Esempio: Join Studenti ed Esami

Matricola	Nome	Cognome
1	Luca	Rossi
2	Maria	Bianchi
3	Giorgio	Viola
4	Silvia	Neri
5	Enzo	Verdi

$\bowtie$

Matricola	Voto
1	25
3	30
2	23
1	29
4	29
1	26
5	30
4	30

=

Nome	Cognome	Matricola	Voto
Luca	Rossi	1	25
Giorgio	Viola	3	30
Maria	Bianchi	2	23
Luca	Rossi	1	29
Silvia	Neri	4	29
Luca	Rossi	1	26
Enzo	Verdi	5	30
Silvia	Neri	4	30

Problemi: Matricola, Nome, cognome, voto degli studenti:

- con (almeno un) voto maggiore di 28 (quantificatore esistenziale)
- non hanno mai ottenuto un voto maggiore di 28 (differenza)
- Nomi degli studenti che hanno ottenuto solo voti maggiore di 28 (quantificatore universale)

Quantificatore esistenziale - Parte I

Problema: Matricola, Nome, cognome, voto degli studenti: con (almeno un) voto maggiore di 28 (quantificatore esistenziale)

Matricola	Nome	Cognome
1	Luca	Rossi
2	Maria	Bianchi
3	Giorgio	Viola
4	Silvia	Neri
5	Enzo	Verdi

$\bowtie$

Matricola	Voto
1	25
3	30
2	23
1	29
4	29
1	26
5	30
4	30

$=$

Nome	Cognome	Matricola	Voto
Luca	Rossi	1	25
Giorgio	Viola	3	30
~~Maria~~	~~Bianchi~~	~~2~~	~~23~~
Luca	Rossi	1	29
Silvia	Neri	4	29
Luca	Rossi	1	26
Enzo	Verdi	5	30
Silvia	Neri	4	30

Luca Rossi va eliminato?

?

Join e selezione

Esempio di trasformazione su quantificatore esistenziale

Matricola, Nome, cognome, materia, data, voto degli studenti con voto maggiore di 28.

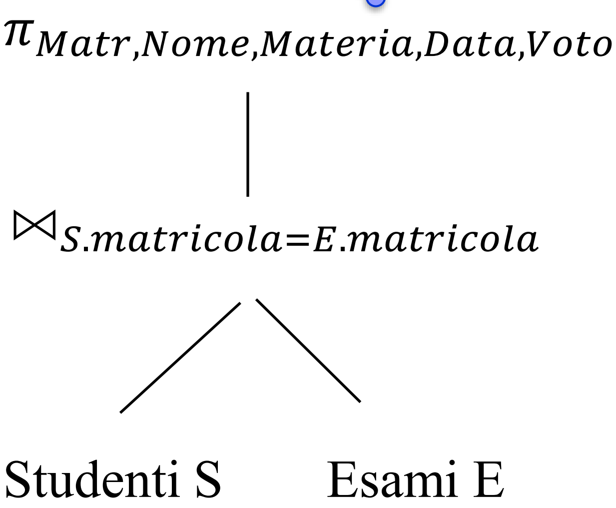
In che punto applichiamo la restrizione sul voto?

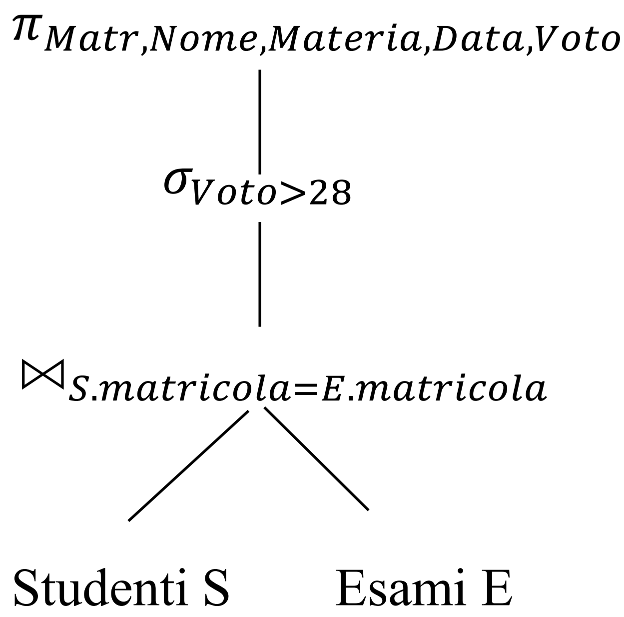
Studenti

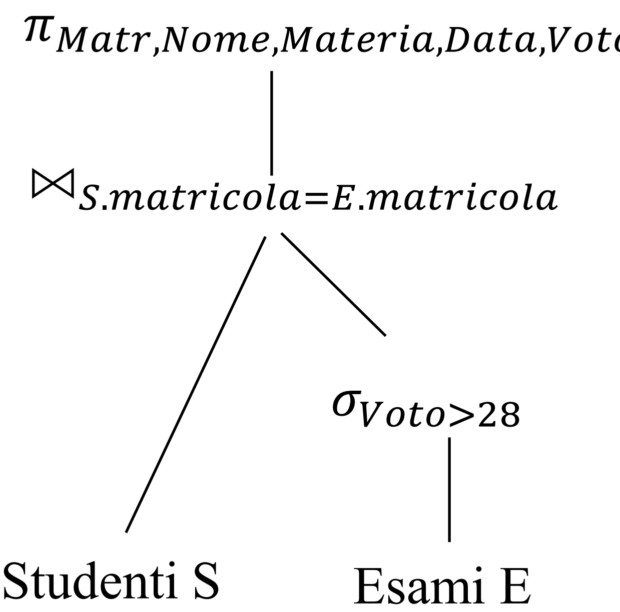
Nome	<u>Matricola</u>	Provincia	AnnoNascita

