

A.A. 2006 - 2007
Basi di dati 1
Algebra relazionale

Trasparenze tratte da

Basi di dati I, AA 06-07, prof. Batini
Atzeni et al. Basi di dati McGraw-Hill

Linguaggi per basi di dati

- operazioni sullo schema
 - DDL: data definition language
- operazioni sui dati
 - DML: data manipulation language
 - interrogazione ("query")
 - aggiornamento

Linguaggi di interrogazione per basi di dati relazionali

- Procedurali
 - specificano le modalità di generazione del risultato ("come")
- Dichiarativi
 - specificano le proprietà del risultato ("che cosa")

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

3

Linguaggi di interrogazione

- Algebra relazionale: procedurale
- Calcolo relazionale:
 - dichiarativo (teorico)
- SQL (Structured Query Language):
 - parzialmente dichiarativo
 - (adottato nei DBMS)
- QBE (Query by Example):
 - dichiarativo (adottato nei DBMS)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

4

Noi approfondiremo

- Algebra relazionale a lezione ed esercitazione
- SQL a lezione ed esercitazione
- QBE accennato a esercitazione, quando parleremo del sistema Access

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

5

Algebra relazionale

- Insieme di operatori
 - su relazioni
 - che producono relazioni
 - e possono essere composti

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

6

Operatori dell'algebra relazionale

- 1. Unione, Intersezione, Differenza
- 2. Ridenominazione
- 3. Selezione
- 4. Proiezione
- 5. Join (Join naturale, Prodotto cartesiano, Theta-join)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

7

Unione, intersezione, differenza

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

8

Operatori insiemistici

- le relazioni sono insiemi
- i risultati debbono essere relazioni
- è possibile applicare unione, intersezione, differenza solo a relazioni definite sugli stessi attributi

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

9

Unione: tutte le n-ple dell'una e dell'altra

Laureati

Matricola	Nome	Età
7274	Rossi	42
7432	Neri	54
9824	Verdi	45

Quadri

Matricola	Nome	Età
9297	Neri	33
7432	Neri	54
9824	Verdi	45

Laureati \cup Quadri

Matricola	Nome	Età
7274	Rossi	42
7432	Neri	54
9824	Verdi	45
9297	Neri	33

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

10

Intersezione: le n-ple comuni

Laureati

Matricola	Nome	Età
7274	Rossi	42
7432	Neri	54
9824	Verdi	45

Quadri

Matricola	Nome	Età
9297	Neri	33
7432	Neri	54
9824	Verdi	45

Laureati \cap Quadri

Matricola	Nome	Età
7432	Neri	54
9824	Verdi	45

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

11

Differenza: le n-ple della prima da cui sono tolte le n-ple della seconda

Laureati

Matricola	Nome	Età
7274	Rossi	42
7432	Neri	54
9824	Verdi	45

Quadri

Matricola	Nome	Età
9297	Neri	33
7432	Neri	54
9824	Verdi	45

Laureati – Quadri

Matricola	Nome	Età
7274	Rossi	42

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

12

Una unione sensata ma impossibile, perche'
le relazioni sono su attributi diversi

Paternità

Padre	Figlio
Adamo	Abele
Adamo	Caino
Abramo	Isacco

Maternità

Madre	Figlio
Eva	Abele
Eva	Set
Sara	Isacco

Paternità \cup Maternità

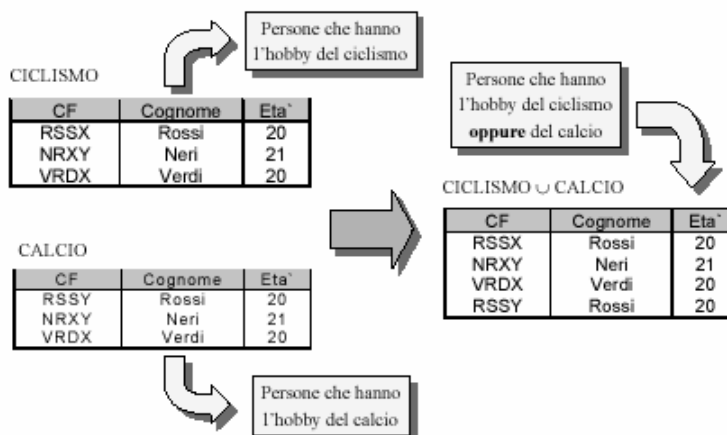
??

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

13

Unione



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

14

Differenza

Matematica

Matricola	Cognome	Eta'
7274	Rossi	20
7432	Neri	21
9824	Verdi	20

Matematica - Fisica

Matricola	Cognome	Eta'
7274	Rossi	20

Fisica

Matricola	Cognome	Eta'
9297	Rossi	20
7432	Neri	21
9824	Verdi	20

Fisica - Matematica

Matricola	Cognome	Eta'
9297	Rossi	20

Differenza non commutativa

$$R-S \neq S-R$$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

15

Differenza

CICLISMO

CF	Cognome	Eta'
RSSX	Rossi	20
NRXY	Neri	21
VRDX	Verdi	20

Persone che hanno
l'hobby del ciclismo

CICLISMO - CALCIO

CF	Cognome	Eta'
RSSX	Rossi	20

Differenza: operazione
non commutativa

CALCIO

CF	Cognome	Eta'
RSSY	Rossi	20
NRXY	Neri	21
VRDX	Verdi	20

Persone che hanno
l'hobby del calcio

CALCIO - CICLISMO

CF	Cognome	Eta'
RSSY	Rossi	20

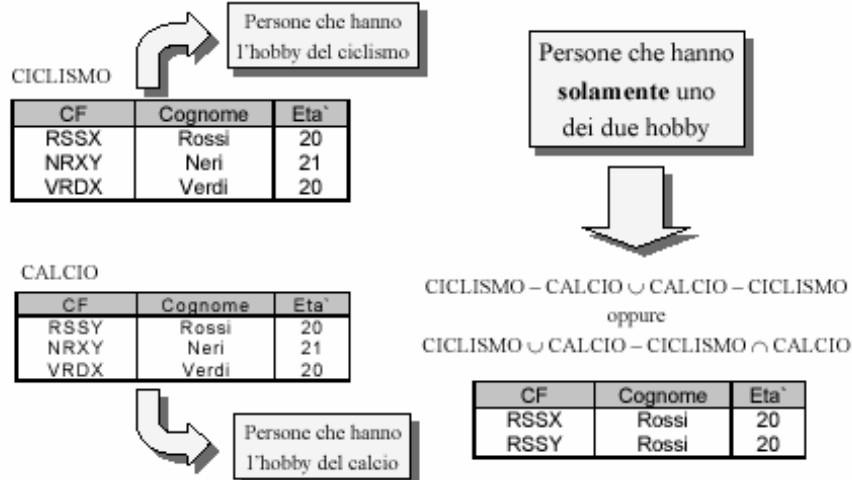
Persone che hanno
solo l'hobby del calcio

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

16

Un semplice esempio di query



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

17

Operatore di ridenominazione

- Per poter applicare operazioni insiemistiche come unione, intersezione, differenza a relazioni su attributi (in parte) diversi, e' necessario ridenominare attributi, in modo da uniformare i nomi
- Questo viene fatto dall' operatore di ridenominazione

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

18

Ridenominazione

- operatore monadico (cioe' con un argomento)
- "modifica lo schema" lasciando inalterata l'istanza dell'operando

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

19

Paternità

Padre	Figlio
Adamo	Abele
Adamo	Caino
Abramo	Isacco

$REN_{Genitore \leftarrow Padre}$ (Paternità)

Genitore	Figlio
Adamo	Abele
Adamo	Caino
Abramo	Isacco

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

20

Paternità

Padre	Figlio
Adamo	Abele
Adamo	Caino
Abramo	Isacco

REN_{Genitore ← Padre} (Paternità)

Genitore	Figlio
Adamo	Abele
Adamo	Caino
Abramo	Isacco

Maternità

Madre	Figlio
Eva	Abele
Eva	Set
Sara	Isacco

REN_{Genitore ← Madre} (Maternità)

Genitore	Figlio
Eva	Abele
Eva	Set
Sara	Isacco

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

21

REN_{Genitore ← Padre} (Paternità)

Genitore	Figlio
Adamo	Abele
Adamo	Caino
Abramo	Isacco

REN_{Genitore ← Padre} (Paternità)



