

Basi di Dati

prof. Letizia Tanca

**Linguaggi formali di
interrogazione per
il Modello Relazionale dei Dati**

Linguaggi di interrogazione

Ricordiamo:

- Permettono di trovare un dato basandosi **sulle sue proprietà**.
- Permettono di trovare dati basandosi su **confronti tra i contenuti** di più tabelle.

CLASSIFICAZIONE DEI LINGUAGGI

- **LINGUAGGI FORMALI**
 - Algebra relazionale
 - Calcolo relazionale
 - Delle tuple
 - Dei domini
 - Datalog
- **LINGUAGGI “COMMERCIALI”**
 - SQL (Structured Query Language)
 - QUEL
 - QBE (Query By Example)

Basi di Dati

3

CLASSIFICAZIONE DEI LINGUAGGI

- **LINGUAGGI DI DEFINIZIONE DEI DATI (DDL)**
 - per creare gli schemi dei dati e definire le loro proprietà
- **LINGUAGGI DI MANIPOLAZIONE DEI DATI (DML)**
 - per aggiornare le istanze dei dati
 - per l'interrogazione dei dati

Basi di Dati

4

ALGEBRA RELAZIONALE

E' UN LINGUAGGIO

- formale
- funzionale
- per formulare interrogazioni

BASATO SU

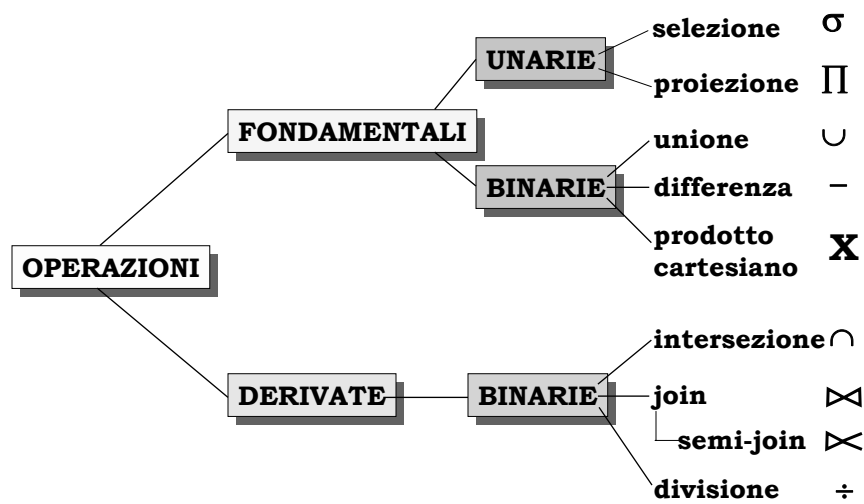
- 5 operazioni fondamentali

DEFINITO DA E. CODD NEL 1970

Basi di Dati

5

OPERAZIONI DELL'ALGEBRA RELAZIONALE



Basi di Dati

6

ESEMPIO

studente

MATR.	NOME	CITTA'	CL
123	Carlo	Bologna	Inf
415	Paola	Torino	Inf
702	Antonio	Roma	Log

corso

C-CORSO	TITOLO	DOCENTE
1	matematica	Barozzi
2	informatica	Meo

esame

MATR.	C-CORSO	DATA	VOTO
123	1	7/9/1997	30
123	2	8/1/1998	28
702	2	7/9/1997	20

Basi di Dati

7

ESEMPIO DI SELEZIONE

$\sigma_{\text{nome} = \text{"Paola"}}$ **STUDENTE**

e' una tabella con

- **SCHEMA**
lo stesso di STUDENTE (stesso grado)
- **ISTANZA**
le tuple di STUDENTE che soddisfano il predicato di selezione (cardinalita' \leq)

MATR.	NOME	CITTA'	CL
415	Paola	Torino	Inf

Basi di Dati

8

SINTASSI DELLA SELEZIONE

$\sigma_P R$

σ	OPERATORE
P	PREDICATO
R	OPERANDO

P ESPRESSIONE BOOLEANA DI PREDICATI SEMPLICI

OPERAZIONI BOOLEANE

- **AND** ($P_1 \wedge P_2$)
- **OR** ($P_1 \vee P_2$)
- **NOT** ($\neg P_1$)