Esami

<u>Materia</u>	<u>Matricola*</u>	Data	Voto



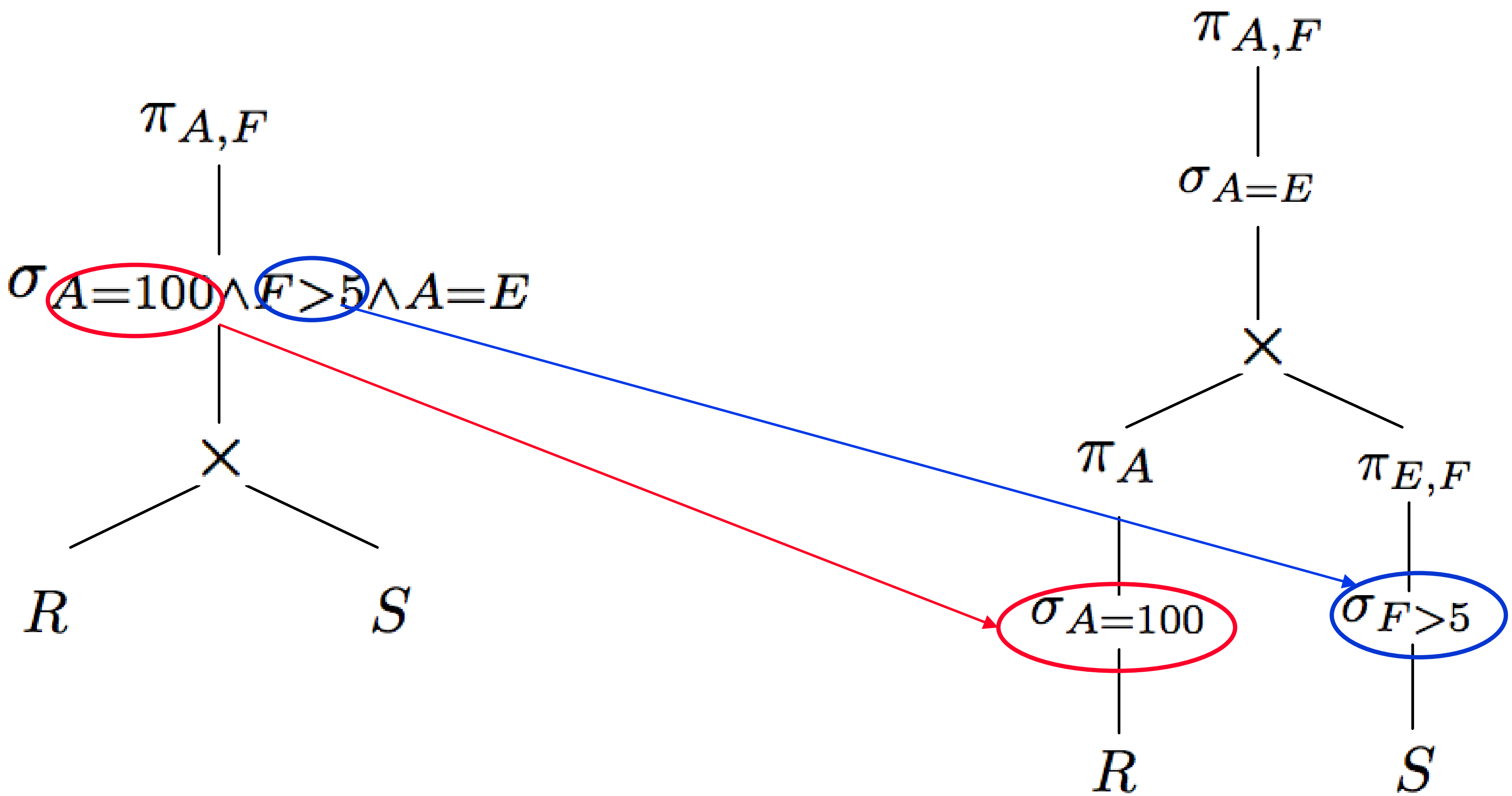




ALBERI LOGICI E TRASFORMAZIONI ALGEBRICHE

- Consideriamo le relazioni R(A, B, C, D) e S(E, F, G) e l'espressione:

$$\pi_{A,F}(\sigma_{A=100 \wedge F>5 \wedge A=E}(R \times S))$$



Esempio 2 – Differenza – Parte I

Problema: studenti che non hanno mai ottenuto un voto maggiore di 28 (differenza)

Matricola	Nome	Cognome
1	Luca	Rossi
2	Maria	Bianchi
3	Giorgio	Viola
4	Silvia	Neri
5	Enzo	Verdi

$\bowtie$

Matricola	Voto
1	25
3	30
2	23
1	29
4	29
1	26
5	30
4	30

=

Nome	Cognome	Matricola	Voto
~~Luca~~	~~Rossi~~	~~1~~	~~25~~
~~Giorgio~~	~~Viola~~	~~3~~	~~30~~
Maria	Bianchi	2	23
~~Luca~~	~~Rossi~~	~~1~~	~~29~~
~~Silvia~~	~~Neri~~	~~4~~	~~29~~
~~Luca~~	~~Rossi~~	~~1~~	~~26~~
~~Enzo~~	~~Verdi~~	~~5~~	~~30~~
~~Silvia~~	~~Neri~~	~~4~~	~~30~~

