

Proiezione

Dati una relazione $r(X)$ e un sottoinsieme Y di X , la proiezione di r su Y si indica con

$$\Pi_Y (r) \quad \text{o} \quad \text{PROJ}_Y (r)$$

ed è l'insieme di tuple su Y ottenute dalle tuple di r considerando solo i valori su Y .

$$\Pi_Y (r) = \{ t[Y] \mid t \in r \}$$

Una proiezione ha un numero di tuple *minore o uguale* rispetto alla relazione r cui è applicata: limitare il numero di attributi può portare a creare duplicati. Quindi il numero di tuple è uguale se e solo se Y è superchiave per r .

Proiezione

Operatore monadico

Produce un risultato che

- possiede parte degli attributi dell'operando
- contiene tuple cui contribuiscono tutte le tuple dell'operando

- **sintassi**

$\text{PROJ}_{\text{ListaAttributi}} (\text{Operando})$

- **semantica**

il risultato contiene le tuple che si ottengono restringendo tutte le tuple dell'operando agli attributi nella lista (ed eliminando gli eventuali duplicati)

Impiegati

Matricola	Cognome	Filiale	Stipendio
7309	Rossi	Roma	55
5998	Neri	Milano	64
9553	Rossi	Roma	44
5698	Neri	Napoli	64

- visualizzare matricola e cognome di tutti gli impiegati

Matricola	Cognome
7309	Rossi
5998	Neri
9553	Rossi
5698	Neri

PROJ Matricola, Cognome (Impiegati)

- visualizzare cognome e filiale di tutti gli impiegati

Cognome	Filiale
Neri	Napoli
Neri	Milano
Rossi	Roma

PROJ Cognome, Filiale (Impiegati)

Join

Combinando selezione e proiezione, si può estrarre qualunque elemento (es. lo stipendio dell'impiegato con matricola 7309) da **una sola** relazione (nel nostro caso, la tabella Impiegati). Non si possono però correlare informazioni presenti in relazioni diverse.

Il **join** è l'operatore più interessante (potente) dell'algebra relazionale in quanto permette di correlare dati presenti in relazioni diverse.

Join (naturale)

E' l'operatore più caratteristico dell'algebra relazionale, che evidenzia la proprietà del modello relazionale di essere basato su valori. Non ha un corrispettivo nella teoria degli insiemi.

L'operatore di join (naturale) correla dati contenuti in relazioni diverse.

Il suo risultato è una relazione definita sull'unione degli schemi degli operandi, le cui tuple sono ottenute combinando le tuple degli operandi **che hanno valori uguali** su attributi comuni (con lo stesso nome).

Join



Il join naturale $r_1 \diamond r_2$ di $r_1(X_1)$ e $r_2(X_2)$ è una relazione definita su $X_1 \cup X_2$ (che si può scrivere $X_1 X_2$)



$$r_1 \diamond r_2 = \{ t \text{ su } X_1 X_2 \mid t[X_1] \in r_1 \text{ e } t[X_2] \in r_2 \}$$



Il grado della relazione ottenuta è minore o uguale alla somma dei gradi delle due relazioni in quanto gli attributi omonimi compaiono una sola volta.

Per questo motivo la definizione implica, normalmente, che una tupla del risultato sia ottenuta concatenando ogni tupla della prima relazione con ogni tupla della seconda in cui gli attributi comuni hanno lo stesso valore.

Join

Due casi limite:

- **join vuoto**, le due tabelle hanno alcuni attributi in comune ma nessuna tupla di un operando è combinabile con nessuna tupla dell'altro.
 - **prodotto cartesiano** degli operandi: ciascuna delle tuple di un operando è combinabile con tutte le tuple dell'altro. Può verificarsi se:
 - esistono attributi comuni che assumono lo stesso valore per ogni tupla di entrambi gli operandi (e, quindi, non contengono informazione significativa)
- oppure se
- non ci sono attributi comuni.

Join

Se $X_1 \cap X_2$ è vuoto il join naturale equivale al *prodotto cartesiano* fra le relazioni.

Dimostrazione

Le due relazioni non hanno attributi comuni $\Rightarrow$ ogni tupla della prima relazione può concatenarsi con ogni tupla della seconda, soddisfacendo la definizione di join.

Se le due relazioni hanno lo stesso schema ($X_1 = X_2$) il join naturale equivale all'intersezione fra le relazioni.

Dimostrazione

Ogni tupla può «concatenarsi» solo con una tupla uguale $\Rightarrow$ è presente nel risultato finale se e solo se è presente in entrambe le relazioni.

Join



Se ciascuna tupla di ciascun operando contribuisce ad almeno una tupla del risultato il join si dice *completo*.



Se per alcune tuple non è verificata la corrispondenza e quindi non contribuiscono al risultato, le tuple si dicono *dangling*.



Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparto	Capo
A	Mori
B	Bruni

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

- ogni tupla contribuisce al risultato: il join è **completo**

Un Join non completo

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparto	Capo
B	Mori
C	Bruni

Impiegato	Reparto	Capo
Neri	B	Mori
Bianchi	B	Mori

Un Join vuoto

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparto	Capo
D	Mori
C	Bruni

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

Join: proprietà

- $r_1 \diamond r_2$ contiene un numero di tuple compreso fra zero (*risultato = insieme vuoto*) e il prodotto di $|r_1|$ e $|r_2|$ (*risultato = prodotto cartesiano*)
- Se $r_1 \diamond r_2$ è **completo** allora contiene un numero di tuple pari almeno al massimo fra $|r_1|$ e $|r_2|$ (*nessuna tupla resta esclusa, quindi tutte le tuple della relazione di cardinalità maggiore partecipano sicuramente al join*)
- Se $X_1 \cap X_2$ contiene una chiave per r_2 , allora $r_1(X_1) \diamond r_2(X_2)$ contiene al più $|r_1|$ tuple (*ogni tupla di r_1 si associa al più a una tupla di r_2*).

Join: proprietà

- se il join coinvolge una chiave di r_2 e un vincolo di integrità referenziale da r_1 a r_2 , allora il numero di tuple è pari a $|r_1|$
(ogni tupla di r_1 si associa ad una e una sola tupla di r_2)
- $r_1 \diamond r_2 = r_2 \diamond r_1$ il join è *commutativo*
- $(r_1 \diamond r_2) \diamond r_3 = r_1 \diamond (r_2 \diamond r_3)$ il join è *associativo*
Quindi sequenze di join possono essere scritte senza parentesi