

25. Algebra e operatori relazionali

database

algebra e operatori relazionali

Oltre a definire la struttura della base di dati, il modello relazionale prevede anche un insieme di operazioni di base per il trattamento dei dati, ed in particolare per consentire il recupero di informazioni significative dai dati distribuiti tra le varie relazioni/tabelle; si tratta di operazioni che applicandosi ad una o più tabelle producono ancora una tabella che a sua volta può essere ulteriormente elaborata utilizzando gli stessi operatori.

Questo insieme di operazioni è parte di una struttura algebrica nota come **Algebra Relazionale** introdotta da Edgar Codd come esempio di linguaggio minimale per il calcolo su relazioni, cioè su tabelle.

Una sequenza di operazioni che coinvolge una o più tabelle viene definita *espressione* dell'algebra relazionale, il cui risultato è ancora una volta una tabella.

algebra e operatori relazionali

Riprendiamo un esempio già visto in precedenza; sia dato un piccolo database di due sole tabelle dove riportiamo i dati dei clienti e quelli degli agenti che li seguono.

Tra le due tabelle esiste un'associazione 1:N sull'attributo comune *CodAgente* (chiave primaria per *Agenti* e chiave esterna per *Clienti*).

CodAgente	NomeAgente	CodZona
AK 0098	Carlo Bressi	AB 2
PM 0753	Pietro Conti	PP 5
JP 1123	Alba Sergi	KR 1
Agenti		

CodAgente	CodCli	Nominativo	Indirizzo	Prov
AK 0098	C 002	Antonia Rossi	Via Roma, 126	RM
AK 0098	A 011	Giacomo Alfieri	Via Napoli, 29	NA
JP 1123	A 004	Franco Pisori	Via Venezia, 38	VE
AK 0098	C 014	Maria Petrocelli	Via Bologna, 78	BO
PM 0753	B 003	Piero Giovanelli	Via Palermo, 23	NA
JP 1123	A 012	Carla Vecchi	Via Salerno, 88	MI
Clienti				

algebra e operatori relazionali

Cerchiamo un elenco dei soli clienti della provincia di Napoli con i rispettivi agenti, informazioni che come si può notare si trovano distribuite in due tabelle diverse.

La prima cosa da fare dunque è quella di fondere insieme le due tabelle in questione in un'unica relazione che riporti i dati di ciascun agente insieme con quelli dei rispettivi clienti; ovviamente questa operazione introduce un certo livello di ridondanza sulla tabella prodotta, ma si tratta solo di una tabella temporanea utilizzata esclusivamente per le operazioni in corso.

A questo proposito, abbiamo già notato che le due tabelle non sono indipendenti tra loro, ma legate da un'associazione *uno-a-molti* stabilita dalla presenza di un campo in comune, *CodAgente* che è chiave primaria per *Agenti* e chiave esterna per *Clienti*.

algebra e operatori relazionali

Sfruttiamo l'operatore relazionale binario di **CONGIUNZIONE** tra le due tabelle per fondere insieme le righe di entrambe là dove il campo in comune *CodAgente* contiene le stesse informazioni ottenendo come risultato ancora una tabella (temporanea).



algebra e operatori relazionali

A questo punto selezioniamo i soli record che nel campo *Prov* contengono il valore "NA", applicando alla tabella temporanea appena ottenuta l'operatore relazionale unario di **SELEZIONE**.

CodAgente	NomeAgente	CodZona	CodCli	Nominativo	Indirizzo	Prov
AK 0098	Carlo Bressi	AB 2	C 002	Antonia Rossi	Via Roma, 126	RM
AK 0098	Carlo Bressi	AB 2	A 011	Giacomo Alfieri	Via Napoli, 29	NA
JP 1123	Alba Sergi	KR 1	A 004	Franco Pisori	Via Venezia, 38	VE
AK 0098	Carlo Bressi	AB 2	C 014	Maria Petrocelli	Via Bologna, 78	BO
PM 0753	Pietro Conti	PP 5	B 003	Piero Giovanelli	Via Palermo, 23	NA
JP 1123	Alba Sergi	KR 1	A 012	Carla Vecchi	Via Salerno, 88	MI