Selezioniamo gli studenti che hanno preso voti >28

Faccio la differenza fra gli studenti e questo insieme

Esempio 2 – Differenza – Parte II

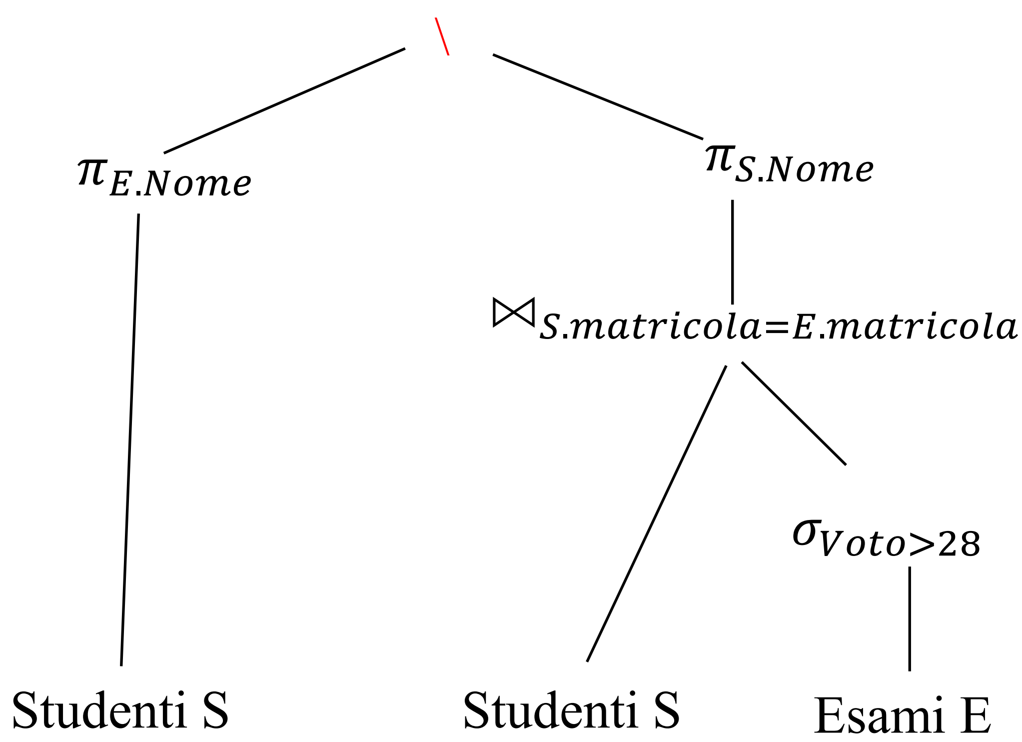
Nome degli studenti che non hanno mai ottenuto un voto maggiore di 28.

Studenti

Nome	Matricola	Provincia	AnnoNascita

Esami

Materia	Matricola*	Data	Voto



Esempio 3 – Quantificatore universale – Parte I

Problema: studenti che hanno ottenuto solo voti maggiore di 28 (quantificatore universale)

Output: Giorgio, Silvia,Enzo

Matricola	Nome	Cognome
1	Luca	Rossi
2	Maria	Bianchi
3	Giorgio	Viola
4	Silvia	Neri
5	Enzo	Verdi

⋈

Matricola	Voto
1	25
3	30
2	23
1	29
4	29
1	26
5	30
4	30

=

Nome	Cognome	Matricola	Voto
~~Luca~~	~~Rossi~~	~~1~~	~~25~~
Giorgio	Viola	3	30
~~Maria~~	~~Bianchi~~	~~2~~	~~23~~
~~Luca~~	~~Rossi~~	~~1~~	~~29~~
Silvia	Neri	4	29
~~Luca~~	~~Rossi~~	~~1~~	~~26~~
Enzo	Verdi	5	30
Silvia	Neri	4	30

Selezioniamo gli studenti che hanno preso voti <=28

Faccio come prima la differenza fra gli studenti e questo insieme?