REN_{Genitore ← Madre} (Maternità)

REN_{Genitore ← Madre} (Maternità)

Genitore	Figlio
Eva	Abele
Eva	Set
Sara	Isacco

Genitore	Figlio
Adamo	Abele
Adamo	Caino
Abramo	Isacco
Eva	Abele
Eva	Set
Sara	Isacco

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

22

Impiegati	Cognome	Ufficio	Stipendio
	Rossi	Roma	55
	Neri	Milano	64

Operai	Cognome	Fabbrica	Salario
	Bruni	Monza	45
	Verdi	Latina	55

REN Sede, Retribuzione ← Ufficio, Stipendio (Impiegati)

REN Sede, Retribuzione ← Fabbrica, Salario (Operai)

Cognome	Sede	Retribuzione
Rossi	Roma	55
Neri	Milano	64
Bruni	Monza	45
Verdi	Latina	55

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

23

2. Selezione

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

24

Selezione

- operatore monadico
- permette di selezionare un sottoinsieme delle ennuple
- produce un risultato che
 - ha lo stesso schema dell'operando
 - contiene un sottoinsieme delle ennuple dell'operando,
 - quelle che soddisfano una condizione espressa dall' operatore

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

25

Impiegati

Matricola	Cognome	Filiale	Stipendio
7309	Rossi	Roma	55
5998	Neri	Milano	64
9553	Milano	Milano	44
5698	Neri	Napoli	64

- impiegati che
 - guadagnano più di 50
 - guadagnano più di 50 e lavorano a Milano
 - hanno un cognome uguale al nome della filiale presso cui lavorano

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

26

Selezione: sintassi

- sintassi

SEL Condizione (Operando)

- *Condizione*: espressione booleana (come quelle dei vincoli di ennuola)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

27

Sintassi della condizione (anche detta formula proposizionale)

- Data una relazione $r(X)$, e' una formula ottenuta combinando con i connettivi OR, AND, e NOT condizioni atomiche del tipo
- $A \text{ CONFR } B$ oppure $A \text{ CONFR } c$, dove
- 1. CONFR e' un operatore di confronto ($=, >$, ecc.)
- 2. A e B sono attributi in X sui cui valori CONFR abbia senso
- 3. c e' una costante per cui il confronto CONFR sia definito

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

28

Esempi

Impiegati			
Matricola	Cognome	Filiale	Stipendio
7309	Rossi	Roma	55
5998	Neri	Milano	64
9553	Milano	Milano	44
5698	Neri	Napoli	64

- impiegati che
 - guadagnano più di 50
 - STIPENDIO > 50
 - guadagnano più di 50 e lavorano a Milano
 - STIPENDIO > 50 AND FILIALE = 'Milano'
 - hanno un cognome uguale al nome della filiale presso cui lavorano
 - COGNOME = FILIALE

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

29

Altri esempi

- ETA' > 20
- ETA' > 20 AND STIPENDIO < 40
- SESSO = "M" OR ETA' < 50

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

30

Selezione: semantica

- il risultato contiene le ennuple dell'operando che soddisfano la condizione (cioe' su cui la condizione e' vera)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

31

- impiegati che guadagnano più di 50

Impiegati

Matricola	Cognome	Filiale	Stipendio
7309	Rossi	Roma	55
5998	Neri	Milano	64
5698	Neri	Napoli	64

SEL_{Stipendio > 50} (Impiegati)

9/10/2001

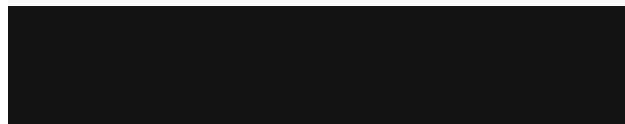
Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

32

- impiegati che guadagnano più di 50 e lavorano a Milano

Impiegati

Matricola	Cognome	Filiale	Stipendio
5998	Neri	Milano	64



SEL `Stipendio > 50 AND Filiale = 'Milano'` (Impiegati)

9/10/2001

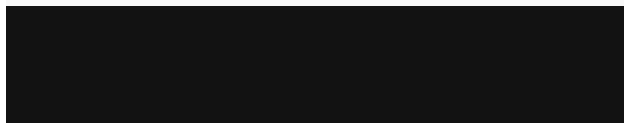
Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

33

- impiegati che hanno il cognome uguale al nome della filiale presso cui lavorano

Impiegati

Matricola	Cognome	Filiale	Stipendio
9553	Milano	Milano	44



SEL `Cognome = Filiale` (Impiegati)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

34

Selezione con valori nulli

Impiegati

Matricola	Cognome	Filiale	Età
7309	Rossi	Roma	32
5998	Neri	Milano	45
9553	Bruni	Milano	NULL

SEL $Età > 40$ (Impiegati)

- la condizione atomica è vera solo per valori non nulli

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

35

Un risultato non desiderabile

Impiegati

Matricola	Cognome	Filiale	Età
7309	Rossi	Roma	32
5998	Neri	Milano	45
9553	Bruni	Milano	NULL

$SEL_{Età > 30}(\text{Impiegati}) \cup SEL_{Età \leq 30}(\text{Impiegati})$
produce come risultato una tabella uguale o
diversa da Persone?

Diversa!

- Perché? Perché le selezioni vengono valutate separatamente!

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

36

Altro risultato non desiderabile

Impiegati

Matricola	Cognome	Filiale	Età
7309	Rossi	Roma	32
5998	Neri	Milano	45
9553	Bruni	Milano	NULL

- Ma anche

$SEL_{Età > 30 \vee Età \leq 30} (Impiegati)$
 $\neq Impiegati$

- Perché? Perché anche le condizioni atomiche vengono valutate separatamente!

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

37

$SEL_{Età > 40} (Impiegati)$

- la condizione atomica è vera solo per valori non nulli
- per riferirsi ai valori nulli esistono forme apposite di condizioni:

IS NULL
IS NOT NULL

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

38

Valutazione della selezione

- A questo punto:

$$\begin{aligned} & \text{SEL}_{\text{Età} > 30}(\text{Persone}) \cup \text{SEL}_{\text{Età} \leq 30}(\text{Persone}) \cup \\ & \quad \text{SEL}_{\text{Età IS NULL}}(\text{Persone}) \\ & = \\ & \text{SEL}_{\text{Età} > 30 \vee \text{Età} \leq 30 \vee \text{Età IS NULL}}(\text{Persone}) \\ & = \\ & \text{Persone} \end{aligned}$$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

39

**Trovare gli impiegati con piu' di 40 anni
o di cui non si sa l'eta'**

Impiegati

Matricola	Cognome	Filiale	Età
5998	Neri	Milano	45
9553	Bruni	Milano	NULL

SEL_{(Età > 40) OR (Età IS NULL)} (Impiegati)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

40

Proiezione

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

41

Impiegati

Matricola	Cognome	Filiale	Stipendio
7309	Neri	Napoli	55
5998	Neri	Milano	64
9553	Rossi	Roma	44
5698	Rossi	Roma	64

- Risolve un'altra esigenza del tipo
 - per tutti gli impiegati voglio:
 - matricola e cognome
 - cognome e filiale

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

42

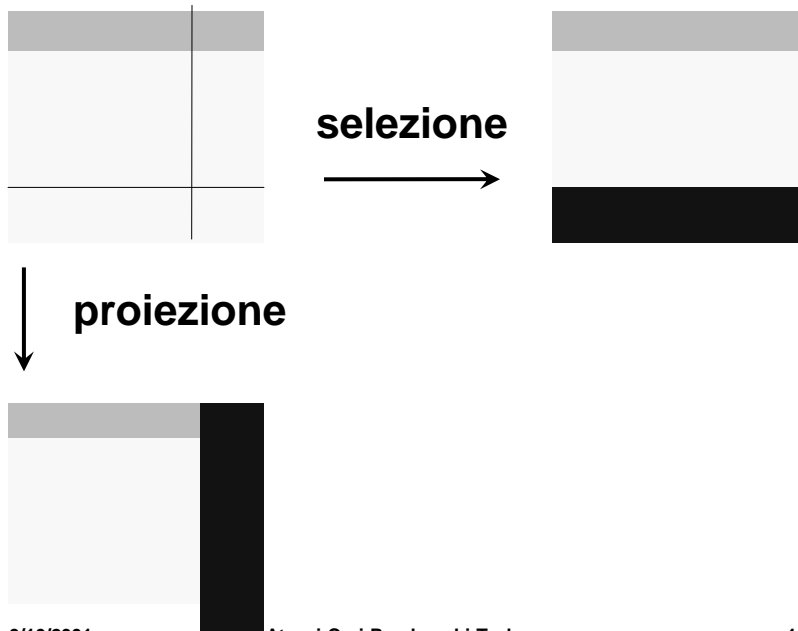
Selezione e proiezione

- operatori "ortogonali"
- selezione:
 - decomposizione orizzontale
- proiezione:
 - decomposizione verticale