PREDICATI SEMPLICI

- **TRUE, FALSE**
- **termine comparatore termine**

COMPARATORE

=, ≠, <, ≤, >, ≥

TERMINE

- **COSTANTE, ATTRIBUTO**
- **ESPRESSIONE ARITMETICA DI COSTANTI E ATTRIBUTI**

ESEMPIO DI SELEZIONE

σ (CITTÀ = "Torino") \vee ((CITTÀ = "Roma") \wedge \neg (CORSO="Log")) **studente**

studente

MATR.	NOME	CITTA'	CL
123	Carlo	Bologna	Inf
415	Paola	Torino	Inf
762	Antonio	Roma	Log

Esempio di Proiezione

$\Pi_{\text{NOME, CL}} \text{studente}$

e' una tabella con

- **SCHEMA**
solo gli attributi NOME e CL (grado \leq)
- **ISTANZA**
la restrizione delle tuple di studente sugli attributi NOME e CL (cardinalita' \leq)

NOME	CL
Carlo	Inf
Paola	Inf
Antonio	Log

Basi di Dati

11

SINTASSI DELLA PROIEZIONE

$\Pi_{\text{ATTR}_1, \dots, \text{ATTR}_n} R$

- **formalmente la proiezione elimina i duplicati**
- **nei sistemi in uso l'eliminazione dei duplicati va richiesta esplicitamente**

$\Pi_{\text{CL}} \text{STUDENTE}$

CL
Inf
Log

Basi di Dati

12

Esempio di schema di base di dati

PRESIDENTI (NOME-P, DATA-N, DATA-M, PARTITO, STATO, NOME-M)
CONGRESSI (# CONGRESSO, %S-REP, %C-REP, %S-DEM, %C-DEM)
AMMINISTRAZIONI (# AMMIN, DATA-IN, VICE-PRES, DATA-N-VP, NOME-P,
DATA-N-P)
ELEZIONI (ANNO, VOTI-PRES, NOME-P, DATA-N-PRES, NOME-PERD,
DATA-N-PERD, VOTI-PERD)
STATI (STATO, POPOLAZ, # AMMIN.)
PRESID-CONGR (NOME-P, DATA-N, # CONGR)

Interrogazioni

- *Trovare l'anno di nascita del presidente J.F. Kennedy*
- *Trovare le persone che sono state presidenti OPPURE vicepresidenti in amministrazioni inaugurate dopo il 1880*
- *Trovare le persone che sono state presidenti E ANCHE vicepresidenti in qualche amministrazione inaugurata dopo il 1880*
- *Trovare le persone che sono state presidenti MA MAI vicepresidenti in amministrazioni inaugurate dopo il 1880*

OPERAZIONI NON ALGEBRICHE

- **ASSEGNAZIONE**

Serve a dare un nome al risultato di un'operazione algebrica

- **TORINESI** = $\sigma_{(CITTÀ = \text{"Torino"})}$ **STUDENTI**
- **INFORMATICI** = $\sigma_{(CL = \text{"inf"})}$ **STUDENTI**

- **RIDENOMINAZIONE**

Permette di modificare i nomi degli attributi

- $\rho_{\text{genitore} \leftarrow \text{padre}}$ **PATERNITÀ**
- $\rho_{\text{genitore} \leftarrow \text{madre}}$ **MATERNITÀ**

in questo modo e' possibile rendere compatibili domini originariamente diversi

Basi di Dati

15

Le operazioni INSIEMISTICHE: UNIONE, DIFFERENZA e INTERSEZIONE

- $TABELLA1 \cup TABELLA2$
- $TABELLA1 - TABELLA2$
- $TABELLA1 \cap TABELLA2$

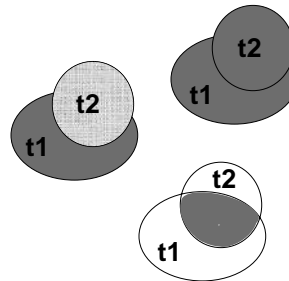


TABELLA1 e TABELLA2

devono essere *compatibili*:

- stesso grado
- domini ordinatamente dello stesso tipo
- vale: $r \cap s = r - (r - s)$ perciò l'intersezione è un operatore **DERIVATO**

Basi di Dati

16

ESEMPIO DI UNIONE

INFORMATICI \cup TORINESI

e' una tabella con

- **SCHEMA**
lo stesso di INFORMATICI
- **ISTANZA**
l'unione delle tuple di INFORMATICI e TORINESI