Tabella Temporanea

CodAgente	NomeAgente	CodZona	CodCli	Nominativo	Indirizzo	Prov
AK 0098	Carlo Bressi	AB 2	A 011	Giacomo Alfieri	Via Napoli, 29	NA
PM 0753	Pietro Conti	PP 5	B 003	Piero Giovanelli	Via Palermo, 23	NA

Tabella Temporanea

algebra e operatori relazionali

Infine applichiamo a questa ultima l'operatore relazionale unario di **PROIEZIONE** sui campi *NomeAgente* e *Nominativo* per ottenere le informazioni richieste in una semplice tabella finale riassuntiva; si tratta ancora di una tabella temporanea che all'occorrenza può essere memorizzata insieme con le altre già presenti nel database.



algebra e operatori relazionali

Come accennato in precedenza, una caratteristica sostanziale dell'algebra relazionale è la **proprietà di chiusura**: gli operatori cioè si applicano a tabelle per produrre come risultato nuove tabelle.

Questa proprietà è fondamentale in quanto permette di applicare un qualsiasi operatore al risultato di una qualsiasi operazione.

I linguaggi basati sull'algebra relazionale forniscono una sintassi specifica per rappresentare i tradizionali operatori insiemistici (unione, intersezione, differenza, prodotto cartesiano) e altri tipici del calcolo con le relazioni (selezione/restrizione, proiezione, congiunzione, divisione).

algebra e operatori relazionali

L'algebra relazionale ha sei *operatori primitivi* e alcuni *operatori derivati* che, in linea di principio, possono essere definiti come combinazione di quelli primitivi; gli operatori possono essere di tipo *unario*, se operano su una sola tabella, *binario*, se operano su due tabelle. Vediamone alcuni tra i più importanti.

Operatori primitivi:

- **unione** (binario);
- **differenza** (binario);
- **prodotto cartesiano** (binario);
- **selezione** o **restrizione** (unario);
- **proiezione** (unario);
- **ridenominazione** (unario).

Operatori derivati:

- **intersezione** (binario);
- **congiunzione** o **join** (binario);
- **divisione** (binario).

algebra e operatori relazionali

Nella tabella seguente sono riportati sia i simboli sia i nomi degli operatori dell'algebra relazionale corrispondenti:

Attenzione ai simboli, molto simili tra loro, dei seguenti operatori:

- *join*: ⋈;
- *left outer join*: ⋈_L;
- *right outer join*: ⋈_R;
- *full outer join*: ⋈_F.

Operazione	Operatore	Simbolo
Proiezione	PROJECT	π
Selezione	RESTRICT	σ
Ridenominazione	RENAME	ρ
Unione	UNION	\cup
Prodotto cartesiano	TIMES	\times
Congiunzione (<i>join</i>)	JOIN	\bowtie
<i>Left outer join</i>	LEFT OUTER JOIN	\bowtie_L
<i>Right outer join</i>	RIGHT OUTER JOIN	\bowtie_R
<i>Full outer join</i>	FULL OUTER JOIN	\bowtie_F
Divisione	DIVIDE	\div
Intersezione	INTERSECT	\cap
Differenza	MINUS	$-$

algebra e operatori relazionali

Prodotto cartesiano:

Il prodotto cartesiano di due tabelle R ed S **non necessariamente compatibili**

R **TIMES** S oppure $T \leftarrow R \times S$

produce, come tabella risultato, un insieme le cui righe sono costituite dalla concatenazione di ogni riga di R con tutte le righe di S :