Giorgio	Viola
Silvia	Neri
Enzo	Verdi

Esempio 3 – quantificatore universale – Parte II

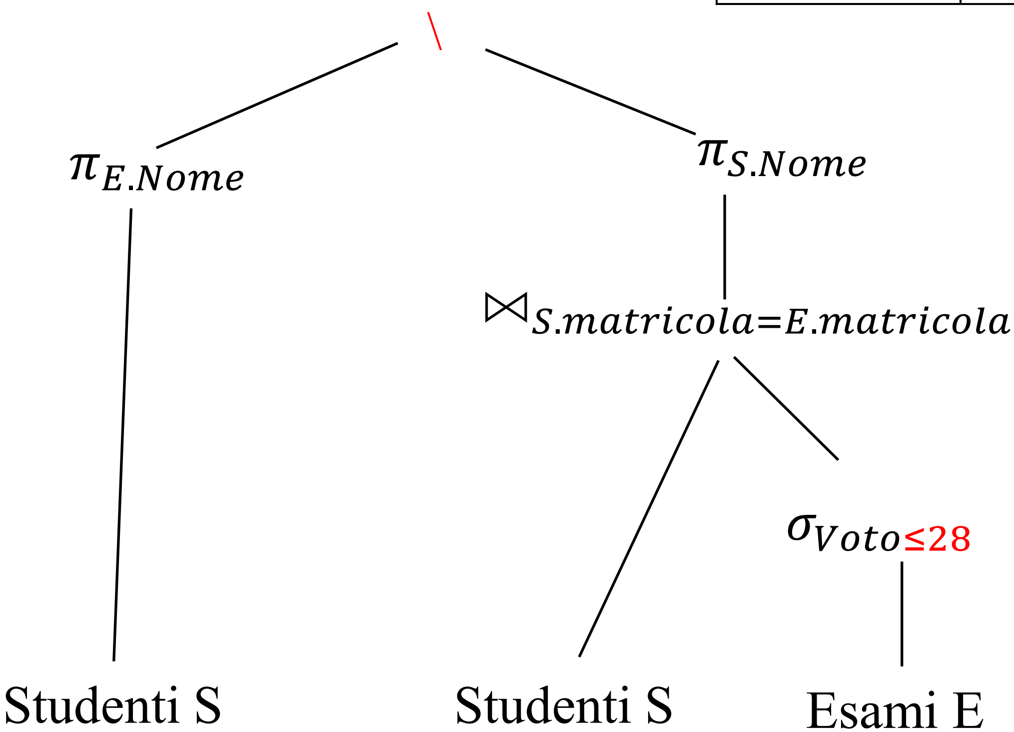
Nomi degli studenti che hanno ottenuto solo voti maggiore di 28.

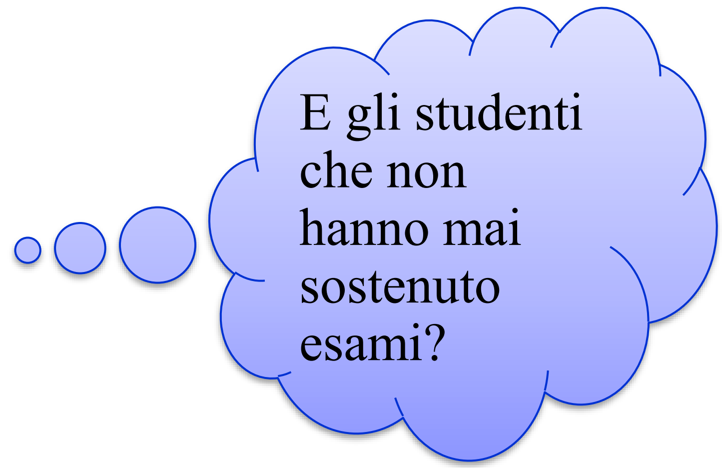
Studenti

Nome	<u>Matricola</u>	Provincia	AnnoNascita

Esami

<u>Materia</u>	<u>Matricola*</u>	Data	Voto





Esempio 3 – Quantificatore universale – Parte I

Studenti che hanno sostenuto almeno un esame

Problema: studenti che hanno ottenuto solo voti maggiore di 28 (**escludendo quelli che non hanno sostenuto esami**).

Matricola	Nome	Cognome
1	Luca	Rossi
2	Maria	Bianchi
3	Giorgio	Viola
4	Silvia	Neri
5	Enzo	Verdi

⋈

Matricola	Voto
1	25
3	30
2	23
1	29
4	29
1	26
4	30

=

Nome	Cognome	Matricola	Voto
Luca	Rossi	1	25
Giorgio	Viola	3	30
Maria	Bianchi	2	23
Luca	Rossi	1	29
Silvia	Neri	4	29
Luca	Rossi	1	26
Silvia	Neri	4	30

Nota che Enzo non ha sostenuto esami!

Esami degli Studenti con voto <=28

Nome	Cognome	Matricola	Voto
Luca	Rossi	1	25
Maria	Bianchi	2	23
Luca	Rossi	1	26

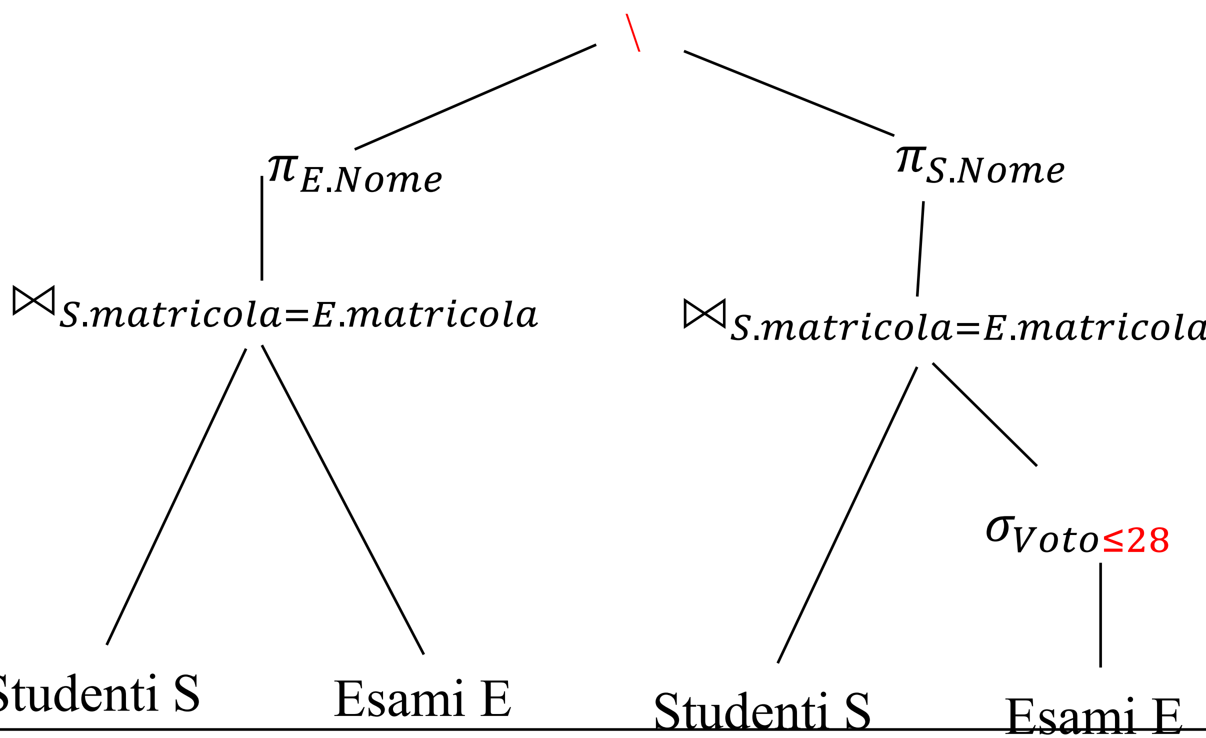
Faccio la differenza fra gli studenti che hanno sostenuto almeno un esame e questo insieme