MATR.	NOME	CITTA'	CL
123	Carlo	Bologna	Inf
415	Paola	Torino	Inf

Basi di Dati

17

ESEMPIO DI DIFFERENZA

INFORMATICI – TORINESI

e' una tabella con

- **SCHEMA**
lo stesso di INFORMATICI
- **ISTANZA**
la differenza delle tuple di INFORMATICI e TORINESI

MATR.	NOME	CITTA'	CL
123	Carlo	Bologna	Inf

Basi di Dati

18

ESEMPIO DI INTERSEZIONE

INFORMATICI \cap TORINESI

e' una tabella con

- **SCHEMA**
lo stesso di INFORMATICI
- **ISTANZA**
l'intersezione delle tuple di INFORMATICI e
TORINESI

MATR.	NOME	CITTA'	CL
415	Paola	Torino	Inf

Basi di Dati

19

PRODOTTO CARTESIANO (*quasi un'operazione insiemistica*)

$R \times S$

e' una tabella con

- **SCHEMA**
tutti gli attributi di R e di S
(grado ($R \times S$) = grado (R) + grado (S))
- **ISTANZA**
tutte le possibili coppie di tuple di R e di S
(card ($R \times S$) = card (R) * card (S))

Basi di Dati

20

ESPRESSIONI ALGEBRICHE

- Concatenazione di piu' operazioni algebriche
- Esprimono interrogazioni in modo formale
- Consentono di estrarre informazioni dalla base di dati

UN ALTRO ESEMPIO DI BASE DI DATI

- **CORRENTISTA (NOMECORR, VIA, CITTA')**
- **CLIENTE (CORRENTISTA, IMPIEGATO)**
- **FILIALE (NOMEFIL, PATRIMONIO, CITTA')**
- **DEPOSITI (FILIALE, NUM-DEP, CORRENTISTA, SALDO)**
- **PRESTITI (FILIALE, NUM-PRES, CORRENTISTA, QUANTITA')**

SINTASSI DEL JOIN

$$R_1 \bowtie_P R_2$$

- E' equivalente alla seguente espressione

$$\sigma_P (R_1 \times R_2)$$
- **SINTASSI DEL PREDICATO DI JOIN**
 - espressione congiuntiva di predicati semplici
 - $ATTR_1 \text{ COMP } ATTR_2$
 - $ATTR_1 \in r_1$
 - $ATTR_2 \in r_2$
 - $COMP =, \neq, <, \leq, >, \geq$
 - attributi omonimi sono resi non ambigui usando la notazione puntata: $R_j.ATTR_j$

Basi di Dati

25

EQUI-JOIN E JOIN NATURALE

- **EQUI-JOIN**
 - il predicato ammette solo confronti di uguaglianza
- **JOIN NATURALE**
 - equi-join di tutti gli attributi omonimi
 - si omette il predicato
 - si elimina la colonna ripetuta

STUDENTE \bowtie **ESAME**

MATR.	NOME	CITTA'	CL	C-CORSO	DATA	VOTO
123	Carlo	Bologna	Inf	1	7/9/1997	30
123	Carlo	Bologna	Inf	2	8/1/1998	28
702	Antonio	Roma	Log	2	7/9/1997	20

Basi di Dati

26

SINTASSI DEL JOIN NATURALE

$$S \bowtie R = \Pi_{R \cup S} (R \bowtie_{R.A1=S.A1 \wedge, \dots, \wedge R.AN=S.AN} S)$$

- $\{A1, \dots, An\} = R \cap S$
- $R \cap S = \Phi \Rightarrow \bowtie \equiv \times$
- \bowtie E' ASSOCIATIVA

Basi di Dati

27

JOIN NATURALE DI TRE TABELLE

studente \bowtie esame \bowtie corso

MATR.	NOME	CITTA'	CL	C-CORSO	DATA	VOTO	TITOLO	DOCENTE
123	Carlo	Bologna	Inf	1	7/9/1997	30	matematica	Barozzi
123	Carlo	Bologna	Inf	2	8/1/1998	28	informatica	Meo
702	Antonio	Roma	Log	2	7/9/1997	20	informatica	Meo