R		
x	y	z
p	1	2
p	2	1
q	1	2
r	2	5
r	3	3

S	
k	l
2	t
3	u
4	v

T				
x	y	z	k	l
p	1	2	2	t
p	1	2	3	u
p	1	2	4	v
p	2	1	2	t
p	2	1	3	u
p	2	1	4	v
q	1	2	2	t
q	1	2	3	u
q	1	2	4	v
r	2	5	2	t
r	2	5	3	u
r	2	5	4	v
r	3	3	2	t
r	3	3	3	u
r	3	3	4	v

algebra e operatori relazionali

Selezione (*Restrizione*):

Produce un sottoinsieme «**orizzontale**» (per righe) di una tabella selezionando tutte le righe della stessa che soddisfino una specificata condizione logica; ad es. il comando

RESTRICT R **WHERE** $z=2$ oppure $T \leftarrow \sigma_{(z=2)} R$

restituisce una nuova tabella T in cui sono presenti solo le righe della tabella R in cui l'attributo z assume valore 2:

R				T		
x	y	z		x	y	z
p	1	2	→	p	1	2
p	2	1		q	1	2
q	1	2	↗			
r	2	5				
r	3	3				

algebra e operatori relazionali

Selezione (*Restrizione*):

Ad es. il comando seguente:

RESTRICT Ricoveri **WHERE** Reparto="Ortopedia" **OR** Reparto="Cardiologia"

$T \leftarrow \sigma_{(\text{Reparto}=\text{"Ortopedia"} \text{ OR } \text{Reparto}=\text{"Cardiologia"})} \text{Ricoveri}$

seleziona i pazienti dei soli reparti di *Ortopedia* e *Cardiologia* in una tabella *Ricoveri* che registra i pazienti ospiti di un'azienda ospedaliera.

Paziente	DataRicovero	Reparto	NoteDiagnosi
Atonia Rossi	04/01/2003	Ortopedia	...
Franco Pisori	11/05/2003	Chirurgia	...
Pietro Giovanelli	11/05/2003	Ortopedia	...
Giacomo Alfieri	22/10/2003	Medicina	...
Gilberto Sergi	22/10/2003	Medicina	...
Carla Vecchi	22/10/2003	Cardiologia	...

Paziente	DataRicovero	Reparto	NoteDiagnosi
Atonia Rossi	04/01/2003	Ortopedia	...
Pietro Giovanelli	11/05/2003	Ortopedia	...
Carla Vecchi	22/10/2003	Cardiologia	...

OrtopediaCardiologia

algebra e operatori relazionali

Proiezione:

Produce un sottoinsieme «**verticale**» (cioè per colonne) di una tabella selezionando solo le colonne relative agli attributi specificati ed **eliminando eventuali righe replicate**

PROJECT R **OVER** k, l oppure $T \leftarrow \pi_{(k, l)} R$

restituisce una nuova tabella T in cui sono presenti solo le colonne della tabella R identificate dagli attributi k ed l :

R					T	
x	y	z	k	l	k	l
p	1	2	2	t	2	t
q	1	2	2	t	5	t
r	2	5	5	t	3	u
r	3	3	3	u		

algebra e operatori relazionali

Proiezione:

Ad es. data la tabella in figura, produciamo una nuova tabella con nominativo e codice fiscale dei pazienti registrati:

PROJECT Ricoveri **OVER**
Paziente, DataRicovero

$T \leftarrow \pi_{(\text{Paziente}, \text{DataRicovero})} \text{Ricoveri}$

Paziente	DataRicovero	Reparto	NoteDiagnosi
Atonia Rossi	04/01/2003	Ortopedia	...
Pietro Giovanelli	11/05/2003	Ortopedia	...
Carla Vecchi	22/10/2003	Cardiologia	...

Ricoveri	
Paziente	DataRicovero
Atonia Rossi	04/01/2003
Pietro Giovanelli	11/05/2003
Carla Vecchi	22/10/2003
Ricoveri	

algebra e operatori relazionali

Proiezione:

NB: Quando la lista di attributi non comprende la chiave primaria (o una qualche chiave candidata), è possibile che nel risultato si presentino anche tuple duplicate; eseguiamo ad es. l'operazione seguente:

PROJECT Ordini **OVER**
CodArticolo,
DescrArticolo, Prezzo

Vengono prodotte due tuple identiche.

