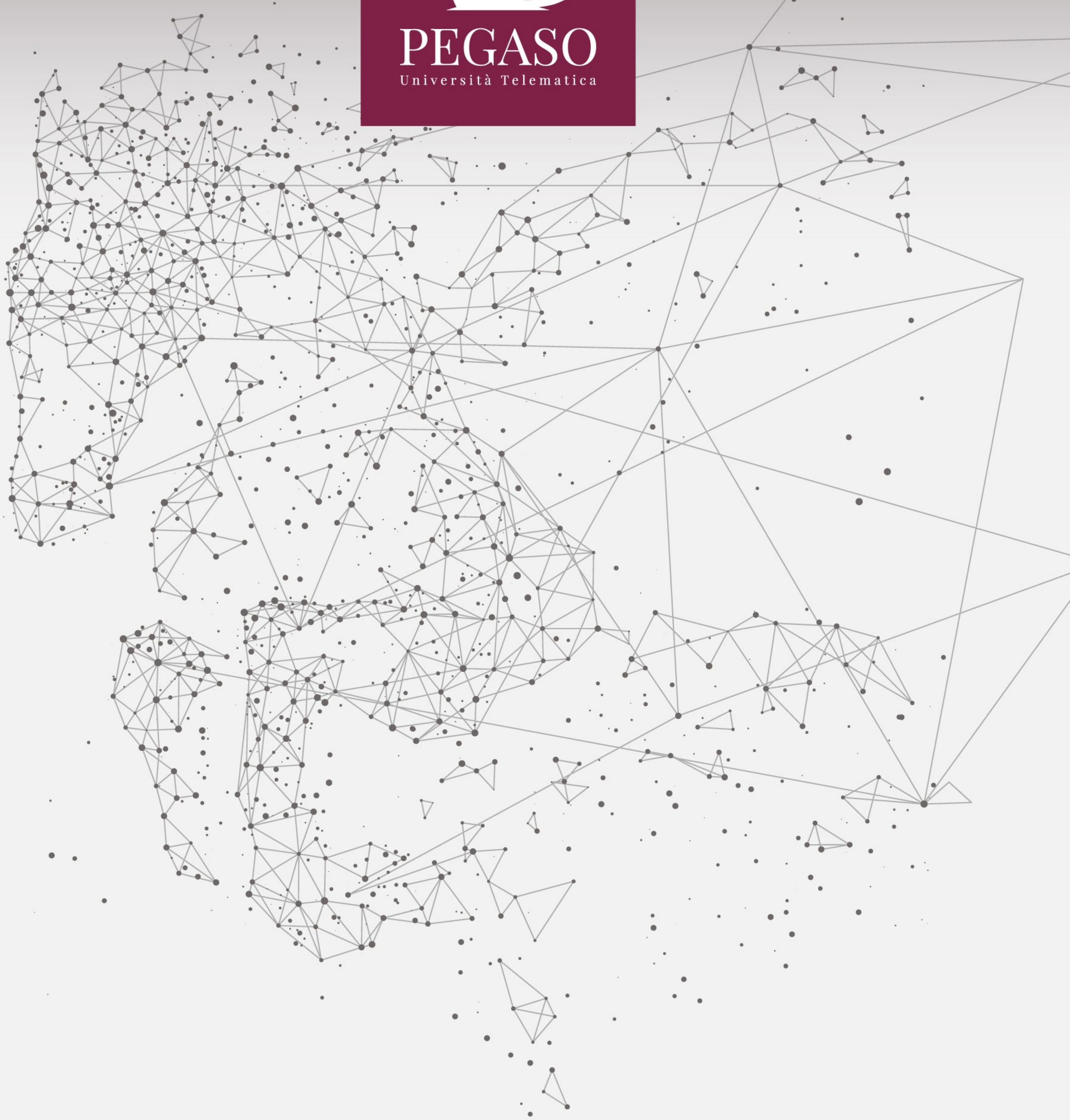
PEGASO
Università Telematica



Indice

1. Introduzione

In questa unità didattica si affronta lo studio delle ultime particolarità dell'operatore Join dell'algebra relazionale: il theta-Join, e l'equi-Join. Si affronta lo studio di due casi estremi, quelli in cui, date due relazioni, i domini sono coincidenti o disgiunti. Infine si illustrano alcune equivalenze nell'uso di insiemi di operatori come la selezione ed il join naturale al fine di apprendere in modo più approfondito le proprietà di tali operatori. Esempi concreti corredano tutti gli argomenti trattati per consolidare le conoscenze acquisite.

