie R_2$ è una relazione su $X_1 X_2$ (Unione di X_1 con X_2):

$$R_1 \bowtie R_2 = \{ \forall t \in X_1 X_2 \mid \exists t_1 \in R_1 \text{ and } t_2 \in R_2 \text{ con } t[X_1] = t_1 \text{ e } t[X_2] = t_2 \}$$

Tornando all'esempio precedente, possiamo comprendere meglio tale operatore, facendo ricorso alla rappresentazione insiemistica dell'operatore. In Figura 3 abbiamo l'insieme Numero, dominio dell'attributo Numero, dove ciascun elemento è collegato sia all'insieme dell'attributo Voto che all'insieme dell'attributo Candidato. La relazione viene costruita attraverso l'attributo comune Numero ed è rappresentata in Figura 4. L'attributo Numero funge anzi da chiave primaria della nuova relazione. Come si vedrà in seguito, questa situazione di coinvolgimento di chiavi primarie nel Join naturale accade molto spesso.

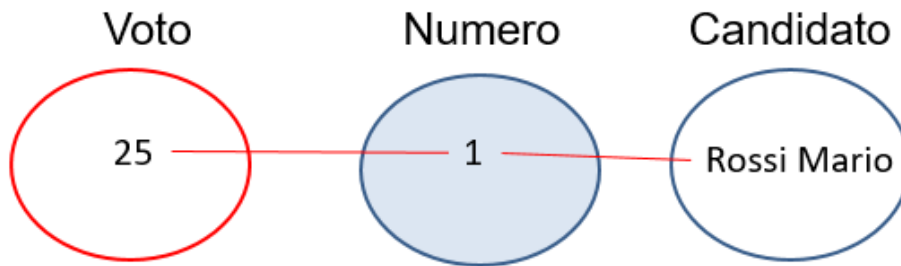


Figura 3: aspetto insiemistico del Join naturale.

Date le due relazioni dell'esempio precedente, $r_1(\underline{\text{Numero}}, \text{Voto})$ e $r_2(\underline{\text{Numero}}, \text{Candidato})$ ed applicando la definizione di Join naturale si ottiene la relazione risultante:

- $r_1 \bowtie r_2 = r(\underline{\text{Numero}}, \text{Voto}, \text{Candidato})$

Numero	Candidato	Voto
1	Mario Rossi	25
2	Nicola Russo	13
3	Mario Bianchi	27
4	Remo Neri	28

Figura 4: relazione risultato del Join naturale.

Riassumendo:

- Le tuple del risultato sono ottenute combinando tuple degli operandi con valori uguali sugli attributi comuni;
- Ogni tupla che compare nel risultato del join naturale di r_1 e r_2 , istanze rispettivamente di $R_1(X_1)$ e $R_2(X_2)$, è ottenuta come combinazione di una tupla di r_1 con una tupla di r_2 sulla base dell'uguaglianza dei valori degli attributi comuni (cioè quelli in $X_1 \cap X_2$).

- Inoltre, lo schema del risultato è l'unione degli schemi degli operandi.

Esempi

In questo paragrafo vediamo brevemente due esempi significativi di applicazione del Join naturale.

Il primo esempio riguarda l'applicazione del Join naturale tra le due relazioni illustrate in Figura 5 al fine di reperire, per ciascun corso, gli studenti che lo hanno superato e con che voto.

Esami				Corsi			
Matricola	CodCorso	Voto	Lode	CodCorso	Titolo	Docente	Anno
29323	483	28	NO	483	Analisi	Biondi	1
39654	729	30	Sì	729	Analisi	Neri	1
29323	913	26	NO	913	Sistemi Informativi	Castani	2
35467	913	30	NO				

Figura 5: le due relazioni operande dell'esempio proposto.

A partire quindi dalle suddette tabelle, vediamo che hanno in comune l'attributo CodCorso, il quale è anche chiave primaria della relazione Corsi, evidenziati con cornice rossa nelle tabelle. Applicando la definizione di Join naturale, ricaviamo la relazione illustrata in Figura 6 dove con cornice rossa è stato evidenziato l'attributo intersezione.

Esami				Corsi			
Matricola	CodCorso	Voto	Lode	Titolo	Docente	Anno	
29323	483	28	NO	Analisi	Biondi	1	
39654	729	30	Sì	Analisi	Neri	1	
29323	913	26	NO	Sistemi Informativi	Castani	2	
35467	913	30	NO	Sistemi Informativi	Castani	2	

Figura 6: relazione risultato dell'esempio.

Il secondo esempio riguarda la relazione *Infrazioni*, contenente le infrazioni al codice della strada commesse da singole autovetture, raffigurata in Figura 7 e la relazione *Autovetture*, rappresentata in Figura 8.