L'operazione di proiezione deve dunque rimuovere tutte le tuple ripetute (nel nostro caso verrà rimossa una copia della tupla <0098, Cacciavite a stella, 2,50>) in modo che il risultato sia ancora una relazione valida, cioè un insieme di tuple.

<u>CodCli</u>	<u>CodArticolo</u>	<u>DataOrdine</u>	DescrArticolo	Prezzo
C 002	0098	12/11/2003	Cacciavite a stella	€ 2,50
A011	0128	10/10/2003	Termometro	€ 12,75
A004	1123	09/07/2003	Serie brugole	€ 10,00
C 014	0098	01/06/2003	Cacciavite a stella	€ 2,50
Ordini				

<u>CodArticolo</u>	DescrArticolo	Prezzo
0098	Cacciavite a stella	€ 2,50
0128	Termometro	€ 12,75
1123	Serie brugole	€ 10,00
0098	Cacciavite a stella	€ 2,50
TabellaFinale		

algebra e operatori relazionali

Selezione e *Proiezione* svolgono funzioni che potremmo definire ortogonali fra loro, visto che la *Selezione* genera decomposizioni orizzontali della tabella di partenza, mentre la proiezione genera decomposizioni verticali della stessa.



algebra e operatori relazionali

Congiunzione (join):

L'operatore **JOIN** può essere applicato ad una coppia di tabelle *R* ed *S* per produrre una tabella le cui righe sono costituite dal prodotto cartesiano delle righe di *R* ed *S* che soddisfano una specificata condizione (a differenza dell'operatore prodotto cartesiano che produce come risultato tutte le possibili combinazioni tra le righe di due tabelle).

Il formato generale del comando è:

```
JOIN R WITH S WHERE (R.<attributo> = S.<attributo>)
```

dove i due *R.<attributo>* ed *S.<attributo>* devono essere definiti sullo stesso dominio.

algebra e operatori relazionali

Congiunzione (join):

Per esempio il comando:

JOIN R **WITH** S **WHERE** $(z=k)$ oppure $T \leftarrow R \bowtie_{(z=k)} S$

produce la seguente tabella T a partire dalle tabelle R ed S selezionando dal prodotto cartesiano di R ed S le sole righe per le quali il valore dell'attributo z è uguale al valore dell'attributo k :

R			S		T				
x	y	z	k	l	x	y	z	k	l
p	1	2	2	t	p	1	2	2	t
p	2	1	3	u	q	1	2	2	t
q	1	2	4	v	r	2	5	5	t
r	2	5	5	t	r	3	3	3	u
r	3	3	4	t					

algebra e operatori relazionali

Congiunzione (join):

Ad es. date le tabelle *Pazienti* e *Ricoveri*, l'operazione:

```
JOIN Pazienti WITH Ricoveri  
ON Pazienti.CodFiscale  
= Ricoveri.CodFiscale
```

produce la tabella *RicoveriPazienti*
mostrata in figura.

CodFiscale	Paziente	DataNascita
PSRFRN11C22D333E	Franco Pisori	12/11/1980
TSSGFR11C22D333E	GianFranco Tassin	24/07/1969
GVNPTR11C22D333E	Pietro Giovanelli	09/09/1982
BPPRGI11C22D333E	Beppe Regi	07/12/1978
PTRSMN11C22D333E	Simona Petri	10/11/1970
VCCCRL11C22D333E	Carla Vecchi	03/09/1962
SRGGCM11C22D333E		

CodFiscale	DataRicovero	Reparto	NoteDiagnosi
PSRFRN11C22D333E	04/01/2003	Ortopedia	...
PSRFRN11C22D333E	11/05/2003	Chirurgia	...
GVNPTR11C22D333E	11/05/2003	Ortopedia	...
PSRFRN11C22D333E	22/10/2003	Medicina	...
GVNPTR11C22D333E	22/10/2003	Medicina	...
VCCCRL11C22D333E	22/10/2003	Chirurgia	...