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

43



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

44

Proiezione

- operatore monadico
- produce un risultato che
 - ha una parte degli attributi dell'operando
 - contiene ennuple a cui contribuiscono tutte le ennuple dell'operando

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

45

Proiezione, sintassi e semantica

- sintassi
 $PROJ_{ListaAttributi} (Operando)$
- semantica
 - il risultato contiene le ennuple dell'operando ristrette ai soli attributi nella ListaAttributi

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

46

- Trovare matricola e cognome di tutti gli impiegati

Matricola	Cognome	
7309	Neri	
5998	Neri	
9553	Rossi	
5698	Rossi	

PROJ Matricola, Cognome (Impiegati)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

47

- Trovare cognome e filiale di tutti gli impiegati

	Cognome	Filiale
	Neri	Napoli
	Neri	Milano
	Rossi	Roma

Attenzione!
Dobbiamo
eliminare le
ennuple ripetute
Una relazione e'
un insieme, e un
insieme non ha
elementi uguali

PROJ Cognome, Filiale (Impiegati)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

48

Cardinalità delle proiezioni

- La cardinalità di una relazione è il numero delle sue ennuple; si indica con $|R|$
- una proiezione
 - contiene al più tante ennuple quante l'operando
 - può contenerne di meno (vedi esempio precedente)

Impiegati

Matricola	Cognome	Filiale	Stipendio
7309	Neri	Napoli	55
5998	Neri	Milano	64
9553	Rossi	Roma	44
5698	Rossi	Roma	64

PROJ_{Cognome} (Impiegati)

Cognome
Neri
Rossi

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

49

Cardinalità delle proiezioni (π)

Studenti

Cognome	Nome	CdL	Univ
Rossi	Marco	Fisica	MI
Neri	Massimo	Fisica	FI
Viola	Mary	Matematica	TO
Blu	Sara	Matematica	TO

$\pi_{CdL, Univ} (Studenti)$

CdL	Univ
Fisica	MI
Fisica	FI
Matematica	TO

Unione delle
due tuple
uguali

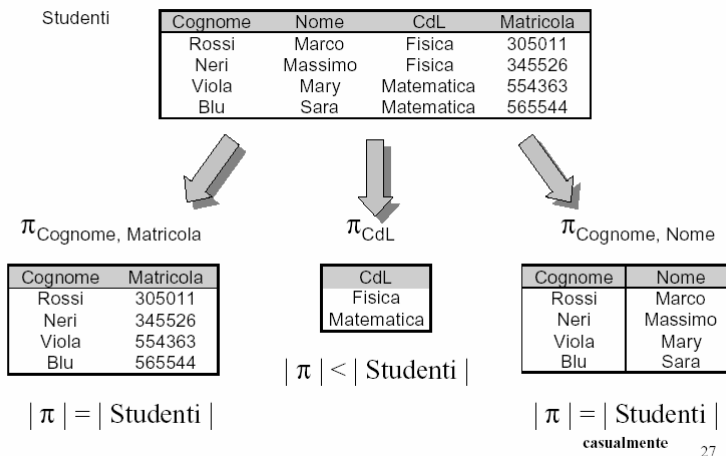
SQL si comporta diversamente

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

50

Cardinalità delle proiezioni (π)



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

51

Cardinalità delle proiezioni

- Vale la proprietà'
- se X è una superchiave di R , allora $\text{PROJ}_X(R)$ contiene esattamente tante ennuple quante R
- Per la definizione di superchiave (ogni superchiave compare una volta sola nella relazione)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

52

Selezione e proiezione

- Combinando selezione e proiezione, possiamo estrarre interessanti informazioni da una relazione

Impiegati

Matricola	Cognome	Filiale	Stipendio
7309	Neri	Napoli	55
5998	Neri	Milano	64
9553	Rossi	Roma	44
5698	Rossi	Roma	64

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

53

Selezione e proiezione

- Combinando selezione e proiezione, possiamo estrarre interessanti informazioni da una relazione

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

54

- matricola e cognome degli impiegati che guadagnano più di 50

Matricola	Cognome	
7309	Rossi	
5998	Neri	
5698	Neri	

PROJ_{Matricola,Cognome} (SEL_{Stipendio > 50} (Impiegati))

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

55

- Combinando selezione e proiezione, possiamo estrarre informazioni da una relazione
- non possiamo però correlare, mettere insieme, informazioni presenti in relazioni diverse

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

56

Join

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

57

Join

- il join è l'operatore più interessante dell'algebra relazionale
- permette di correlare, mettere insieme, integrare dati che si trovano in relazioni diverse

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

58

Prove scritte in un concorso pubblico

- I compiti sono anonimi e ad ognuno è associata una busta chiusa con il nome del candidato
- Ciascun compito e la relativa busta vengono contrassegnati con uno stesso numero


9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

59

Dati disponibili

1	25
2	13
3	27
4	28


Compiti

1	Mario Rossi
2	Nicola Russo
3	Mario Bianchi
4	Remo Neri


Buste chiuse

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

60

Schema corrispondente e Join delle due rel.

Numero	Voto
1	25
2	13
3	27
4	28

Numero	Candidato
1	Mario Rossi
2	Nicola Russo
3	Mario Bianchi
4	Remo Neri

Numero	Candidato	Voto
1	Mario Rossi	25
2	Nicola Russo	13
3	Mario Bianchi	27
4	Remo Neri	28

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

61

Join naturale

- operatore binario (generalizzabile)
- produce un risultato
 - sull'unione degli attributi degli operandi
 - con ennuple costruite ciascuna a partire da una ennupla di ognuno degli operandi

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

62

Join naturale, sintassi e semantica

- $R_1(X_1), R_2(X_2)$
- $R_1 \text{ JOIN } R_2$ è una relazione su X_1X_2
-

$\{t \text{ su } X_1X_2 \mid \text{esistono } t_1 \in R_1 \text{ e } t_2 \in R_2$
 con $t[X_1] = t_1 \text{ e } t[X_2] = t_2\}$

Quindi contribuiscono le ennuple che
 hanno gli stessi valori negli attributi
 comuni

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
 Basi di dati, Capitolo 3

63

Nello schema precedente

Numero	Voto	Numero	Candidato
1	25	1	Mario Rossi
2	13	2	Nicola Russo
3	27	3	Mario Bianchi
4	28	4	Remo Neri

Numero	Candidato	Voto
1	Mario Rossi	25
2	Nicola Russo	13
3	Mario Bianchi	27
4	Remo Neri	28

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
 Basi di dati, Capitolo 3

64

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

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

- ogni ennupla contribuisce al risultato:
- join completo

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

65

Un join non completo

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

Impiegato	Reparto	Capo
Neri	B	Mori
Bianchi	B	Mori

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

66

Un join vuoto

Impiegato	Reparto	Reparto	Capo
Rossi	A	D	Mori
Neri	B	C	Bruni
Bianchi	B		

Impiegato	Reparto	Capo
-----------	---------	------

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

67

Un join completo, con $n \times m$ ennuple

Impiegato	Reparto	Reparto	Capo
Rossi	B	B	Mori
Neri	B	B	Bruni

Impiegato	Reparto	Capo
Rossi	B	Mori
Rossi	B	Bruni
Neri	B	Mori
Neri	B	Bruni

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

68

Cardinalità del join - proprietà - 1

- Il join di R_1 e R_2 contiene un numero di ennuple compreso fra zero e il prodotto di $|R_1|$ e $|R_2|$
- $R_1(A,B)$, $R_2(B,C)$
- in generale
$$0 \leq |R_1 \text{ JOIN } R_2| \leq |R_1| \times |R_2|$$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