<u>Codice</u>	<u>Data</u>	<u>Agente</u>	<u>Articolo</u>	<u>Stato</u>	<u>Targa</u>
143256	25/10/2017	567	44	I	AB 234 ZK
987554	26/10/2017	456	34	I	AB 234 ZK
987557	26/10/2017	456	34	I	CB 123 AA
630876	15/10/2017	456	53	F	CB 123 AA
539856	12/10/2017	567	44	F	CB 123 AA

Figura 7: la relazione Infrazioni.

<u>Stato</u>	<u>Targa</u>	<u>Proprietario</u>	<u>Indirizzo</u>
I	C B 123 AA	Verdi Piero	Via Tigli....
I	DE 834 ZZ	Verdi Piero	Via Tigli...
I	AB 234 ZK	Bini Luca	Via Aceri...
F	CB 123 AA	Beau Marcel	Rue Louis XIV...

Figura 8: la relazione Autovetture.

Applicando la definizione di Join naturale, possiamo generare una relazione la quale contiene l'unione degli attributi con i due attributi *Stato* e *Targa* in comune. Si noti che anche in questo caso, il legame avviene attraverso una chiave primaria (*Stato*+*Targa*), evidenziati con cornice rossa. Pertanto il risultato finale è illustrato in

Figura 9.

<u>Codice</u>	<u>Data</u>	<u>Ag</u>	<u>Art</u>	<u>Stato</u>	<u>Targa</u>	<u>Propr.</u>	<u>Ind.</u>
143256	25/10/2017	567	44	I	AB 234 ZK	Bini Luca	Via Aceri...
987554	26/10/2017	456	34	I	AB 234 ZK	Bini Luca	Via Aceri...
987557	26/10/2017	456	34	I	CB 123 AA	Verdi Piero	Via Tigli...
630876	15/10/2017	456	53	F	CB 123 AA	Beau Marcel	Rue Luis...
539856	12/10/2017	567	44	F	CB 123 AA	Beau Marcel	Rue Luis...

Figura 9: risultato del Join naturale tra la relazione Infrazioni e la relazione Autovetture.

4. Join completo

Si supponga di avere, per una industria, le relazioni illustrate in Figura 10. Come si può notare, si hanno due relazioni, Impiegato e Reparto, legate tra loro attraverso il codice del reparto. Se voglio quindi conoscere per ogni impiegato il suo capo reparto, applico il Join alle due relazioni, ottenendo quindi la relazione di Figura 11.

Impiegato	Reparto	Reparto	Capo
Rossi	A	A	Mori
Neri	B	B	Bruni
Bianchi	B		

Figura 10: uno schema di database per una industria.

Impiegato	Reparto	Capo
Rossi	A	Mori
Neri	B	Bruni
Bianchi	B	Bruni

Figura 11: il risultato del Join tra Impiegato e Reparto.

Alla luce di questo esempio, un Join si definisce **Join completo** quando tutte le ennuple delle relazioni operande contribuiscono, come nel caso dell'esempio, alla relazione risultato.

Ne consegue che, date due relazioni con n ed m ennuple rispettivamente, il risultato finale potrà contenere al massimo **$n \times m$** ennuple.

5. Join incompleto

Si parta sempre dall'esempio precedente, ovvero dallo schema di database composto dalle due relazioni Impiegato e Reparto. Si supponga di avere la situazione illustrata in Figura 12, dove il reparto A non esiste nella tabella reparti mentre nessun impiegato appartiene al reparto C.

Impiegato	Reparto	Reparto	Capo
Rossi	A	B	Mori
Neri	B	C	Bruni
Bianchi	B		

Figura 12: caso di Join incompleto.

In questa situazione, non tutte le ennuple contribuiscono al risultato, come si può notare dalla Figura 13 dove solo 2 ennuple partecipano al risultato.

Impiegato	Reparto	Capo
Neri	B	Mori
Bianchi	B	Mori

Figura 13: relazione risultante in un Join incompleto.

6. Join vuoto

Come ultimo caso particolare, consideriamo il caso di Figura 14 nessun impiegato appartiene ad alcun reparto presente nella tabella Reparti. In questo caso il risultato è un **Join vuoto**, ovvero un Join che produce una relazione vuota illustrata in Figura 15.

Impiegato	Reparto	Reparto	Capo
Rossi	A	D	Mori
Neri	B	C	Bruni
Bianchi	B		

Figura 14: caso di Join vuoto.

Impiegato	Reparto	Capo
-----------	---------	------

Figura 15: relazione vuota prodotta da un Join vuoto.

Bibliografia

- Atzeni P., Ceri S., Fraternali P., Paraboschi S., Torlone R. (2018). Basi di Dati. McGraw-Hill Education.
- Batini C., Lenzerini M. (1988). Basi di Dati. In Cioffi G. and Falzone V. (Eds). Calderini. Seconda Edizione.
- <http://www-db.disi.unibo.it/courses/SIG/algebra2.pdf>