CodFiscale	Paziente	DataNascita	DataRicovero	Reparto	NoteDiagnosi
PSRFRN11C22D333E	Franco Pisori	12/11/1980	04/01/2003	Ortopedia	...
PSRFRN11C22D333E	Franco Pisori	12/11/1980	11/05/2003	Chirurgia	...
GVNPTR11C22D333E	Pietro Giovanelli	09/09/1982	11/05/2003	Ortopedia	...
PSRFRN11C22D333E	Franco Pisori	12/11/1980	22/10/2003	Medicina	...
GVNPTR11C22D333E	Pietro Giovanelli	09/09/1982	22/10/2003	Medicina	...
VCCCRL11C22D333E	Carla Vecchi	03/09/1962	22/10/2003	Chirurgia	...

RicoveriPazienti

algebra e operatori relazionali

Congiunzione (join):

Date due tabelle R ed S , l'applicazione dell'operatore `JOIN` non mantiene nel risultato le righe di R che non hanno corrispondenze in S e viceversa.

Per ovviare a questo problema, esiste l'operatore **OUTER JOIN** (*join esterno*) nelle forme **LEFT** (\bowtie), **RIGHT** (\bowtie) e **FULL** (\bowtie), che rende possibile l'inserimento nel risultato delle righe che sarebbero altrimenti ignorate dall'operatore `JOIN`.

L'applicazione dell'operatore `OUTER JOIN` completa con valori nulli le righe che sarebbero eliminate con l'applicazione dell'operatore `JOIN`.

algebra e operatori relazionali

Congiunzione (join):

Per esempio il comando:

LEFT OUTER JOIN R **WITH** S **WHERE** (z=k) oppure $T \leftarrow R \bowtie_{(z=k)} S$

applicato alle seguenti tabelle R ed S produce come risultato la tabella T:

Il risultato finale del left join comprende dunque anche tutte le righe non collegate a sinistra (*left*) dell'operatore di join.

R			S		T				
x	y	z	k	l	x	y	z	k	l
p	1	2	2	t	p	1	2	2	t
p	2	1	3	u	p	2	1	null	null
q	1	2	4	v	q	1	2	2	t
r	2	5	5	t	r	2	5	5	t
r	3	3	4	t	r	3	3	3	u

algebra e operatori relazionali

In virtù della *proprietà di chiusura* di cui godono gli operatori dell'algebra booleana, il risultato di un'operazione su tabelle è ancora una tabella; è possibile dunque applicare più operatori in cascata, realizzando in questo modo un'espressione dell'algebra relazionale.

Ad es., si può dimostrare che un'operazione di selezione (`RESTRICT`) in cui la condizione di selezione sia composta da più clausole elementari collegate tramite l'operatore booleano `AND`, può essere sostituita con una cascata di operazioni di selezione semplice.

algebra e operatori relazionali

Data la tabella *Studenti* che registra gli iscritti ai corsi di laurea di un certo ateneo universitario, vogliamo ottenere un elenco degli studenti nati tra il 1995 ed il 1997.

Matricola	Cognome	Nome	DataNascita	DataIscrizione	Residenza	CorsoLaurea	AnnoCorso
8473	Baldieri	Marco	12/07/1995	02/09/2018	Catanzaro	Scienze Infermieristiche	1
7639	Cinotti	Annarita	02/09/1990	12/10/2017	Cosenza	Fisioterapista	2
8577	Velasco	Ester	09/12/1998	10/10/2018	Crotone	Scienze Infermieristiche	1
4498	Bitonti	Piero	11/10/1993	05/11/2016	Lamezia T.	Scienze Infermieristiche	3
8841	Ragusa	Giacomo	04/05/1995	04/10/2018	Soverato	Tecn. Prevenzione	1
6990	Vecchi	Andrea	14/07/1994	31/10/2016	Cosenza	Fisioterapista	3
8904	Petrocelli	Carmelina	03/06/1997	12/11/2018	Crotone	Scienze Infermieristiche	1
7609	Maurito	Franco	09/09/1987	03/09/2018	Catanzaro	Logopedista	1
5901	Pascale	Irene	13/11/1975	09/11/2017	Lamezia T.	Scienze Infermieristiche	2
Studenti							