Basi di Dati

28

SEMI-JOIN

STUDENTE \bowtie ESAME
STUDENTE.MATR=ESAME.MATR

- e' equivalente alla seguente espressione

$\Pi_{\text{STUDENTE.*}} (\text{STUDENTE } \bowtie \text{ ESAME})$

- produce un sottoinsieme di STUDENTE:
solo gli studenti che hanno dato almeno un esame

studente

MATR.	NOME	CITTA'	CL
123	Carlo	Bologna	Inf
702	Antonio	Roma	Log

Basi di Dati

29

SEMI-JOIN NATURALE

STUDENTE \bowtie ESAME

- E' EQUIVALENTE ALLA SEGUENTE ESPRESSIONE

$\Pi_{\text{STUDENTE.*}} (\text{STUDENTE } \bowtie \text{ ESAME})$

- in questo esempio il risultato e' uguale al caso precedente

studente

MATR.	NOME	CITTA'	CL
123	Carlo	Bologna	Inf
702	Antonio	Roma	Log

Basi di Dati

30

DIVISIONE

$$\mathbf{R \div S}$$

- e' equivalente alla seguente espressione

$$\Pi_{R-S} \mathbf{R} - \Pi_{R-S} ((\Pi_{R-S} \mathbf{R} \times \mathbf{S}) - \mathbf{R})$$

con $\text{Schema}(\mathbf{S}) \subseteq \text{Schema}(\mathbf{R})$

- e' utile per esprimere interrogazioni che contengono il quantificatore universale \forall : *seleziona le tuple di R che contengono TUTTE le tuple di S*

r			
X	Y	W	Z
a	b	c	d
a	b	e	f
b	c	e	f
e	d	c	d
e	d	e	f
a	b	d	e

\div

S	
W	Z
c	d
e	f

=

X	Y
a	b
e	d

Basi di Dati

31

ESEMPIO DI USO DELLA DIVISIONE

TROVARE I NOMI DEI CORRENTISTI CHE HANNO UN DEPOSITO IN TUTTE LE FILIALI DI Milano

- TUTTE LE FILIALI DI Milano

$$r_1 = \Pi_{\text{NOME}} (\sigma_{\text{CITTA}='Milano'} \text{ filiale})$$

- I CLIENTI DI CIASCUNA FILIALE

$$r_2 = \Pi_{\text{FILIALE}, \text{CORRENTISTA}} \text{ depositi}$$

- I CLIENTI DI r_2 CHE SONO APPAIATI CON TUTTE LE FILIALI DI r_1

$$r_3 = r_2 \div r_1$$

Basi di Dati

32

PRECEDENZA DEGLI OPERATORI

- La precedenza funziona circa come in algebra elementare
- gli operatori unari (come il meno unario in algebra) hanno priorità massima
- il join e il prodotto cartesiano sono “come il prodotto”
- per il resto meglio mettere le parentesi

ESEMPIO DI ESPRESSIONE COMPLESSA

ESTRARRE I NOMI DEGLI STUDENTI CHE HANNO SOSTENUTO GLI ESAMI DI *informatica* E *matematica* LO STESSO GIORNO

$$\Pi_{\text{NOME}} (\text{STUDENTE} \bowtie (\text{ESAME} \bowtie \sigma_{\text{TITOLO}=\text{"INFORMATICA"}} \text{CORSO}) \bowtie_{\text{MATR}=\text{MATRMAT} \wedge \text{DATA}=\text{DATAMAT}} \rho_{\text{MATRMAT.DATAMAT} \leftarrow \text{MATR.DATA}} \Pi_{\text{MATR.DATA}} (\text{ESAME} \bowtie \sigma_{\text{TITOLO}=\text{"MATEMATICA"}} \text{CORSO}))$$

Esempio di schema di base di dati

PRESIDENTI (NOME-P, DATA-N, DATA-M, PARTITO, STATO, NOME-M)

CONGRESSI (# CONGRESSO, %S-REP, %C-REP, %S-DEM, %C-DEM)

AMMINISTRAZIONI (# AMMIN, DATA-IN, VICE-PRES, DATA-N-VP, NOME-P, DATA-N-P)

ELEZIONI (ANNO, VOTI-PRES, NOME-P, DATA-N-PRES, NOME-PERD, DATA-N-PERD, VOTI-PERD)

STATI (STATO, POPOLAZ, # AMMIN.)