Esempio 4 – quantificatore universale – Parte III

Nomi degli studenti che hanno ottenuto solo voti maggiore di 28 (escludendo quelli che non hanno sostenuto esami).

Studenti	Nome	<u>Matricola</u>	Provincia	AnnoNascita

Esami	<u>Materia</u>	<u>Matricola*</u>	Data	Voto



Non Distributività della proiezione rispetto alla differenza

- In generale $\boxed{\pi_A(R_1 - R_2) <> \pi_A(R_1) - \pi_A(R_2)}$
- Se R1 e R2 sono definite su AB, e contengono tuple uguali su A e diverse su B

Imp1

Impiegato	Capo
Neri	Mori
Bianchi	Bruni
Verdi	Bini

Imp2

Impiegato	Capo
Neri	Rossi
Bianchi	Bordeaux
Verdi	Blu

$$\pi_A (Imp1 - Imp2) \equiv \pi_A (Imp1) - \pi_A (Imp2) ?$$

Dipende da chi è A….

Se R_1 e R_2 sono definite su AB e contengono tuple uguali su A e diverse su B, NO

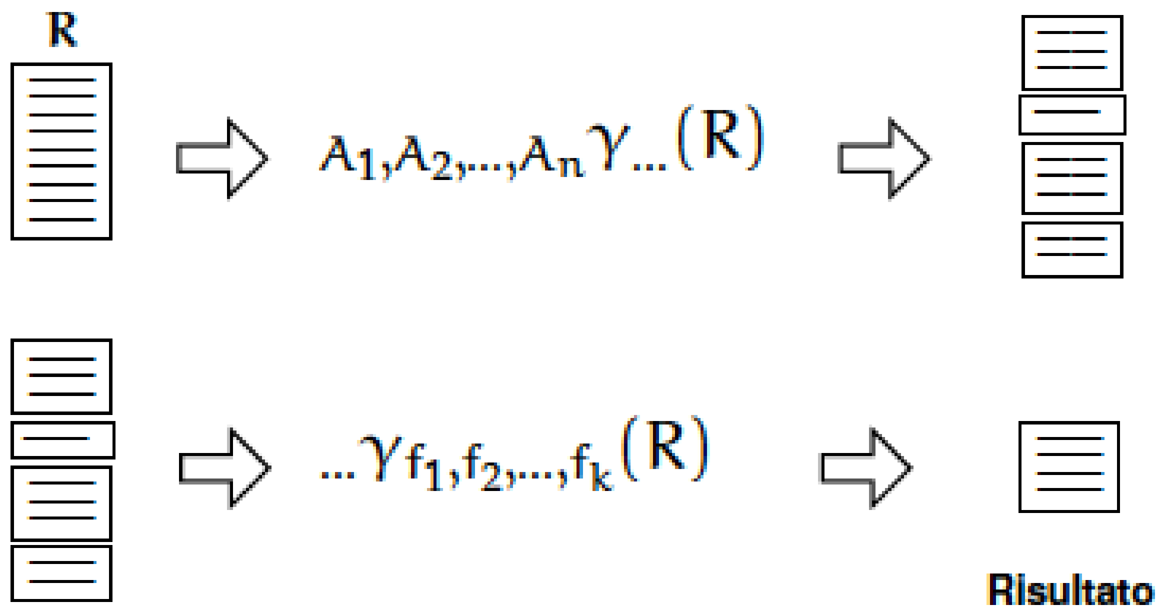
ALGEBRA RELAZIONALE: RAGGRUPPAMENTO

- Raggruppamento: ${}_{\{A_i\}}\gamma_{\{f_i\}}(R)$
- Gli A_i sono attributi di R e le f_i sono espressioni che usano funzioni di aggregazione (min, max, count, sum, …)

- Il valore del raggruppamento è una relazione calcolate come segue
  - Si partizionano le ennuple di R mettendo nello stesso gruppo tutte le ennuple con valori uguali degli A_i
  - Si calcolano le espressioni f_i per ogni gruppo
  - Per ogni gruppo si restituisce una sola ennupla con attributi i valori degli A_i e delle espressioni f_i