Studenti nati tra il 1995 ed il 1997.

algebra e operatori relazionali

RESTRICT Studenti **WHERE**

DataNascita >= "01/01/1995" AND DataNascita <= "31/12/1997"

equivalente a:

RESTRICT

RESTRICT Studenti **WHERE** DataNascita >= "01/01/1995"

WHERE

DataNascita <= "31/12/1997"

Selezione esterna

Selezione interna

In cui si può notare come l'operatore esterno di selezione venga applicato alla tabella risultante dalla selezione più interna.

algebra e operatori relazionali

In pratica con la prima selezione (quella interna) vengono scelte le sole righe relative agli studenti nati dopo il 01/01/1995:

Matricola	Cognome	Nome	DataNascita	DataIscrizione	Residenza	CorsoLaurea	AnnoCorso
8473	Baldieri	Marco	12/07/1995	02/09/2018	Catanzaro	Scienze Infermieristiche	1
6577	Velasco	Ester	09/12/1998	10/10/2018	Crotone	Scienze Infermieristiche	1
8841	Ragusa	Giacomo	04/05/1995	04/10/2018	Soverato	Tecn. Prevenzione	1
8904	Petrocelli	Carmelina	03/06/1997	12/11/2018	Crotone	Scienze Infermieristiche	1
PrimaSelezione							

a questa viene applicata poi in cascata la seconda selezione (quella esterna) per selezionare tutte le righe relative agli studenti nati entro il 1997:

Matricola	Cognome	Nome	DataNascita	DataIscrizione	Residenza	CorsoLaurea	AnnoCorso
8473	Baldieri	Marco	12/07/1995	02/09/2018	Catanzaro	Scienze Infermieristiche	1
8841	Ragusa	Giacomo	04/05/1995	04/10/2018	Soverato	Tecn. Prevenzione	1
8904	Petrocelli	Carmelina	03/06/1997	12/11/2018	Crotone	Scienze Infermieristiche	1
Seconda Selezione							

algebra e operatori relazionali

È ovvio che possono esserci diversi livelli di nidificazione degli operatori relazionali combinati in cascata tra loro.

Per avere ad es. un elenco con cognome, nome e città di residenza degli studenti residenti nel catanzarese (nel nostro esempio, in Catanzaro oppure Soverato), potremmo prima effettuare una selezione per Residenza = Catanzaro o Soverato, quindi una proiezione per i soli campi richiesti:

PROJECT

RESTRICT Studenti **WHERE**

Residenza = "Catanzaro" **OR** Residenza = "Soverato"