69

Cardinalità del join - proprietà - 2

- $R_1(A,B)$, $R_2(\underline{B},C)$
- se il join coinvolge una chiave di R_2 , allora il numero di ennuple è compreso fra zero e $|R_1|$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

70

Cardinalità del join - proprietà - 3

- $R_1(A, B)$, $R_2(B, C)$
- se B è chiave in R_2 ed esiste vincolo di integrità referenziale fra B (in R_1) e R_2 :
 $|R_1 \text{ JOIN } R_2| = |R_1|$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

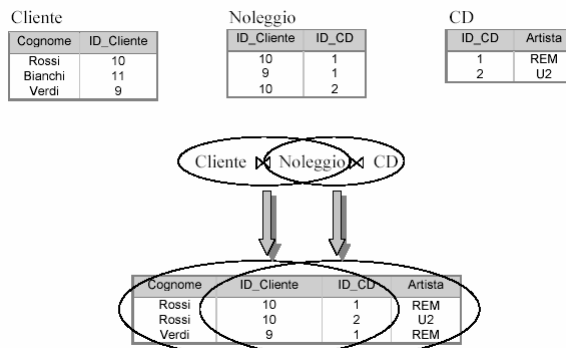
71

Proprietà del Join

- Il join è commutativo e associativo

$$r_1 \bowtie r_2 = r_2 \bowtie r_1$$

$$r_1 \bowtie (r_2 \bowtie r_3) = (r_1 \bowtie r_2) \bowtie r_3$$



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

72

Join, un problema

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

Impiegato	Reparto	Capo
Neri	B	Mori
Bianchi	B	Mori

- alcune ennuple non contribuiscono al risultato: vengono "tagliate fuori"

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

73

Join esterno

- Il join esterno estende, con valori nulli, le ennuple che verrebbero escluse da un join del tipo precedente (interno)
- esiste in tre versioni:
 - sinistro, destro, completo

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

74

Join esterno

- Sinistro → LEFT: mantiene tutte le ennuple del primo operando, estendendole con valori nulli, se necessario
- Destro → RIGHT: ... del secondo operando ...
- Completo → FULL: ... di entrambi gli operandi ...

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

75

Esempio Join esterno

Esame

Studente	Materia
Rossi	Fisica
Verdi	Fisica
Bianchi	Informatica

Corsi

Professore	Materia
Neri	Matematica
Scuri	Informatica

Esame ⋈_{LEFT} Corsi

Studente	Materia	Professore
Rossi	Fisica	NULL
Verdi	Fisica	NULL
Bianchi	Informatica	Scuri

Esame ⋈_{RIGHT} Corsi

Studente	Materia	Professore
NULL	Matematica	Neri
Bianchi	Informatica	Scuri

Esame ⋈_{FULL} Corsi

Studente	Materia	Professore
Rossi	Fisica	NULL
Verdi	Fisica	NULL
Bianchi	Informatica	Scuri
NULL	Matematica	Neri

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

76

Impiegati

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparti

Reparto	Capo
B	Mori
C	Bruni

Impiegati JOIN_{LEFT} Reparti

Impiegato	Reparto	Capo
Neri	B	Mori
Bianchi	B	Mori
Rossi	A	NULL

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

77

Impiegati

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparti

Reparto	Capo
B	Mori
C	Bruni

Impiegati JOIN_{RIGHT} Reparti

Impiegato	Reparto	Capo
Neri	B	Mori
Bianchi	B	Mori
NULL	C	Bruni

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

78

Impiegati

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparti

Reparto	Capo
B	Mori
C	Bruni

Impiegati JOIN_{FULL} Reparti

Impiegato	Reparto	Capo
Neri	B	Mori
Bianchi	B	Mori
Rossi	A	NULL
NULL	C	Bruni

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

79

Prodotto cartesiano

- un join naturale su relazioni che non hanno attributi in comune
- contiene sempre un numero di ennuple pari al prodotto delle cardinalità degli operandi (le ennuple sono tutte combinabili)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

80

Impiegati

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparti

Codice	Capo
A	Mori
B	Bruni

Impiegati JOIN 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

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

81

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

$SEL_{Condizione} (R_1 JOIN R_2)$

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

$R_1 JOIN_{Condizione} R_2$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

82

Perché "theta-join"?

- La condizione C è spesso una congiunzione (AND) di atomi di confronto $A_1 \vartheta A_2$ dove ϑ è uno degli operatori di confronto ($=, >, <, \dots$)
- se l'operatore è sempre l'uguaglianza ($=$) allora si parla di equi-join

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

83

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

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

84

Impiegati

Impiegato	Reparto
Rossi	A
Neri	B
Bianchi	B

Reparti

Reparto	Capo
A	Mori
B	Bruni

Impiegati JOIN Reparti

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

85

ESEMPIO

Codice	Data	Agente	Art	Prov	Numero
343535	12/05/99	567	44	FI	804546
343344	14/05/99	456	22	SI	345678
342233	22/06/99	456	44	FI	804546
332112	27/07/99	456	53	PI	768930
334545	07/08/99	567	44	GR	546788

INFRAZIONI

è una chiave

Prov	Numero	Proprietario	Indirizzo
FI	804546	Paolo Rossi	Via Bixio 5 Prato
SI	345678	Piero Berti	Via Abete 6 Siena
PI	768930	Luca Bianchi	Piazza Po 16 Pisa
GR	546788	Anna Verdi	Viale Europa 4 Firenze

AUTOVEICOLI

Codice	Data	Agente	Art	Prov	Numero	Proprietario	Indirizzo
343535	12/05/99	567	44	FI	804546	Paolo Rossi	Via Bixio 5 Prato
343344	14/05/99	456	22	SI	345678	Piero Berti	Via Abete 6 Siena
342233	22/06/99	456	44	FI	804546	Paolo Rossi	Via Bixio 5 Prato
332112	27/07/99	456	53	PI	768930	Luca Bianchi	Piazza Po 16 Pisa
334545	07/08/99	567	44	GR	546788	Anna Verdi	Viale Europa 4 Firenze

INFRAZIONI

AUTOVEICOLI

Il join è basato su una chiave della relazione AUTOVEICOLI, quindi ogni tupla di INFRAZIONI si combina con una sola tupla di AUTOVEICOLI

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

86

Join naturale e equijoin

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

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

87

Join naturale ed equi-join

Impiegati

Reparti

Impiegato Reparto

Reparto Capo

Impiegati JOIN Reparti

**PROJ_{Impiegato,Reparto,Capo} (SEL_{Reparto=Codice}
(Impiegati JOIN REN_{Codice ← Reparto} (Reparti)))**

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

88

Quanti join abbiamo studiato?

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

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

89

Join e ridenominazione

$$r(X) \text{ JOIN } r(X) = r(X)$$

Per eseguire il join di una relazione con se stessa in modo significativo bisogna usare la ridenominazione

ricordiamo che la ridenominazione

Modifica il nome di un sottoinsieme degli attributi di una relazione lasciando inalterato il contenuto delle relazione

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

90

Data la relazione Genitori (genitore, figlio) creare la relazione Nonni (nonno, nipote)

genitori

Genitore figlio	
Luca	Anna
Maria	Anna
Giorgio	Luca
Silvia	Maria
Enzo	Maria

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

91

Data la relazione Genitori (genitore, figlio) creare la relazione Nonni (nonno, nipote)

genitori

Genitore figlio	
Luca	Anna
Maria	Anna
Giorgio	Luca
Silvia	Maria
Enzo	Maria

REN_{nonno, genitore} <-- genitore, figlio (genitori)

Nonno genitore

Luca	Anna
Maria	Anna
Giorgio	Luca
Silvia	Maria
Enzo	Maria

REN_{nonno, genitore} <-- genitore, figlio (genitori) join genitori

Nonno Genitore figlio

Giorgio	Luca	Anna
Silvia	Maria	Anna
Enzo	Maria	Anna

Nonni= (PROJ_{nonno, nipote} REN_{nipote<-- figlio}

(Ren_{nonno, genitore} <-- genitore, figlio (genitori) join genitori))