2. Join

In questa unità didattica si affronta lo studio delle ultime particolarità dell'operatore Join dell'algebra relazionale: il theta-Join, e l'equi-Join. Si affronta lo studio di due casi estremi, quelli in cui, date due relazioni, i domini sono coincidenti o disgiunti. Infine si illustrano alcune equivalenze nell'uso di insiemi di operatori come la selezione ed il join naturale al fine di apprendere in modo più approfondito le proprietà di tali operatori. Esempi concreti corredano tutti gli argomenti trattati per consolidare le conoscenze acquisite.

In questa sezione illustriamo alcune proprietà dell'operatore join naturale.

Prima proprietà

- Date due relazioni: $R_1(X_1)$, $R_2(X_2)$, si ha:

$$(\pi_{X1}(R_1 \bowtie R_2)) \subseteq R_1$$

Seconda proprietà

- Data una relazione $R(X)$, con $X = X_1 \cup X_2$, si ha:

$$(\pi_{X1}(R) \bowtie (\pi_{X2}(R)) \supseteq R$$

Si dimostra semplicemente utilizzando le definizioni dei due operatori relazionali $\bowtie$ e π.

Proprietà Commutativa

Date due relazioni R_1 ed R_2, il Join naturale gode della seguente proprietà commutativa:

$$R_1 \bowtie R_2 = R_2 \bowtie R_1$$

Proprietà Associativa

Date tre relazioni, R_1, R_2 ed R_3, il Join naturale gode anche della proprietà associativa:

$$R1 \bowtie (R2 \bowtie R3) = (R1 \bowtie R2) \bowtie R3$$

Join n-ario

Date n Relazioni $R_1 \ldots R_n$, si ha:

$$\mathop{\bowtie}_{i=1}^{n} R_i =: R_1 \bowtie R_2 \bowtie \ \ldots \ R_n$$

La dimostrazione segue in modo immediato dalla definizione degli operatori relazionali coinvolti.

3. Casi estremi

Consideriamo adesso due casi particolari riguardanti gli insiemi di attributi X_1 e X_2:

- Caso insiemi disgiunti;
- Caso insiemi coincidenti.

Insiemi disgiunti

In questo caso, si deve effettuare un join naturale su relazioni **senza attributi** in comune. Il risultato di tale operazione è una nuova relazione la quale:

- Contiene sempre un numero di ennuple pari al **prodotto** delle cardinalità degli operandi (le ennuple sono tutte combinabili). Si dimostra seguendo la definizione di Join naturale.
- Ne consegue che Il Join diviene quindi un **Prodotto Cartesiano.**

Vediamo un esempio. Date le due relazioni Impiegati e Reparti di Figura 1:

Impiegati

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparti

Codice	Capo
A	Mori
B	Bruni

Figura 1: le due relazioni Impiegati e Reparti.

Come risultato si ha una nuova relazione, contenente **tutte** le possibili ennuple, ovvero si ha il Prodotto **Cartesiano** tra X_1 e X_2, illustrato in Figura 2.

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

Figura 2: relazione prodotto cartesiano.

Insiemi coincidenti

In questo caso di X_1 coincidente con X_2, il Join naturale coincide con l'intersezione:

$$\mathbf{R_1(X_1) \bowtie R_2(X_2) = R_1(X_1) \cap R_2(X_2)}$$

Tale risultato si dimostra in modo immediato utilizzando la definizione dell'operatore relazionale Join.

4. Theta-Join

Dato il Join naturale su due domini X_1 ed X_2 disgiunti, il risultato del Join naturale, come osservato precedentemente, è il prodotto cartesiano, ovvero l'insieme di tutte le coppie che si possono formare prendendo una ennupla da R_1 ed una ennupla da R_2.

Tale operazione ha in pratica senso, nella maggior parte dei casi, solo se seguito da operazione di selezione. Quindi, date due relazioni R_1 ed R_2, con soluzione il prodotto cartesiano, si ha che l'operazione più comune e utile risulta la seguente:

$$\sigma_{Condizione}\ (R_1 \bowtie R_2)$$

Questa operazione viene chiamata **theta-join** e indicata con:

$$R_1 \bowtie_{Condizione} R_2$$

Perché Theta-Join?:

- La condizione **C** è spesso una congiunzione (**AND**) di atomi di confronto $\mathbf{A_1 \vartheta\ A_2}$ dove ϑ è uno degli operatori di confronto (=, >, <, …).

<u>Esempio</u>

Date le due relazioni Impiegati (Impiegato, Reparto) e Reparti (Codice, Capo), illustrate in Figura 3, dove non esistono attributi in comune, vogliamo conoscere i dipendenti che afferiscono allo stesso reparto.