PRESID-CONGR (NOME-P, DATA-N, # CONGR)

Basi di Dati

35

Interrogazioni

- *Trovare l'anno di nascita del presidente J.F. Kennedy*
- *Trovare gli anni in cui è stato eletto un presidente repubblicano proveniente dall'Illinois*
- *Trovare i numeri di congressi presieduti dal presidente eletto nel 1955*
- *Trovare i perdenti delle elezioni vinte da qualche presidente di nome Roosevelt*
- *Trovare i nomi delle mogli dei presidenti provenienti dalla California eletti dopo il 1960*
- *Trovare le persone che sono state presidenti OPPURE vicepresidenti in amministrazioni inaugurate dopo il 1880*
- *Trovare le persone che sono state presidenti E ANCHE vicepresidenti in qualche amministrazione inaugurata dopo il 1880*
- *Trovare le persone che sono state presidenti MA MAI vicepresidenti in amministrazioni inaugurate dopo il 1880*

Basi di Dati

36

RAPPRESENTAZIONE DELLE ESPRESSIONI TRAMITE ALBERI

Ogni espressione dell'algebra
relazionale può essere rappresentata
mediante un albero sintattico che
esprime l'ordine di valutazione degli
operatori

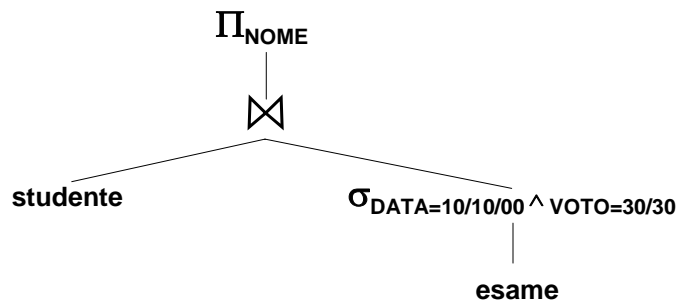
Basi di Dati

37

ESEMPIO DI ALBERO

*Nomi degli studenti che hanno sostenuto un
esame il 10/10/00 con voto pari a 30/30*

$\Pi_{\text{NOME}} (\text{STUDENTE} \bowtie \sigma_{\text{DATA}=10/10/00 \wedge \text{VOTO}=30/30} \text{ESAME})$

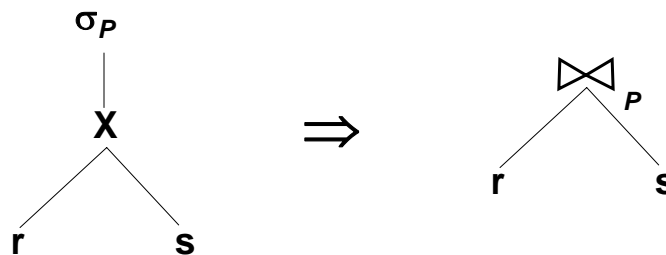


Basi di Dati

38

EQUIVALENZA DI ESPRESSIONI ALGEBRICHE

1 - ELIMINAZIONE DEI PRODOTTI CARTESIANI

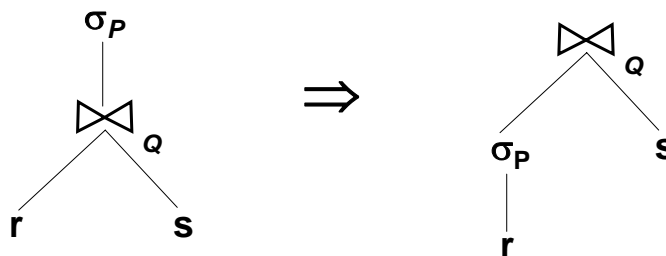


Basi di Dati

39

EQUIVALENZA DI ESPRESSIONI ALGEBRICHE

2 - PUSH DELLA SELEZIONE RISPETTO AL JOIN



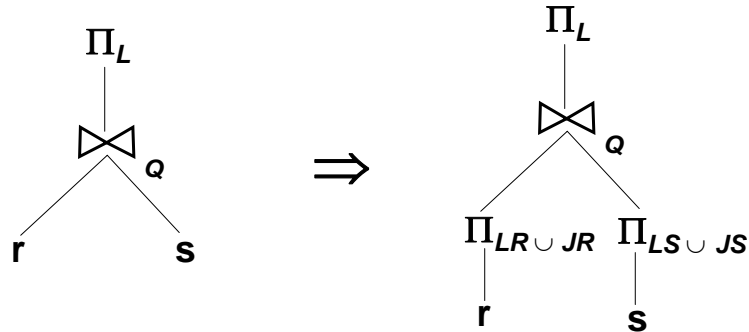
**VALE SE P SI APPLICA AI SOLI ATTRIBUTI DI R:
altrimenti possiamo suddividere P nei suoi "congiunti"**