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

92

Interrogazioni in algebra relazionale

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

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

Esempi di interrogazioni

Partiamo dallo schema e istanza qui sotto

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

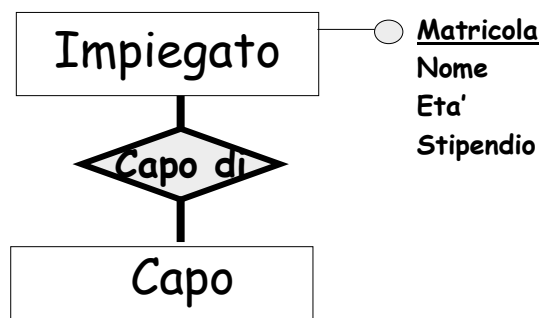
Supervisione	Impiegato	Capo
	7309	5698
	5998	5698
	9553	4076
	5698	4076
	4076	8123

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

95

Struttura concettuale dello schema:
apparentemente e'



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

96

In realta' la struttura e' piu' complessa

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

Anche i capi

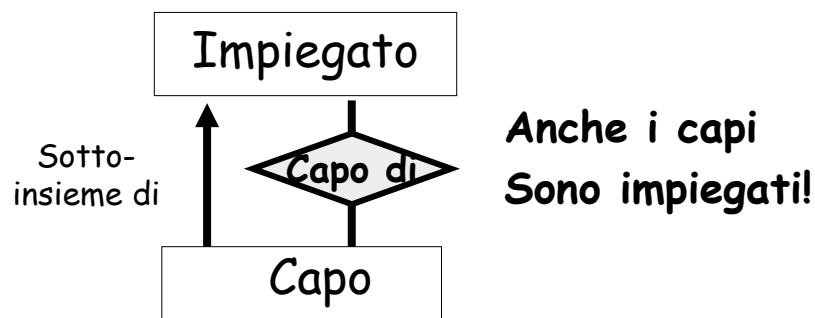
Sono impiegati!

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

97

Struttura concettuale dello schema:
occorre aggiungere una relazione IS-A



Anche i capi
Sono impiegati!

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

98

Programma

- Esprimere 6 interrogazioni, via via piu' complesse:
- 1. Trovare matricola, nome, età e stipendio degli impiegati che guadagnano più di 40 mila €
- 2. Trovare matricola, nome ed età degli impiegati che guadagnano più di 40mila €
- 3. Trovare le matricole dei capi degli impiegati che (gli impiegati!) guadagnano più di 40 mila €
- 4. Trovare nome e stipendio dei capi degli impiegati che (gli impiegati!) guadagnano più di 40 mila euro
- 5. Trovare gli impiegati che guadagnano più del proprio capo, mostrando matricola, nome e stipendio dell'impiegato e del capo
- 6. Trovare le matricole dei capi i cui impiegati guadagnano tutti più di 40 mila €

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

99

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

$SEL_{Stipendio > 40}(Impiegati)$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

100

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

SEL_{Stipendio > 40}(Impiegati)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

101

- Seconda interrogazione
- 2. Trovare matricola, nome ed età degli impiegati che guadagnano più di 40mila €
- Qui abbiamo bisogno di una proiezione per eliminare gli attributi non richiesti

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

102

Matricola	Nome	Età
7309	Rossi	34
5698	Bruni	43
4076	Mori	45
8123	Lupi	46

$$PROJ_{Matricola, Nome, Et\grave{a}}$$

$$(SEL_{Stipendio > 40}(Impiegati))$$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

103

- 3. Trovare le matricole dei capi degli impiegati che (gli impiegati!) guadagnano più di 40 mila €
- Quali sono gli attributi coinvolti?

Impiegati (Matricola, Nome, Et\grave{a}', Stipendio)
 Supervisione (Impiegato, Capo)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

104

Eseguiamo "visivamente" la interrogazione

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

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

105

- Trovare le matricole dei capi degli impiegati che (gli impiegati!) guadagnano più di 40 mila €
- Qui abbiamo bisogno di un join perche' l'interrogazione riguarda attributi di entrambe le relazioni

Impiegati (Matricola, Nome, Eta', Stipendio)
Supervisione (Impiegato, Capo)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

106

- Proviamo a individuare un problema intermedio, che poi porta alla soluzione completa →
- Trovo prima gli impiegati che guadagnano più di 40mila €

$SEL_{Stipendio > 40}(Impiegati)$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

107

Quale tabella viene selezionata?

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

$SEL_{Stipendio > 40}(Impiegati)$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

108

Soluzione

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

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

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

109

2. Poi trovo i capi cercando tali impiegati nella
relazione Supervisione
Impiegati (Matricola, Nome, Eta', Stipendio)
Supervisione (Impiegato, Capo)

Supervisione JOIN Impiegato=Matricola

(SEL_{Stipendio>40}(Impiegati))

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

110

Nuova relazione

Impiegati-Capi (Impiegato, Nome, Eta',
Stipendio, Capo)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

111

- E infine trovo le matricole dei capi

$PROJ_{Capo}$

(Supervisione JOIN $Impiegato=Matricola$

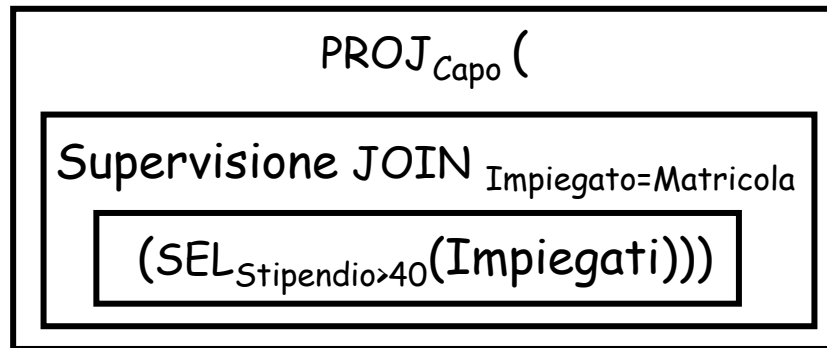
($SEL_{Stipendio>40}(Impiegati)$)))

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

112

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



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

113

Relazione risultato

CAPO
5698
4076
8123

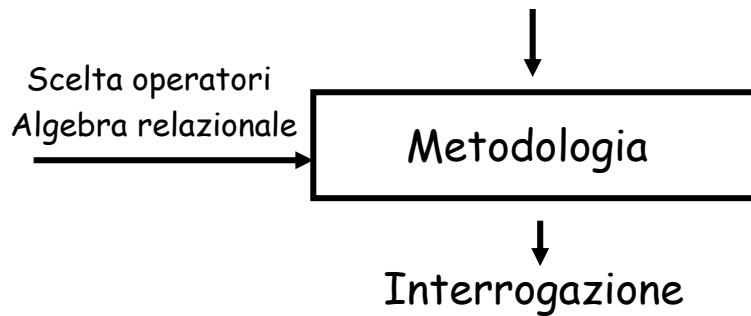
9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

114

Possiamo cominciare a definire una metodologia per costruire interrogazioni

Specifica della interrogazione
in linguaggio naturale



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

115

Metodologia

1. Individua le relazioni coinvolte nella specifica della interrogazione, attraverso gli attributi citati e le condizioni
2. Individua i tipi di operazioni necessarie
3. Individua un possibile ordinamento delle operazioni che porta ad ottenere il risultato richiesto

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

116

Per le prime tre interrogazioni

Specifica della interrogazione	Relazioni	Operazioni
1. Trovare matricola, nome, età e stipendio degli impiegati che guadagnano più di 40 mila	Impiegato	Selezione
2. Trovare matricola, nome ed età degli impiegati che guadagnano più di 40mila €	Impiegato	Selezione Proiezione
3. Trovare le matricole dei capi degli impiegati che guadagnano più di 40mila €	Impiegato Supervisione	Selezione Join Proiezione

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

117

Nuova interrogazione

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

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

118

Eseguiamo "visivamente" la interrogazione

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

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

119

Soluzione

1. Qui ci possiamo semplificare la vita utilizzando la interrogazione precedente, che produce una relazione con un attributo *CAPO*.