OVER

Cognome, Nome, Residenza

Proiezione esterna

Selezione interna

algebra e operatori relazionali

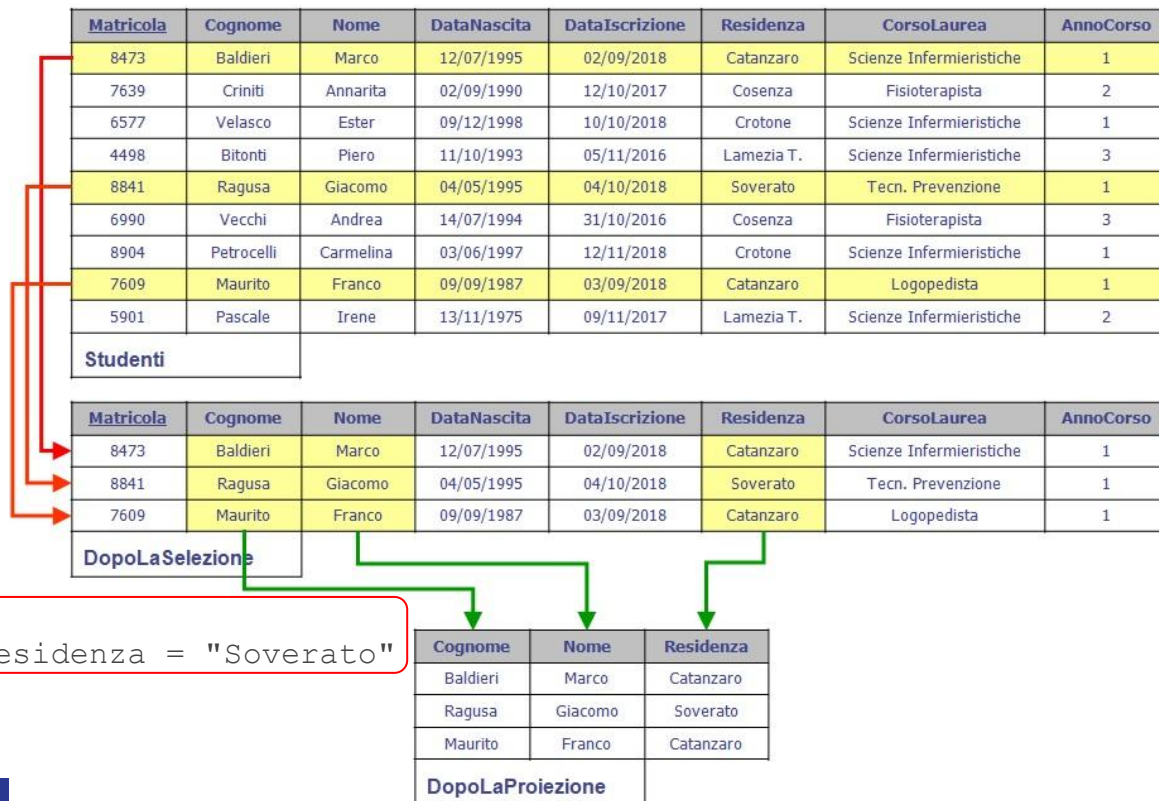
L'operazione di proiezione (PROJECT) viene applicata alla tabella risultante dall'operazione interna di selezione (RESTRICT).

PROJECT

RESTRICT Studenti **WHERE**
Residenza = "Catanzaro" **OR** Residenza = "Soverato"

OVER

Cognome, Nome, Residenza



algebra e operatori relazionali

Si può ottenere lo stesso risultato proiettando prima le colonne che ci interessano, e selezionando poi le sole righe relative agli studenti richiesti, dalla tabella risultante .

L'operazione di selezione (RESTRICT) viene applicata alla tabella risultante dall'operazione interna di selezione (PROJECT).

RESTRICT

PROJECT Studenti **OVER**
Cognome, Nome, Residenza

WHERE

Residenza = "Catanzaro" **OR**
Residenza = "Soverato"

Matricola	Cognome	Nome	DataNascita	DataIscrizione	Residenza	CorsoLaurea	AnnoCorso
8473	Baldieri	Marco	12/07/1995	02/09/2018	Catanzaro	Scienze Infermieristiche	1
7639	Criniti	Annarita	02/09/1990	12/10/2017	Cosenza	Fisioterapista	2
6577	Velasco	Ester	09/12/1998	10/10/2018	Crotone	Scienze Infermieristiche	1
4498	Bitonti	Piero	11/10/1993	05/11/2016	Lamezia T.	Scienze Infermieristiche	3
8841	Ragusa	Giacomo	04/05/1995	04/10/2018	Soverato	Tecn. Prevenzione	1
6990	Vecchi	Andrea	14/07/1994	31/10/2016	Cosenza	Fisioterapista	3
8904	Petrocelli	Carmelina	03/06/1997	12/11/2018	Crotone	Scienze Infermieristiche	1
7609	Maurito	Franco	09/09/1987	03/09/2018	Catanzaro	Logopedista	1
5901	Pascale	Irene	13/11/1975	09/11/2017	Lamezia T.	Scienze Infermieristiche	2