Basi di Dati

40

EQUIVALENZA DI ESPRESSIONI ALGEBRICHE

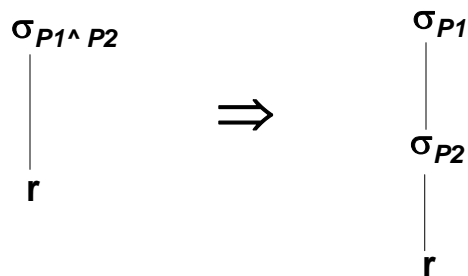
3 - PUSH DELLA PROIEZIONE RISPETTO AL JOIN



- $J = JR \cup JS$ SONO GLI ATTRIBUTI "COINVOLTI NEL JOIN"
- $LR = L - \text{SCHEMA}(s)$
- $LS = L - \text{SCHEMA}(r)$

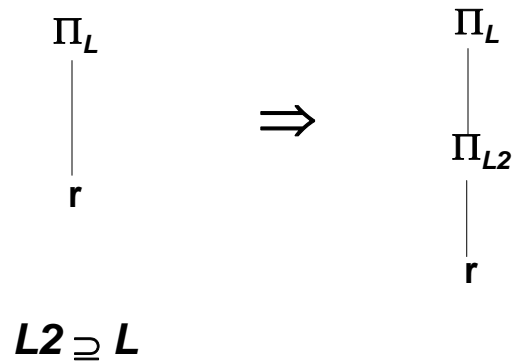
EQUIVALENZA DI ESPRESSIONI ALGEBRICHE

4 - ATOMIZZAZIONE DELLE SELEZIONI



EQUIVALENZA DI ESPRESSIONI ALGEBRICHE

5 - IDEMPOTENZA DELLA PROIEZIONE

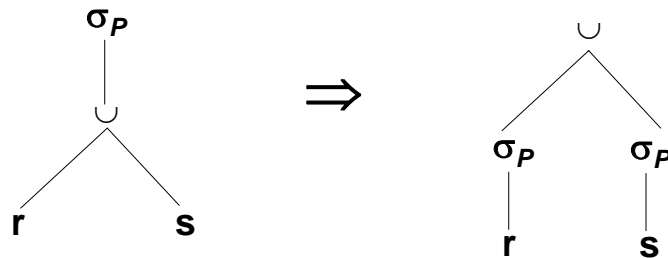


Basi di Dati

43

EQUIVALENZA DI ESPRESSIONI ALGEBRICHE

6 - PUSH DELLA SELEZIONE RISPETTO ALL'UNIONE

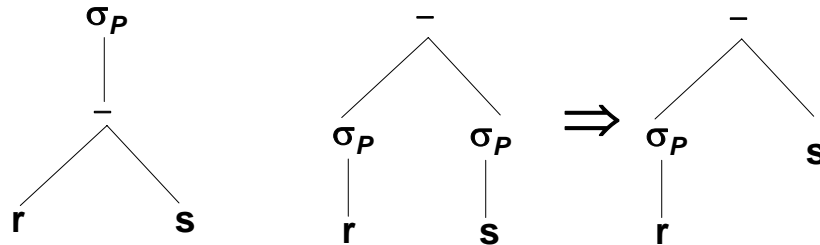


Basi di Dati

44

EQUIVALENZA DI ESPRESSIONI ALGEBRICHE

7 - PUSH DELLA SELEZIONE RISPETTO ALLA DIFFERENZA



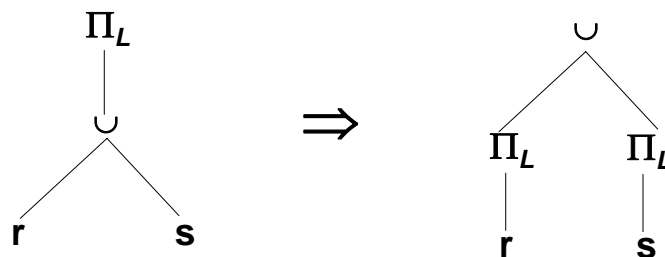
Nota: la selezione commuta con tutte e tre le operazioni insiemistiche

Basi di Dati

45

EQUIVALENZA DI ESPRESSIONI ALGEBRICHE

8 - PUSH DELLA PROIEZIONE RISPETTO ALL'UNIONE



Nota: la proiezione NON commuta con le altre operazioni insiemistiche

Basi di Dati

46

EQUIVALENZA DI ESPRESSIONI ALGEBRICHE

La proiezione NON commuta con le altre operazioni insiemistiche

- **Differenza:** $R_1(A, B), R_2(A, B)$
 $\Pi_A(R_1 - R_2)$ è diverso da $\Pi_A R_1 - \Pi_A R_2$
 (v. slide successiva)