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

120

- Trovare nome e stipendio dei capi degli impiegati che (gli impiegati) guadagnano più di 40 mila € - Prima parte

```

Impiegati JOINMatricola=Capo
  PROJCapo(Supervisione
    JOINImpiegato=Matricola(SELStipendio>40(Impiegati))))

```

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

121

Soluzione: passo 2

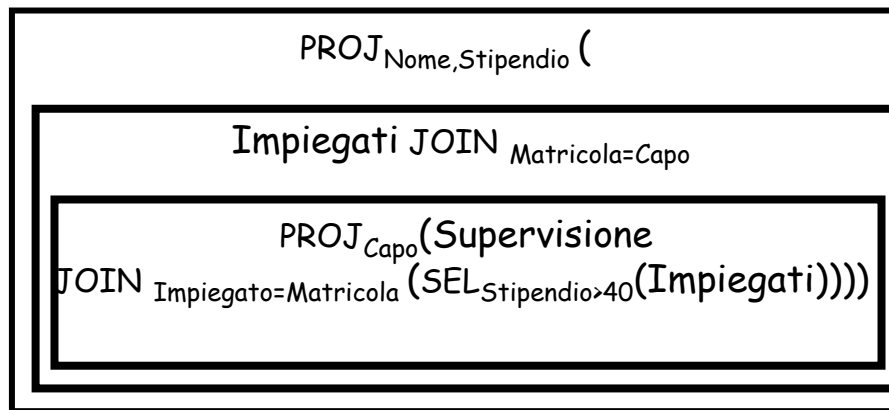
2. La relazione con attributo CAPO va messa in JOIN con la relazione Impiegato, e poi vanno estratti gli attributi NOME e STIPENDIO

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

122

- Trovare nome e stipendio dei capi degli impiegati che (gli impiegati) guadagnano più di 40 mila €



9/10/2001

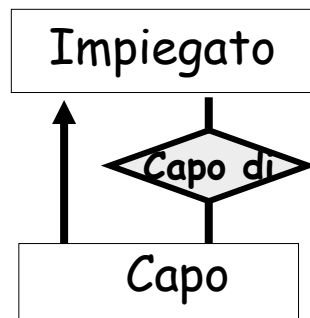
Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

123

Per le prime quattro interrogazioni

Specifica della interrogazione	Relazioni	Operazioni
1. Trovare matricola, nome, età e stipendio degli impiegati che guadagnano più di 40 mila	Impiegato	Selezione
2. Trovare matricola, nome ed età degli impiegati che guadagnano più di 40mila €	Impiegato	Selezione Proiezione
3. Trovare le matricole dei capi degli impiegati che guadagnano più di 40mila €	Impiegato Supervisione	Selezione Join Proiezione
Trovare nome e stipendio dei capi degli impiegati che <u>(gli impiegati)</u> guadagnano più di 40 mila €	Impiegato Supervisione Impiegato	Selezione Join Proiezione Join Proiezione
9/10/2001	Atzeni-Ceri-Paraboschi-Torlone, Basi di dati, Capitolo 3	124

Struttura concettuale dello schema:



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

125

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

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

126

Eseguiamo "visivamente" la interrogazione

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

In rosso

Cio' che va prodotto

In output

5998	5698
9553	4076
5698	4076
4076	8123

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

127

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

Qui il problema e' diverso, perche' nella
selezione finale abbiamo bisogno di una
relazione, risultato di JOIN, in cui
compaiono sia Impiegati che Capi, con il
nome e stipendio

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

128

Per le prime quattro interrogazioni

Specifica della interrogazione	Relazioni	Operazioni
Trovare nome e stipendio dei capi degli impiegati che <u>(gli impiegati)</u> guadagnano più di 40 mila €	Impiegato Supervisione Impiegato	Selezione Join Proiezione Join Proiezione
Trovare gli impiegati che <u>guadagnano più</u> del proprio capo, mostrando matricola, <u>nome e stipendio dell'impiegato e del capo</u>	Impiegato Supervisione Impiegato	Selezione Join Proiezione Ridenominazione Join Proiezione

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

129

Soluzione: primo passo

Abbiamo bisogno di due relazioni, una per fare la selezione e un'altra per gli impiegati, con nome e stipendio. Costruiamo:

1. una relazione che mette insieme impiegati e capi, e
2. un' altra che rinomina gli attributi degli impiegati

REN<sub>MatrC, NomeC, StipC, EtàC ←
Matr, Nome, Stip, Età (Impiegati)</sub>

(Supervisione JOIN_{Impiegato=Matricola}
Impiegati)))

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

130

Soluzione: secondo passo
ora dobbiamo 1. unirle con un join, 2. selezionare e 3.
proiettare

PROJ_{Matr, Nome, Stip, MatrC, NomeC, StipC}

(SEL_{Stipendio > StipC}**(**

REN_{MatrC, NomeC, StipC, EtàC ←}
_{Matr, Nome, Stip, Età}**(Impiegati)**

JOIN_{MatrC=Capo}

(Supervisione JOIN_{Impiegato=Matricola}
Impiegati)))

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

131

Soluzione

PROJ_{Matr, Nome, Stip, MatrC, NomeC, StipC}

(SEL_{Stipendio > StipC}**(**

REN_{MatrC, NomeC, StipC, EtàC ←}
_{Matr, Nome, Stip, Età}**(Impiegati)**

JOIN_{MatrC=Capo}

(Supervisione JOIN_{Impiegato=Matricola}
Impiegati)))

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

132

- 6. Trovare le matricole dei capi i cui impiegati guadagnano tutti più di 40 mila €
- Non si può esprimere direttamente nell'algebra (mancano i "quantificatori universali", manca l'equivalente del tutti)
- Però si può riformulare

Togliere dai capi (DIFF) quelli per i quali
almeno un impiegato guadagna meno di, oppure
40 mila euro

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
 Basi di dati, Capitolo 3

133

Proviamo a tradurre la frase che esprime la
 interrogazione in un insieme di frasi più semplici,
 che corrispondono a operatori dell'algebra
 relazionale

- 1. Trova tutti i capi, poi
- 2. Trova gli impiegati che guadagnano meno o 40 mila euro,
- 3. Poi trova i loro capi
- 4. Togli dal primo insieme il secondo

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
 Basi di dati, Capitolo 3

134

- Togliere dai capi (DIFF) quelli per i quali almeno un **impiegato guadagna meno di, oppure 40 mila euro**

1. Trova i capi

PROJ_{Capo} (Supervisione)

(SEL_{Stipendio < 40}(Impiegati)))

2. Trova gli impiegati che guadagnano meno di 40

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

135

- Togliere dai capi (DIFF) quelli per i quali almeno un **impiegato guadagna meno di, oppure 40 mila euro**

PROJ_{Capo} (Supervisione)

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

3. Trova i capi degli impiegati che Guadagnano meno o uguale a 40.000

- Togliere dai capi (DIFF) quelli per i quali almeno un **impiegato guadagna meno di, oppure 40 milioni**



4. Togli dal primo il
secondo

9/10/2001

Atzeni-
Basi di dati, Capitolo 3

Conservazione e perdita di
informazioni: il caso di proiezioni e
join nell'algebra relazionale

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

138

Join e proiezioni

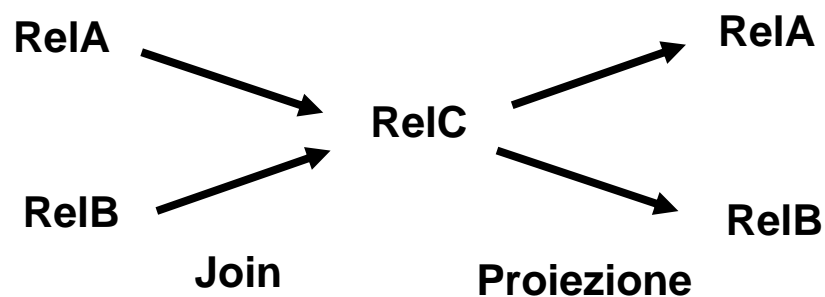
- Join e proiezioni sono operazioni complementari:
- Le proiezioni "spezzano" relazioni in frammenti, perciò separano informazioni "che stavano insieme"
- I join ricompongono frammenti in relazioni più grandi, quindi ricompongono informazioni che erano separate