ALGEBRA RELAZIONALE: RAGGRUPPAMENTO

$$_{\{A_i\}}\gamma_{\{f_i\}}(R)$$



ESECUZIONE DEL RAGGRUPPAMENTO

- Per ogni candidato: numero degli esami, voto minimo, massimo e medio:

$$_{\{\text{Candidato}\}}\gamma_{\{\text{count}(*),\text{min}(\text{Voto}),\text{max}(\text{Voto}),\text{avg}(\text{Voto})\}}(Esami)$$

Materia	Candidato	Voto	Docente
DA	1	20	10
LFC	2	30	20
MTI	1	30	30
LP	2	20	40

→

Materia	Candidato	Voto	Docente
DA	1	20	10
MTI	1	24	30
LFC	2	30	20
LP	2	20	40

Candidato	Count(*)	min(Voto)	max(Voto)	avg(Voto)
1	2	20	24	22
2	2	20	30	25

Esempio su due attributi

Clienti

IDCLIENTE	COGNOME	NOME
192	Zavoli	Luigi
97	Grassi	Maria
114	Di Santo	Luigi
42	Zavoli	Luigi
5	Di Santo	Luigi
138	Grassi	Maria
12	Zavoli	Maria

Raggruppamento per nome

```
NOME                 COUNT(*)
-------------------- ----------
Maria                         3
Luigi                         4
```

Raggruppamento per cognome

	COGNOME	COUNT(*)
1	Zavoli	3
2	Grassi	2
3	Di Santo	2

Raggruppamento per cognome, nome

	COGNOME	NOME	COUNT(*)
1	Zavoli	Luigi	2
2	Di Santo	Luigi	2
3	Zavoli	Maria	1
4	Grassi	Maria	2

Concettualmente diversi: prima si raggruppa per cognome (risp. Nome) e la partizione ottenuta si partiziona ulteriormente per nome (risp. Cognome).

Il risultato è uguale perché la partizione finale (cognome-nome) è uguale a quella (nome-cognome)

Raggruppamento per nome, cognome

NOME	COGNOME	COUNT(*)
Maria	Grassi	2
Luigi	Zavoli	2
Luigi	Di Santo	2
Maria	Zavoli	1

RAGGRUPPAMENTO: misura e dimensione

- Posso raggruppare per voto?

Materia	Candidato	Voto	Docente
DA	1	20	10
LFC	2	30	20
MTI	1	30	30
LP	2	20	40

- Cosa succede se calcolo min(voto)?

RAGGRUPPAMENTO su chiave primaria

- Posso raggruppare per matricola?

Nome	<u>Matricola</u>	Provincia	AnnoNascita
Isaia	071523	PI	1982
Rossi	067459	LU	1984
Bianchi	079856	LI	1983
Isaia	075649	PI	1984

- Cosa succede se calcolo count(*)?

RAGGRUPPAMENTO su chiave primaria

Studenti

Nome	<u>Matricola</u>	Provincia	AnnoNascita

Esami

<u>Materia</u>	<u>Matricola*</u>	Data	Voto

Ridenominazione nella proiezione

Per ogni studente, visualizzare nome, cognome e numero degli esami sostenuti

$$\pi_{S.Nome, S.Cognome, \color{red}{count(*) as\ nEsami}}$$

$$_{\{nome,\ cognome\}}\gamma_{\{count(*)\}}$$

$$\bowtie_{S.matricola=E.matricola}$$

Studenti S   Esami E

Funziona?

