

Universita' di Milano Bicocca
Corso di Basi di dati 1 in eLearning
C. Batini
5. Algebra Relazionale
5.6 Join - 2

Prodotto cartesiano

Prodotto cartesiano

Un join naturale su relazioni che non hanno attributi in comune.

Indicheremo il prodotto cartesiano con il simbolo \bowtie

- Domanda 5.6.1: Se n_1 e' il numero di n-ple di una relazione R_1 e n_2 di una relazione R_2 , quante n-ple ha $R_1 \bowtie R_2$?

Prodotto cartesiano

Risposta:

$n1 * n2$, perché nel prodotto cartesiano tutte le coppie di ennuple sono combinabili tra loro, e quindi esso contiene sempre un numero di ennuple uguale al prodotto delle cardinalità degli operandi.

Impiegati

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparti

Codice	Capo
1	Mori
2	Bruni

Impiegati \bowtie Reparti

Impiegato	Reparto	Codice	Capo
Rossi	A	1	Mori
Rossi	A	2	Bruni
Neri	B	1	Mori
Neri	B	2	Bruni
Bianchi	B	1	Mori
Bianchi	B	2	Bruni

Il theta-join

Il prodotto cartesiano, in pratica, ha senso solo se seguito da selezione:

$SEL_{\text{Condizione}} (R_1 \times R_2)$

- L'operazione viene chiamata theta-join e indicata con

$R_1 \text{ JOIN}_{\text{Condizione}} R_2$

Perché "theta-join"?

- La condizione C è una congiunzione (**AND**) di espressioni di confronto $A_1 \vartheta A_2$ dove ϑ (**Theta**) è uno degli operatori di confronto ($=, >, <, \dots$)

Esempio di theta-join

Impiegati

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparti

Codice	Capo
A	Mori
B	Bruni

$SEL_{\text{Reparto}=\text{Codice}} (\text{Impiegati} \bowtie \text{Reparti})$

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

L'equi-join

Se l'operatore utilizzato nella condizione è l'uguaglianza (=) allora si parla di **equi-join**, esprimibile con la sintassi

$R1 \text{ JOIN}_{A1 = A2} R2$

Esempio di equi-join

Impiegati

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparti

Codice	Capo
A	Mori
B	Bruni

Impiegati JOIN_{Reparto=Codice} Reparti

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

Lo stesso risultato poteva essere ottenuto nel caso seguente con il join naturale, perché il Reparto è indicato nelle due relazioni con lo stesso nome

Impiegati

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparti

Reparto	Capo
A	Mori
B	Bruni

Impiegati JOIN Reparti

Join naturale e equijoin

- Il join naturale utilizza implicitamente i nomi degli attributi per stabilire la condizione, l'equijoin li indica esplicitamente nella condizione.
- I DBMS tipicamente non permettono il join naturale (solo le ultime versioni di SQL)
- Il join naturale puo' essere simulato per mezzo degli altri operatori ...

Quanti join abbiamo studiato?

- Join naturale
- Join completi e incompleti
- Prodotto cartesiano
- THETA Join
- Equi join

Approfondimento

- Per approfondire la problematica dei join, studia la sezione 3.1.5 del libro di riferimento, dove trovi un insieme di formalizzazioni, considerazioni ed esempi che ti permettono di comprendere meglio l'operazione piu' complessa dell'algebra.
- Inoltre, e' particolarmente importante per questa parte studiare a fondo gli esempi ed esercizi proposti nelle esercitazioni.

Facciamo ora il punto sui principali
concetti dell'algebra relazionale

Interrogazioni in algebra relazionale

- Una interrogazione è una funzione che, applicata a istanze di basi di dati produce relazioni. Meglio:

Dato uno schema R di base di dati, una interrogazione è una funzione che per ogni istanza r di R produce una relazione su un dato insieme di attributi X .

Metodo per produrre interrogazioni in algebra relazionale - 1

1. Dalla espressione della interrogazione in linguaggio naturale, trova prima l'insieme R di relazioni su cui applicare gli operatori.
2. Procedi secondo la strategia "dall'interno all'esterno":
 - 2.1. Prima cerca di capire se servono dei **join** (e quali join) sulle relazioni in R che devono essere collegate per costruire il risultato finale.

Metodo per produrre interrogazioni in algebra relazionale - 2

2.2. Poi cerca di capire se vi sono **operazioni insiemistiche**, che non alterano la struttura della relazione ma possono avvicinare al risultato finale.

2.3 Poi individua eventuali **Selezioni** necessarie per rappresentare analoghe selezioni esistenti nella interrogazione in linguaggio naturale.