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

139

Join e proiezioni: caso 1 composizione seguita da decomposizione



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

140

Join e proiezioni: caso 1

Impiegato Reparto		Reparto CapoRep	
Rossi	A	B	Mori
Neri	B	C	Bruni
Bianchi	B		

Impiegato	Reparto	CapoRep
Neri	B	Mori
Bianchi	B	Mori

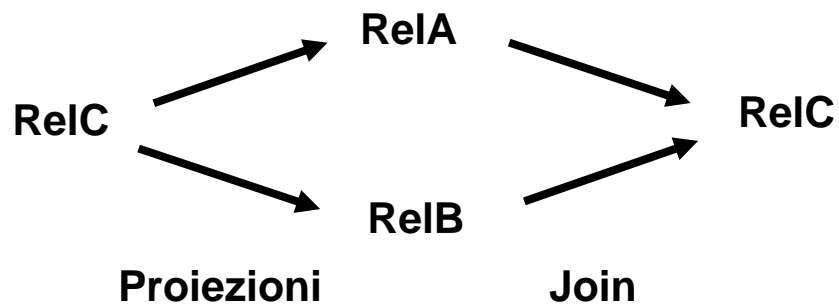
Impiegato Reparto		Reparto CapoRep	
Neri	B	B	Mori
Bianchi	B		

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

141

Join e proiezioni: caso 2 decomposizione seguita da composizione



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

142

Join e proiezioni: caso 2 - esempio 1
ogni reparto ha un solo capo

Impiegato	Reparto	CapoRep
Neri	B	Mori
Bianchi	A	Rossi

Impiegato	Reparto
Neri	B
Bianchi	A

Reparto	CapoRep
B	Mori
A	Rossi

Impiegato	Reparto	CapoRep
Neri	B	Mori
Bianchi	A	Rossi

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

143

Join e Proiezioni: caso 2 - esempio 2
i reparti possono avere piu' capi

Impiegato	Reparto	Capodilmp
Neri	B	Mori
Bianchi	B	Bruni
Verdi	A	Bini

Impiegato	Reparto
Neri	B
Bianchi	B
Verdi	A

Reparto	Capodilmp
B	Mori
B	Bruni
A	Bini

Impiegato	Reparto	Capodilmp
Neri	B	Mori
Bianchi	B	Bruni
Neri	B	Bruni
Bianchi	B	Mori
Verdi	A	Bini

9/10/2001

Basi di dati, Capitolo 3

144

Join e proiezioni: proprietà

- $R_1(X_1), R_2(X_2)$

Caso 1: $PROJ_{X_1}(R_1 \text{ JOIN } R_2) \subseteq R_1$

- $R(X), \quad X = X_1 \cup X_2$

Caso 2: $(PROJ_{X_1}(R)) \text{ JOIN } (PROJ_{X_2}(R)) \supseteq R$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

145

Quando accade nel caso 2 che
 $(PROJ_{X_1}(R)) \text{ JOIN } (PROJ_{X_2}(R)) = R$?

- Cioè, quando accade che separando (decomponendo) uno schema e poi ricomponendolo otteniamo lo stesso risultato?
- Proprietà di DECOMPOSIZIONE SENZA PERDITA:
- Se $X_0 = X_1 \cap X_2$ è chiave di R_1 o di R_2 allora decomponendo e ricomponendo otteniamo il risultato di partenza

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

146

Tornando ai due esempi del caso 2

Esempio 1

- R (Impiegato, Reparto, Caporeparto)
- Decomposta in
 - R1 (Impiegato, Reparto)
 - R2(Reparto, Caporeparto)
- REPARTO e' chiave della seconda relazione
- Vale la decomposizione senza perdita

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

147

Join e proiezioni: caso 2 - esempio 1

Impiegato	Reparto	CapoRep
Neri	B	Mori
Bianchi	A	Rossi

Impiegato	Reparto	Reparto	CapoRep
Neri	B	B	Mori
Bianchi	A	A	Rossi

Impiegato	Reparto	CapoRep
Neri	B	Mori
Bianchi	A	Rossi

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

148

Caso 2 - esempio 2

- R1 (Impiegato, Reparto, Capodiimpiegato)
- Decomposto in
 - R1 (Impiegato, Reparto)
 - R2(Reparto, Capodiimpiegato)
- REPARTO non e' chiave di nessuna delle due relazioni
- Si perdono informazioni, nel senso che si perde il legame tra impiegato e capo

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

149

Join e Proiezioni: caso 2 - esempio 2

Impiegato	Reparto	Capodilimp
Neri	B	Mori
Bianchi	B	Bruni
Verdi	A	Bini

Impiegato	Reparto	Reparto	Capodilimp
Neri	B	B	Mori
Bianchi	B	B	Bruni
Verdi	A	A	Bini

Impiegato	Reparto	Capodilimp
Neri	B	Mori
Bianchi	B	Bruni
Neri	B	Bruni
Bianchi	B	Mori
Verdi	A	Bini

9/10/2001

Basi di dati, Capitolo 3

150

Viste (relazioni derivate)

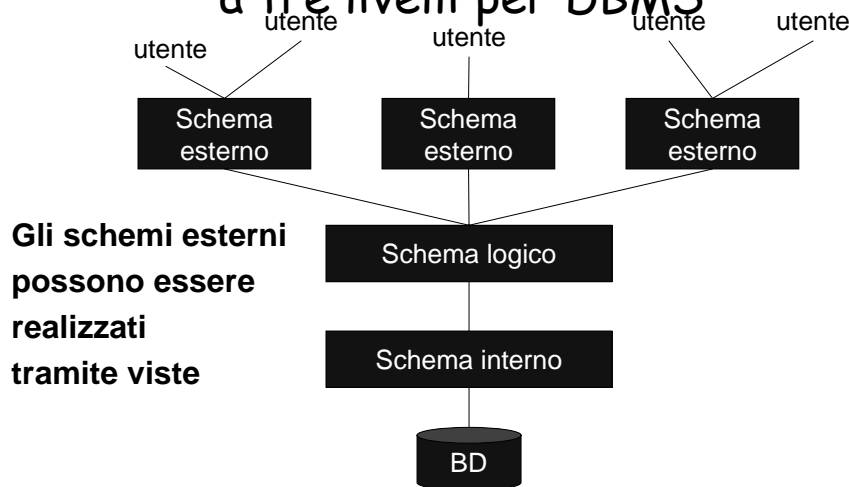
- Rappresentazioni diverse per gli stessi dati (schema esterno)
- Relazioni derivate:
 - relazioni il cui contenuto è funzione del contenuto di altre relazioni (definito per mezzo di interrogazioni)
- Relazioni di base: contenuto autonomo
- Le relazioni derivate possono essere definite su altre derivate, ma ...

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

151

Architettura standard (ANSI/SPARC) a tre livelli per DBMS



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

152

Viste virtuali e materializzate

- Gli schemi esterni possono essere realizzati tramite viste
- Due tipi di relazioni derivate:
 - viste materializzate
 - relazioni virtuali (o viste)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

153

Viste materializzate

- relazioni derivate memorizzate nella base di dati
 - vantaggi:
 - immediatamente disponibili per le interrogazioni
 - svantaggi:
 - ridondanti
 - appesantiscono gli aggiornamenti
 - non sono supportate dai DBMS

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

154

Viste virtuali

- relazioni virtuali (o viste):
 - sono supportate dai DBMS
 - una interrogazione su una vista viene eseguita "ricalcolando" la vista (o quasi)

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

155

Viste, esempio

Afferenza	Impiegato Reparto		Direzione	
			Reparto	Capo
	Rossi	A		
	Neri	B	A	Mori
	Bianchi	B	B	Bruni
	Bianchi	B	B	Bruni

- una vista:
Supervisione =
 $PROJ_{Impiegato, Capo} (Afferenza JOIN Direzione)$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

156

Interrogazioni sulle viste

- Sono eseguite sostituendo alla vista la sua definizione:

$SEL_{Capo='Leoni'}$ (Supervisione)
viene eseguita come
 $SEL_{Capo='Leoni'}($
 $PROJ_{Impiegato, Capo}$ (Afferenza JOIN
Direzione))

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

157

Viste, motivazioni

- Schema esterno: ogni utente vede solo
 - ciò che gli interessa e nel modo in cui gli interessa, senza essere distratto dal resto
 - ciò che e' autorizzato a vedere (autorizzazioni)
 - Strumento di programmazione:
 - si può semplificare la scrittura di interrogazioni: espressioni complesse e sottoespressioni ripetute
 - Utilizzo di programmi esistenti su schemi ristrutturati
- Invece:
- L'utilizzo di viste non influisce sull'efficienza delle interrogazioni

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

158

Viste come strumento di programmazione

- Trovare gli impiegati che hanno lo stesso capo di Rossi

- Senza vista:

```

PROJ Impiegato (Afferenza JOIN Direzione) JOIN
    REN ImpR,RepR ← Imp,Reparto (
    SEL Impiegato='Rossi' (Afferenza JOIN Direzione))
    
```

- Con la vista:

```

PROJ Impiegato (Supervisione) JOIN
    REN ImpR,RepR ← Imp,Reparto (
    SEL Impiegato='Rossi' (Supervisione))
    
```

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

159

Viste e aggiornamenti, attenzione

Afferenza		Direzione	
Impiegato	Reparto	Reparto	Capo
Rossi	A	A	Mori
Neri	B	B	Bruni
Verdi	A	C	Bruni

Supervisione	Impiegato	Capo
	Rossi	Mori
	Neri	Bruni
	Verdi	Mori

- Vogliamo inserire, nella vista, il fatto che Lupi ha come capo Bruni; oppure che Belli ha come capo Falchi; come facciamo?