- **Intersezione:** ovviamente non commuta *perché si deriva da unione e differenza*

La proiezione non commuta con la differenza

Clients banca

CORRENTISTA	IMPIEGATO
Rossi Brambilla Verdi	Ghezzi Ferrari Ferrari

$\Pi_{\text{correntista}}(\text{Clienti Banca} - \text{Clienti Filiale B}) = \{(Rossi), (Verdi)\}$

Clients filiale B

CORRENTISTA	IMPIEGATO
Rossi Brambilla Verdi	Amato Ferrari Lupo

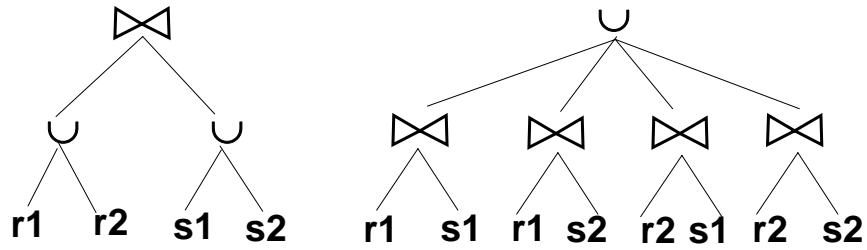
$\Pi_{\text{correntista}} \text{ Clienti Banca} -$

$\Pi_{\text{correntista}} \text{ Clienti Filiale B} = \emptyset$

Il motivo è che l'attributo correntista *non è chiave* di Clienti banca e di Clienti Filiale B. Possiamo affermare:
*se la proiezione include le chiavi delle due relazioni,
 la proiezione commuta con la differenza*

EQUIVALENZA DI ESPRESSIONI ALGEBRICHE

9 - distributività di JOIN rispetto a UNIONE



Basi di Dati

49

FORMULE UTILI

$$r \bowtie r = r$$

$$r \cup r = r$$

$$r - r = \emptyset$$

$$r \bowtie \sigma_P r = \sigma_P r$$

$$r \cup \sigma_P r = r$$

$$r - \sigma_P r = \sigma_{\neg P} r$$

$$\sigma_{P_1} r \bowtie \sigma_{P_2} r = \sigma_{P_1 \wedge P_2} r$$

$$\sigma_{P_1} r \cup \sigma_{P_2} r = \sigma_{P_1 \vee P_2} r$$

$$\sigma_{P_1} r - \sigma_{P_2} r = \sigma_{P_1 \wedge \neg P_2} r$$

$$\sigma_P \emptyset = \emptyset$$

$$\Pi_L \emptyset = \emptyset$$

$$r \times \emptyset = \emptyset$$

$$r \cup \emptyset = r$$

$$r - \emptyset = r$$

$$\emptyset - r = \emptyset$$

$$r \bowtie \emptyset = \emptyset$$

$$r \cap \emptyset = \emptyset$$

Basi di Dati

50

OTTIMIZZAZIONE ALGEBRICA

Tra tutte le rappresentazioni equivalenti conviene scegliere quella meno costosa da eseguire

- Minimizzare la dimensione dei risultati intermedi
 - Utilizzare, dove possibile, le trasformazioni di push (2,3,6,7,8)
 - Usare le trasformazioni di atomizzazione ed idempotenza (4, 5) per generare nuove selezioni e proiezioni
- Cercare di combinare i prodotti cartesiani con le selezioni i cui predicati appartengono ad entrambi gli operandi (diventano join)

ESEMPIO DI OTTIMIZZAZIONE

DATO LO SCHEMA

$R(A,B,C)$

$S(C,D,E)$

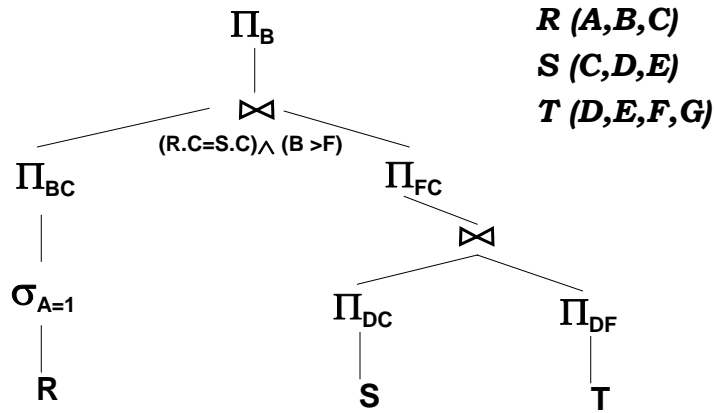
$T(D,E,F,G)$

SI OTTIMIZZAZI L'ESPRESSIONE

$\Pi_B \sigma_{(R.C=S.C) \wedge (S.D=T.D) \wedge (R.A=1) \wedge (R.B>T.F)} (R \times S \times T)$