2.4 Infine definisci su quali attributi fare le proiezioni, ed in particolare la **Proiezione** finale.

Applichiamo il metodo
rispondendo passo per passo
a una domanda

Domanda 5.6.1

Considera lo schema e la istanza qui a destra. Scrivi una interrogazione in algebra relazionale che produca gli studenti, con matricola e cognome, insieme agli esami sostenuti, con voto e titolo del corso

studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

esami

Studente	Voto	Corso
3456	30	04
3456	24	02
9283	28	01
6554	26	01

corsi

Codice	Titolo	Docente
01	Analisi	Mario
02	Chimica	Bruni
04	Chimica	Verdi

Domanda 5.6.1 - 1

1. Dalla espressione della interrogazione in linguaggio naturale, trova prima l'insieme R di relazioni su cui applicare gli operatori.
Risposta:
Tutte e tre, perché in ciascuna c'è qualche attributo da produrre in uscita

studenti

Matricola	Cognome	Nome	Data nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

esami

Studente	Voto	Corso
3456	30	04
3456	24	02
9283	28	01
6554	26	01

corsi

Codice	Titolo	Docente
01	Analisi	Mario
02	Chimica	Bruni
04	Chimica	Verdi

Domanda 5.6.1 - 2

2.1. Prima cerca di capire se servono dei **join** (e quali join) sulle relazioni in R che devono essere collegate per costruire il risultato finale.

Risposta

Si, occorre mettere in equi join Studenti e Esami su $Matricola = Studente$ e Esami e Corsi su $Corso = Codice$, per avere Tutte le informazioni Necessarie collegate in ogni npla

Relazione risultante

Studenti_Esami_Corsi

(Matricola, Cognome, Nome, DataNascita, Studente, Voto, Corso, Codice, Titolo, Docente)

Domanda 5.6.1 - 3

2.2. Poi cerca di capire se vi sono **Operazioni insiemistiche**, che non alterano la struttura della relazione ma avvicinano al risultato finale.

Risposta

In questo caso non servono operazioni insiemistiche, perché abbiamo già in una tabella tutte le informazioni che ci servono

Relazione risultante

Studenti_Esami_Corsi

(Matricola, Cognome, Nome, DataNascita, Studente, Voto, Corso, Codice, Titolo, Docente)

Domanda 5.6.1 - 4

2.3 Poi individua eventuali
Selezioni necessarie per
rappresentare analoghe selezioni
esistenti nella interrogazione in
linguaggio naturale.

Risposta

Non servono neanche selezioni.
Se invece avessimo cercato, ad
esempio, gli studenti con meno di
20 anni, allora avremmo dovuto
fare una selezione con data di
nascita < data attuale - 20 anni

Relazione risultante

Studenti_Esami_Corsi

(Matricola, Cognome, Nome,
DataNascita, Studente, Voto,
Corso, Codice, Titolo, Docente)

Domanda 5.6.1 - 5

2.4 Infine definisci su quali attributi fare le proiezioni, ed in particolare la Proiezione finale.

Risposta
Occorre fare la Proiezione su Matricola, Cognome, Titolo, Voto.

Relazione risultante

Studenti_Esami_Corsi

(Matricola, Cognome, Titolo, Voto)

Domanda 5.6.1 - 6

Interrogazione finale

PROJ Matricola, Cognome, Titolo, Voto (Studenti JOIN
Matricola=Studente Esami JOIN Corso=Codice Corsi)

Seguono ora sei esercizi a crescente
livello di difficoltà

Esercizi

Partiamo dalla seguente istanza e schema associato.

Impiegati

Matricola	Nome	Età	Stipendio
7309	Rossi	34	45
5998	Bianchi	37	38
9553	Neri	42	35
5698	Bruni	43	42
4076	Mori	45	50
8123	Lupi	46	60

Supervisione

Impiegato	Capo
7309	5698
5998	5698
9553	4076
5698	4076
4076	8123

Esercizi 5.6.1-5.6.6

Esprimere 6 interrogazioni, via via più complesse:

5.6.1. Trovare matricola, nome, età e stipendio degli impiegati che guadagnano più di 40 mila €

5.6.2. Trovare matricola, nome ed età degli impiegati che guadagnano più di 40 mila €

5.6.3. Trovare le matricole dei capi degli impiegati che (gli impiegati!) guadagnano più di 40 mila €

5.6.4. Trovare nome e stipendio dei capi degli impiegati che (gli impiegati!) guadagnano più di 40 mila euro

5.6.5 Trovare gli impiegati che guadagnano più del proprio capo, mostrando matricola, nome e stipendio dell'impiegato e del capo

5.6.6. Trovare le matricole dei capi i cui impiegati guadagnano **tutti** più di 40 mila €

Soluzioni

- Risolvete i precedenti esercizi, e poi confrontate le vostre soluzioni con quelle riportate nella dispensa 5.7 EserciziAR associata alle esercitazioni.
- Se avete difficoltà in qualche esercizio, e vi bloccate, leggete la dispensa, cercando suggerimenti su come proseguire nella soluzione dell'esercizio.

Concetti introdotti

- Prodotto cartesiano
- THETA Join
- Equi join