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

160

Viste e aggiornamenti

- "Aggiornare una vista":
 - modificare le relazioni di base in modo che la vista, "ricalcolata" rispecchi l'aggiornamento
- L'aggiornamento sulle relazioni di base corrispondente a quello specificato sulla vista deve essere univoco
- In generale però non è univoco!
- Ben pochi aggiornamenti sono ammissibili sulle viste

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

161

Rappresentazione delle espressioni tramite alberi

Ogni espressione dell'algebra relazionale può essere rappresentata in modo grafico da un albero

Rappresenta l'ordine di valutazione degli operatori

Ogni operatore corrisponde ad un nodo

- operatori unari con un solo ramo in ingresso e uno in uscita
- operatori binari con due rami in ingresso e uno in uscita
- la radice è il nodo

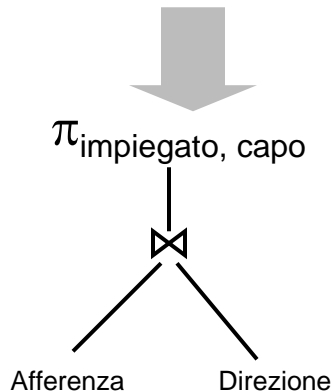
9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

162

Rappresentazione delle espressioni tramite alberi

PROJ_{impiegato, capo} (Afferenza JOIN Direzione))

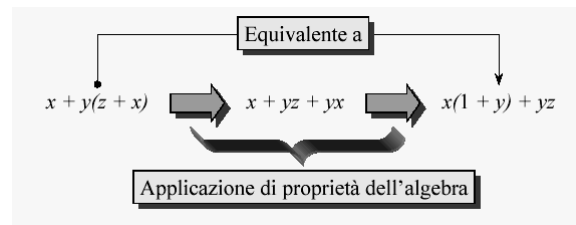


9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

163

Equivalenza di espressioni



- Due espressioni sono equivalenti se producono lo stesso risultato
 - $E_1 \equiv_R E_2$ se $E_1(r) = E_2(r)$ per ogni istanza r della basi di dati con schema R
 - $E_1 \equiv E_2$ se $E_1 \equiv_R E_2$ per ogni schema R

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

164

Equivalenza di espressioni

- Esempio di equivalenza assoluta

$$\pi_{AB}(\sigma_{A>0}(R)) \equiv \sigma_{A>0}(\pi_{AB}(R))$$

- Esempio di equivalenza dipendente dallo schema

$$\pi_{AB}(R_1) \text{ JOIN } \pi_{AC}(R_2) \equiv \pi_{ABC}(R_1 \text{ JOIN } R_2)$$

solo se nello schema **R** l'intersezione fra gli insiemi di attributi di R_1 e R_2 è pari ad A

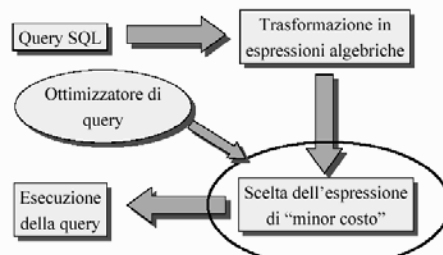
9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

165

Equivalenza di espressioni

- Espressioni equivalenti garantiscono lo stesso risultato, ma la scelta non è però indifferente in termini di "risorse" necessarie.
- Le regole più interessanti sono quelle che permettono di ridurre la dimensione dei risultati intermedi e quelle che portano a una semplificazione dell'espressione.



9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

166

Equivalenza di espressioni

STUDENTI(Nome, Cognome, Indirizzo, Matr)
ESAMI(Mat, Materia, Voto, Data)

Esami di Paolo Rossi

1. $\sigma_{\text{Nome}='Paolo' \text{ and } \text{Cognome}='Rossi'} (\text{STUDENTI} \bowtie \text{ESAMI})$
2. $\sigma_{\text{Nome}='Paolo' \text{ and } \text{Cognome}='Rossi'} (\text{STUDENTI}) \bowtie \text{ESAMI}$

Qual è la più vantaggiosa?

La seconda opera con tabelle più succinte

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

167

Equivalenza di espressioni

Esempio:

$$\pi_{AB}(\sigma_{A > 0}(R)) \equiv \sigma_{A > 0}(\pi_{AB}(R))$$

vale per ogni schema R

$$\pi_{AB}(R_1) \text{ JOIN } \pi_{AC}(R_2) \equiv_R \pi_{ABC}(R_1 \text{ JOIN } R_2),$$

sussiste solo se l'intersezione fra
gli attributi di R_1 e R_2 e' pari ad A

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

168

Equivalenza di espressioni

- Atomizzazione delle selezioni

$$\sigma_{F_1 \wedge F_2} \equiv \sigma_{F_1}(\sigma_{F_2}(E))$$

- Idempotenza delle proiezioni

$$\pi_X(E) \equiv \pi_X(\pi_{XY}(E))$$

- Anticipazione della selezione rispetto al join
(Pushing selections down)

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

se la condizione F fa riferimento solo ad attributi di E_2

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

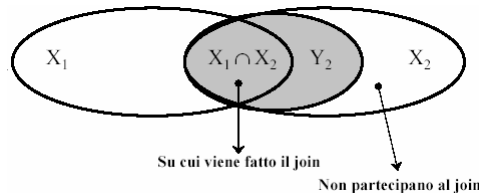
169

Equivalenza di espressioni

- Anticipazione della proiezione rispetto al join

$$\pi_{X_1 Y_2}(E_1 \bowtie E_2) \equiv E_1 \bowtie \pi_{Y_2}(E_2)$$

- E_1 è definita sugli attributi X_1 e E_2 su X_2
- $Y_2 \subseteq X_2$
- $(X_1 \cap X_2) \subseteq Y_2$ - gli attributi coinvolti nel join sono tutti in Y_2



Si possono eliminare subito gli attributi che non compaiono nella relazione finale e non sono coinvolti nel join

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

170

Equivalenza di espressioni

- Inglobamento della selezione in un prodotto cartesiano (definizione di theta-join)

$$\sigma_F(E_1 \bowtie E_2) \equiv E_1 \bowtie_F E_2$$

- Distributività della selezione

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

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

- Distributività della proiezione

$$\pi_X(E_1 \cup E_2) \equiv \pi_X(E_1) \cup \pi_X(E_2)$$

$$\pi_X(E_1 - E_2) \neq \pi_X(E_1) - \pi_X(E_2) \quad (! \text{ attenzione } \neq)$$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

171

Equivalenza di espressioni

- Distributività del join

$$E \bowtie (E_1 \cup E_2) \equiv (E \bowtie E_1) \cup (E \bowtie E_2)$$

- Selezioni con espressioni composte

$$\sigma_{F_1 \vee F_2}(R) \equiv \sigma_{F_1}(R) \cup \sigma_{F_2}(R)$$

$$\sigma_{F_1 \wedge F_2}(R) \equiv \sigma_{F_1}(R) \cap \sigma_{F_2}(R) \equiv \sigma_{F_1}(R) \bowtie_{\sigma_{F_2}}(R)$$

$$\sigma_{F_1 \wedge \neg F_2}(R) \equiv \sigma_{F_1}(R) - \sigma_{F_2}(R)$$

9/10/2001

Atzeni-Ceri-Paraboschi-Torlone,
Basi di dati, Capitolo 3

172