ESEMPIO DI OTTIMIZZAZIONE

$\Pi_B \sigma_{(R.C=S.C) \wedge (S.D=T.D) \wedge (R.A=1) \wedge (R.B>T.F)} (R \times S \times T)$



Basi di Dati

53

ESERCIZI

- Disegnare gli alberi delle interrogazioni sui presidenti
- Ottimizzarli

Basi di Dati

54

Esempio di schema di base di dati

PRESIDENTI (NOME-P, DATA-N, DATA-M, PARTITO, STATO, NOME-M)

CONGRESSI (# CONGRESSO, %S-REP, %C-REP, %S-DEM, %C-DEM)

AMMINISTRAZIONI (# AMMIN, DATA-IN, VICE-PRES, DATA-N-VP, NOME-P, DATA-N-P)

ELEZIONI (ANNO, VOTI-PRES, NOME-P, DATA-N-PRES, NOME-PERD, DATA-N-PERD, VOTI-PERD)

STATI (STATO, POPOLAZ, # AMMIN.)

PRESID-CONGR (NOME-P, DATA-N, # CONGR)

Interrogazioni

- *Trovare l'anno di nascita del presidente J.F. Kennedy*
- *Trovare gli anni in cui è stato eletto un presidente repubblicano proveniente dall'Illinois*
- *Trovare i numeri di congressi presieduti dal presidente eletto nel 1955*
- *Trovare i perdenti delle elezioni vinte da qualche presidente di nome Roosevelt*
- *Trovare i nomi delle mogli dei presidenti provenienti dalla California eletti dopo il 1960*
- *Trovare le persone che sono state presidenti OPPURE vicepresidenti in amministrazioni inaugurate dopo il 1880*
- *Trovare le persone che sono state presidenti E ANCHE vicepresidenti in qualche amministrazione inaugurata dopo il 1880*
- *Trovare le persone che sono state presidenti MA MAI vicepresidenti in amministrazioni inaugurate dopo il 1880*

DIMENSIONI DEI RISULTATI INTERMEDI

Valutare in generale le dimensioni dei risultati intermedi e' operazione complessa

Si usano metodi approssimati facendo ricorso ad operazioni statistiche

Si suppongono:

- uniforme la distribuzione dei valori nei domini
- non correlati i dati relativi ai predicati nelle interrogazioni

DIMENSIONI DEI RISULTATI INTERMEDI

PARAMETRI QUANTITATIVI

- **CARD(r_i)** NUMERO DI TUPLE DI r_i
- **SIZE (t_i)** DIMENSIONE (in byte) DELLA TUPLA DI r_i
- **SIZE (A_j)** DIMENSIONE (in byte) DELL'ATTRIBUTO A_j DI r_i
- **VAL (A_j)** NUMERO DI VALORI DISTINTI DELL'ATTRIBUTO A_j DI r_i
- **$S_i = 1/VAL(A_j)$** LA SELETTIVITA' DEL DOMINIO SUL QUALE E' DEFINITO L'ATTRIBUTO A_j SE I VALORI SONO UNIFORM. DISTRIBUITI

DIMENSIONI DEI RISULTATI INTERMEDI

OPERAZIONE DI SELEZIONE $\sigma_P R$

$$\text{CARD}(\text{risultato}) = S_p * \text{CARD}(R)$$

S_p = SELETTIVITA' DEL PREDICATO

- $P = A \wedge B$ $S_p = S_A * S_B$
- $P = A \vee B$ $S_p = S_A + S_B - S_A * S_B$
- $P = \neg A$ $S_p = 1 - S_A$

$\text{SIZE}(\text{risultato}) = \text{SIZE}(R)$, cioè la tupla del risultato ha la stessa dimensione della tupla originaria

Inoltre, nel risultato sarà $\text{VAL}(A_i) = 1$ se $P \equiv "A_i = v"$