RAGGRUPPAMENTO su chiave primaria

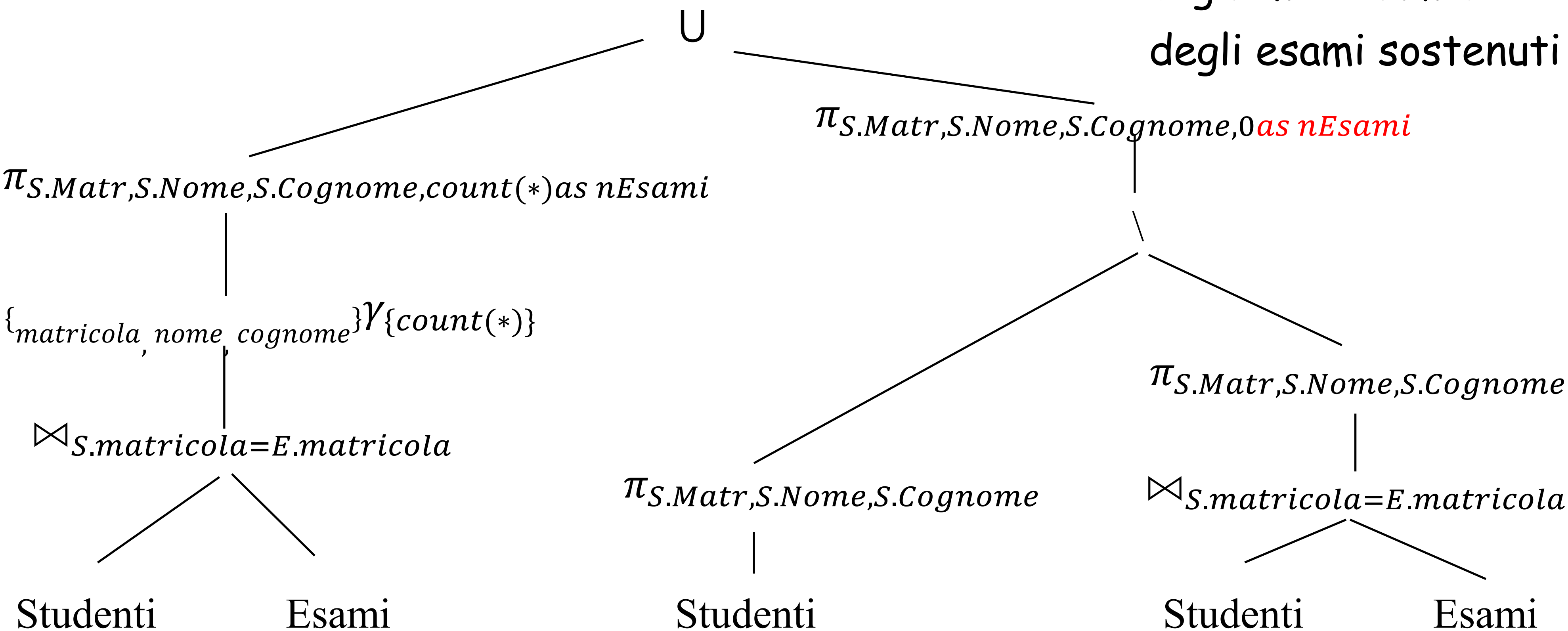
Studenti

Nome	<u>Matricola</u>	Provincia	AnnoNascita

Esami

<u>Materia</u>	<u>Matricola*</u>	Data	Voto

- Per ogni studente, visualizzare matricola, nome, cognome e numero degli esami sostenuti



TRASFORMAZIONI ALGEBRICHE

- Basate su regole di equivalenza fra espressione algebriche
- Consentono di scegliere diversi ordini di join e di anticipare proiezioni e restrizioni.
- Alcuni esempi con la relazione R(A, B, C, D):

$$\pi_A(\pi_{A,B}(R)) \equiv \pi_A(R)$$

$$\sigma_{C_1}(\sigma_{C_2}(R)) \equiv \sigma_{C_1 \wedge C_2}(R)$$

$$\sigma_{C_1 \wedge C_2}(R \times S) \equiv \sigma_{C_1}(R) \times \sigma_{C_2}(S)$$

$$R \times (S \times T) \equiv (R \times S) \times T$$

$$(R \times S) \equiv (S \times R)$$

$$\sigma_C({}_X\gamma_F(R)) \equiv {}_X\gamma_F(\sigma_C(R))$$

Atomizzazione delle selezioni

$$\sigma_{F1 \wedge F2}(E) = \sigma_{F1} (\sigma_{F2}(E))$$

Una selezione congiuntiva può essere sostituita da una cascata di selezioni atomiche

Idempotenza delle Proiezioni

Una proiezione può essere trasformata in una cascata di proiezioni che eliminano i vari attributi in fasi diverse

$$\boxed{\pi_X(E) = \pi_X(\pi_{XY}(E))}$$

se E è definita su un insieme di attributi che contiene Y oltre che X.

Non ha effetto sull'efficienza. Ha effetto sulla leggibilità della query.

Anticipazione della selezione rispetto al Join (Pushing Selection down)

$$\sigma_F(E_1 \bowtie E_2) = E_1 \bowtie \sigma_F(E_2)$$

se F fa riferimento solo ad attributi di E_2.

Aumenta l'efficienza della query perché la selezione riduce il numero delle righe di E2 prima del join.

Anticipazione della proiezione rispetto al join (Pushing Projections down)

$$\boxed{\pi_{X1Y2}(E_1 \bowtie E_2) = E_1 \bowtie \pi_{Y2}(E_2)}$$

Se E_1 e E_2 definite rispettivamente su X_1 e X_2, $Y_2 \subseteq X_2$ e gli attributi in X_2-Y_2 non sono coinvolti nel join.

Non ha effetto sull'efficienza ma sulla leggibilità

Distributività della selezione rispetto all'unione

$$\sigma_F(E_1 \cup E_2) = \sigma_F(E_1) \cup \sigma_F(E_2)$$

Distributività della selezione rispetto alla differenza

$$\sigma_F(E_1 - E_2) = \sigma_F(E_1) - \sigma_F(E_2)$$

Distributività della proiezione rispetto all'unione

$$\boxed{\pi_X(E_1 \cup E_2) = \pi_X(E_1) \cup \pi_X(E_2)}$$

Non Distributività della proiezione rispetto alla differenza

In generale

$$\boxed{\pi_A(R_1 - R_2) <> \pi_A(R_1) - \pi_A(R_2)}$$

Se R_1 e R_2 sono definite sull'insieme di attributi X=AB, e contengono tuple uguali su A e diverse su B

Esempio

Imp1

Impiegato	Capo
Neri	Mori
Bianchi	Bruni
Verdi	Bini

Imp2

Impiegato	Capo
Neri	Rossi
Bianchi	Bordeaux
Verdi	Blu

$$\pi_A\,(\text{Imp1} - \text{Imp2}) \equiv \pi_A\,(\text{Imp1}) - \pi_A\,(\text{Imp2})\ ?$$

Dipende da chi è A....

Se R_1 e R_2 sono definite su AB e contengono tuple uguali su A e diverse su B, NO

Inglobamento di una selezione in un prodotto cartesiano a formare un join

$$\sigma_F(R_1 \bowtie R_2) \equiv R_1 \bowtie_F R_2.$$

Altre equivalenze

- $\sigma_{F_1 \vee F_2}(R) \equiv \sigma_{F_1}(R) \cup \sigma_{F_2}(R)$.
- $\sigma_{F_1 \wedge F_2}(R) \equiv \sigma_{F_1}(R) \cap \sigma_{F_2}(R)$.
- $\sigma_{F_1 \wedge \neg F_2}(R) \equiv \sigma_{F_1}(R) - \sigma_{F_2}(R)$.