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Codice	Capo
A	Mori
B	Bruni

Figura 3: le relazioni Impiegati (a sinistra) e Reparti (a destra).

Calcoliamo prima il Join naturale tra le due tabelle:

$$\mathbf{R3 = (Impiegati \bowtie_{Reparto=Codice} Reparti)}$$

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

Figura 4: il risultato del join tra Impiegati e Reparti.

$$\mathbf{R4=\sigma_{Reparto=Codice}(R3)}$$

Cancelliamo manualmente tutte le ennuple che non soddisfano la condizione, come illustrato in Figura 5.

R4=σReparto=Codice(R3) togliamo le ennuple non conformi:

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

Figura 5: cancellazione delle ennuple non soddisfacenti la condizione.

La relazione risultante è quella illustrata in Figura 6.

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

Figura 6: relazione risultate.

5. Equi-Join

In questa sezione illustriamo un caso particolare del theta-join chiamato equi-Join quando l'operatore di confronto nel theta-join è l'uguaglianza (=).

Esempio

Date le due relazioni: Impiegati(Impiegato, Reparto) e Reparti(Codice, Capo), vogliamo ottenere la relazione:

$$\textbf{Impiegati} \bowtie_{\textbf{Reparto=Codice}} \textbf{Reparti}$$

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Codice	Capo
A	Mori
B	Bruni

Figura 7: le due relazioni Impiegati e Reparti.

Il risultato è quello illustrato in Figura 8.

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

Figura 8: theta-Join.

6. Equivalenza tra operazioni

In questa sezione trattiamo una proprietà molto importante delle operazioni relazionali: l'equivalenza. Diremo che due espressioni relazionali sono **equivalenti** se e solo se **producono lo stesso risultato qualunque sia l'istanza attuale della base di dati**. Tale proprietà risulta importante anche perché i DBMS cercano di eseguire espressioni equivalenti a quelle date, ma meno "costose". Un'equivalenza importante è la **Push selections:** se A è attributo di R_2, si ha:

$$\sigma_{A=10} (R_1 \bowtie R_2) = R_1 \bowtie (\sigma_{A=10} (R_2))$$

Tale equivalenza riduce in modo significativo la dimensione del risultato intermedio (e quindi il costo dell'operazione). Questa equivalenza si dimostra applicando le definizioni dei singoli operatori relazionali.

7. Conclusioni

Data l'importanza dell'operatore Join, vale la pena concludere questo ciclo di unità didattiche teoriche con un riassunto delle principali proprietà[1]:

- Un theta-join consente relazioni di confronto arbitrarie (come ≥).
- Un equi-join è un theta-join che utilizza l'operatore di uguaglianza.
- Un join naturale equivale a attributi che hanno lo stesso nome in ogni relazione.

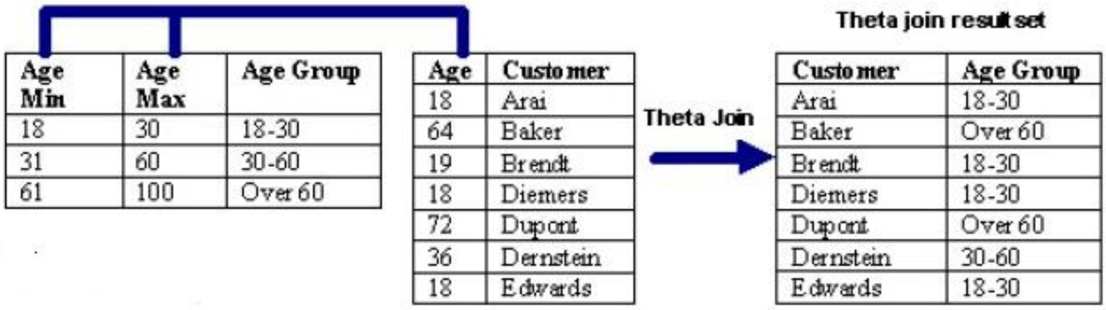


Figura 9: esempio di Theta-Join.

In Figura 9 un esempio di theta-Join tra le relazioni Gruppo (sinistra) e la relazione Customer (destra)[2] dove il theta-Join viene eseguito per avere nella relazione R (Customer, AgeGroup) l'appartenenza di ciascun customer ad un gruppo di età.

[1] https://stackoverflow.com/questions/7870155/difference-between-a-theta-join-equijoin-and-natural-join
[2] https://gerardnico.com/viz/bobj/theta_joins

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.