

Basi di Dati

(Linguaggi per le basi di dati – approfondimento)

Marco Maratea
DeMaCS, Università della Calabria
a.a. 2025/26

Laurea in Informatica

In queste slides presenteremo approfondimenti riguardanti alcune parti riguarda gli esercizi delle lezioni su "Linguaggi per le basi di dati".

Il titolo delle prossime slides corrisponde al numero di slide nel file Linguaggi-partII.pdf oggetto di approfondimento.

Slide 37

Qui mostriamo perchè l'espressione

$PROJ_{Matricola, Nome, Eta} (SEL_{Stipendio > 40}(Impiegati))$

non può essere scritta in maniera equivalente invertendo la selezione e la proiezione.

Questo perchè se proviamo ad invertire i due operatori, ovvero a svolgere l'espressione

$SEL_{Stipendio > 40}(PROJ_{Matricola, Nome, Eta}(Impiegati))$ per passi: per primo eseguiamo $PROJ_{Matricola, Nome, Eta}(Impiegati)$ e otteniamo:

Matricola	Nome	Eta
7309	Rossi	34
5998	Bianchi	37
9553	Neri	42
5698	Bruni	43
4076	Mori	45
8123	Lupi	46

Ora se applichiamo la selezione, ci accorgiamo che Stipendio non vi è più.

Vediamo ora passo passo come viene eseguita la query:
PROJ_{Capo}(Supervisione JOIN_{Impiegato=Matricola}(SEL_{Stipendio>40} Impiegati))

Passo 1. Eseguiamo *SEL_{Stipendio>40} (Impiegati)*

Il risultato, che è quindi anche la soluzione della query a slide 36, viene riportato di seguito:

Matricola	Nome	Età	Stipendio
7309	Rossi	34	45
5698	Bruni	43	42
4076	Mori	45	50
8123	Lupi	46	60

(Ovviamente, il risultato della query a slide 37 è invece la stessa tabella ma senza la colonna Stipendio)

Passo 2. Eseguiamo il JOIN:

*Supervisione JOIN*_{Impiegato=Matricola}(*SEL*_{Stipendio>40} (*Impiegati*))

Svolgiamo quindi il JOIN tra la tabella Supervisione ed il risultato al Passo 1 tramite gli attributi Impiegato e Matricola.

Il risultato è il seguente:

Matricola	Nome	Età	Stipendio	Impiegato	Capo
7309	Rossi	34	45	7309	5698
5698	Bianchi	43	42	5698	4076
4076	Mori	45	50	4076	8123

Passo 3. Eseguiamo la proiezione su Capo della tabella risultato del passo precedente:

$PROJ_{Capo}(Supervisione JOIN_{Impiegato=Matricola}(SEL_{Stipendio>40} Impiegati))$

Il risultato dell'espressione è il seguente:

Capo
5698
4076
8123

che, se guardiamo le tabelle di partenza, sono esattamente i codici dei capi che hanno (almeno) un impiegato che guadagna più di 40.

Per la soluzione alla query presente in questa slide

"Trovare nome e stipendio dei capi degli impiegati che guadagnano più di 40 milioni"

partiamo dal risultato, che è una tabella, della query precedente, che chiamiamo X, e svolgiamo un JOIN con Impiegati:

$PROJ_{Nome, Stipendio}(Impiegati JOIN_{Matricola=Capo} X)$

Questa volta, a differenza di prima, il JOIN avviene legando l'attributo Matricola di Impiegati con l'attributo Capo di Supervisione, perchè la query ci richiede nome e stipendio dei capi.

Il risultato dell'espressione è il seguente:

Nome	Stipendio
Bruni	42
Mori	50
Lupi	60

che, se guardiamo le tabelle di partenza, sono esattamente i nomi e gli stipendi dei capi che hanno (almeno) un impiegato che guadagna più di 40.

L'importanza della query riportata in slide 40 e risolta di seguito:

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

risiede nel fatto che, per risolverla, dobbiamo confrontare valori (dello stipendio) che stanno in uno stesso attributo.

E, in algebra relazionale, non vi è nessun operatore che permette di fare questo. Quindi, nella soluzione che vedremo di seguito, l'idea è quella di sdoppiare la tabella Impiegati, ridenominare gli attributi, ed a quel punto siamo in grado di comparare i valori degli stipendi.

Slide 40 (2)

Svolgiamo la query per passi.