Si noti infine che valgono proprietà commutativa e associativa di tutti gli operatori binari (unione, intersezione, join) tranne la differenza.

OPERATORI ALGEBRICI NON INSIEMISTICI

- $\pi^{b}_{\{Ai\}}(R)$: proiezione multiinsiemistica (senza eliminazione dei duplicati)
- $\tau_{\{Ai\}}(R)$: ordinamento

Calcolo relazionale su ennuple

- Il calcolo relazionale è un linguaggio che permette di definire il risultato di un'interrogazione come l'insieme di quelle ennuple che soddisfano una certa condizione ϕ.

- L'insieme delle matricole degli studenti che hanno superato qualcuno degli esami elencati nella relazione Materie(Materia), si può definire come:

$$\{t.\mathrm{Matricola} \mid t \in \mathrm{Studenti}, \exists m \in \mathrm{Materie}.\ \exists e \in \mathrm{ProveEsami}.\ e.\mathrm{Candidato} = t.\mathrm{Matricola} \wedge e.\mathrm{Materia} = m.\mathrm{Materia}\}$$

- Che è equivalente a

$$\pi_{\mathrm{Matricola}}(\mathrm{Studenti} \underset{\mathrm{Matricola} = \mathrm{Candidato}}{\bowtie} (\mathrm{ProveEsami} \bowtie \mathrm{Materie}))$$

Esercizio 3 – Trovare le matricole dei capi

IMPIEGATI(<u>Matricola</u>, Nome, Età, Stipendio)

Supervisione(<u>Impiegato*, Capo*</u>)

- Trovare matricole dei capi degli impiegati che guadagnano più di 40 mila euro

Esercizio 3 – Trovare le matricole dei capi - Soluzione

IMPIEGATI(Matricola, Nome, Età, Stipendio)

Supervisione(Impiegato*, Capo*)

- Trovare le matricole dei capi degli impiegati che guadagnano più di 40 mila euro

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

Esercizio 3 – Trovare le matricole dei capi - Soluzione

IMPIEGATI(Matricola, Nome, Età, Stipendio)

Supervisione(Impiegato*, Capo*)

- Trovare matricole dei capi degli impiegati che guadagnano più di 40 mila euro

$$\pi_{matricole}(\pi_{Capo}(Supervisione \bowtie_{Impiegato=Matricola} (\sigma_{Stipendio>40}(Impiegati))))$$

Esercizio 4 – Trovare nome e stipendio dei capi

IMPIEGATI(Matricola, Nome, Età, Stipendio)

Supervisione(Impiegato*, Capo*)

Trovare nome e stipendio dei capi degli impiegati che guadagnano più di 40 mila euro

$$\pi_{Nome,Stipendio}\Big(\text{Impiegati} \bowtie_{Matricola=Capo} \pi_{Capo}\big(\text{Supervisione} \bowtie_{Impiegato=Matricola} (\sigma_{Stipendio>40}(\text{Impiegati}))\big)\Big)$$

Esercizio 5

IMPIEGATI(<u>Matricola</u>, Nome, Età, Stipendio)

Supervisione(<u>Impiegato*, Capo*</u>)

Trovare le matricole dei capi i cui impiegati guadagnano tutti più di 40 mila euro

Esercizio 5

IMPIEGATI(<u>Matricola</u>, Nome, Età, Stipendio)

Supervisione(<u>Impiegato*, Capo*</u>)

- Trovare le matricole dei capi i cui impiegati guadagnano **tutti** più di 40 mila euro

Non si può esprimere direttamente nell'algebra, poiché mancano i "quantificatori universali)

Assumiamo per semplicità che non ci siano valori NULL in Stipendio

π_{Capo} (Supervisione) -

π_{Capo} (Supervisione

$\bowtie_{Impiegato=Matricola}$

$(\sigma_{Stipendio \leq 40}$(Impiegati)))

Capi che hanno almeno un Impiegato che guadagna meno di 40

Esercizio 6

IMPIEGATI(<u>Matricola</u>, Nome, Età, Stipendio)

Supervisione(<u>Impiegato*, Capo*</u>)

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

$$\pi_{Matr,Nome,Stip,MatrC,NomeC,StipC}\Big(\sigma_{Stipendio>StipC}\Big(\rho_{\mathbf{MatrC},NomeC,StipC,EtàC \leftarrow Matr,Nome,Stip,Età}(\mathrm{Impiegati}) \bowtie_{\mathbf{MatrC}=Capo} (\mathrm{Supervisione} \bowtie_{Impiegato=Matricola} \mathrm{Impiegati})\Big)\Big)$$

Esercizio 7

IMPIEGATI(<u>Matricola</u>, Nome, Età, Stipendio)

Supervisione(<u>Impiegato*, Capo*</u>)

Trovare quali sono gli impiegati che hanno stipendio massimo

$$\pi_{Matricola}(Impiegati) -$$

$$\pi_{Matricola}\left(Impiegati \bowtie_{Stip < Stip1} \left(\rho_{Matr1,\, Nome1,\, Eta1,\, Stip1 \leftarrow Matr,Nome,Stip,Età}(Impiegati)\right)\right)$$

Partiamo dagli impiegati che **non** hanno lo stipendio massimo