Passo 1. Eseguiamo il primo join tra Supervisore ed Impiegati legati su Impiegato e Matricola perchè, intuitivamente, ora siamo interessati ai dati degli impiegati:

Supervisione JOIN_{Impiegato=Matricola} Impiegati
il cui risultato è:

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

Chiamiamo Y questa tabella risultante.

Passo 2. Eseguiamo la ridenominazione di tutti gli attributi di Impiegati, perchè in questo caso vogliamo che i dati si riferiscano, intuitivamente, ai capi (Matricola contratto con Matr, Stipendio come Stip, ed una "C" aggiunto alla fine del nome dell'attributo), così poi da poter comparare gli stipendi:

REN_{MatrC, NomeC, StipC, EtaC} ← Matr, Nome, Stip, Eta (Impiegati)
il cui risultato è:

MatrC	NomeC	EtaC	StipC
7309	Rossi	34	45
5998	Bianchi	37	38
9553	Neri	42	35
5698	Bruni	43	42
4076	Mori	45	50
8123	Lupi	46	60

Passo 3. Il prossimo passo è il join tra quest'ultima tabella ridenominata e la tabella Y, ovvero:

$REN_{MatrC, NomeC, StipC, EtaC} \leftarrow Matr, Nome, Stip, Eta$ (Impiegati) JOIN $MatrC=Capo$ Y

su MatrC e Capo perchè ora le informazioni di interesse dopo la ridenominazione sono quelle dei capi, da cui si ottiene:

Matr	Nome	Età	Stip	Impiegato	Capo	MatrC	NomeC	EtaC	StipC
7309	Rossi	34	45	7309	5698	5698	Bruni	43	42
5998	Bianchi	37	38	5998	5698	5698	Bruni	43	42
9553	Neri	42	35	9553	4076	4076	Mori	45	50
5698	Bruni	43	42	5698	4076	4076	Mori	45	50
4076	Mori	45	50	4076	8123	8123	Lupi	46	60

Passo 4. A questo punto si applica prima la selezione sotto condizione che $Stip > StipC$ alla tabella precedente ottenendo:

Matr	Nome	Età	Stip	Impiegato	Capo	MatrC	NomeC	StipC	EtaC
7309	Rossi	34	45	7309	5698	5698	Bruni	43	42

e poi, **Passo 5.**, la proiezione sui 6 attributi di interesse, con risultato:

Matr	Nome	Stip	MatrC	NomeC	StipC
7309	Rossi	45	5698	Bruni	42

che indica che Rossi è l'unico impiegato che guadagna più del suo capo (Bruni).

Per la soluzione alla query presente in questa slide

"Trovare le matricole dei capi i cui impiegati guadagnano tutti più di 40 milioni"

non abbiamo un metodo "diretto": tale soluzione richiederebbe qualche operatore che ci consenta di considerare "tutti" gli elementi che hanno una certa proprietà (il JOIN, come visto prima, intuitivamente ci tira fuori quelli che hanno almeno un impiegato che ...).

Quindi, per scrivere una espressione in algebra relazionale che risolve questa query, dobbiamo utilizzare un metodo indiretto che parte da tutti i capi e toglie quelli che sicuramente NON hanno tutti gli impiegati che guadagnano più di 40.

Slide 41 (2)

Per avere tutti gli impiegati basta scrivere la seguente espressione:

PROJ_{Capo} Supervisione

che ha come risultato:

Capo
5698
4076
8123

Per trovare invece i capi che sicuramente NON hanno tutti gli impiegati che guadagnano più di 40, possiamo sfruttare la soluzione della query a slide 38 e cambiare la condizione: se un capo ha **ALMENO** un impiegato che guadagna 40 o meno, sicuramente NON avrà tutti gli impiegati che guadagnano più di 40:

PROJ_{Capo}(Supervisione JOIN_{Impiegato=Matricola}(SEL_{Stipendio} ≤ 40 Impiegati))

Slide 41 (3)

L'espressione precedente ha come risultato:

Capo
5698
4076

perchè i capi con codice 5698 e 4076 hanno un impiegato che guadagna meno di 40.

A questo punto, facciamo la differenza insiemistica tra le due espressioni nella slide precedente, ovvero tra le due tabelle risultanti, che si può fare visto che sono definite sullo stesso schema, e otteniamo il risultato:

Capo
8123

visto che il capo con codice 8123 è il solo che ha tutti gli impiegati che guadagnano più di 40 (ne ha uno e guadagna più di 40).