

Lezione 6

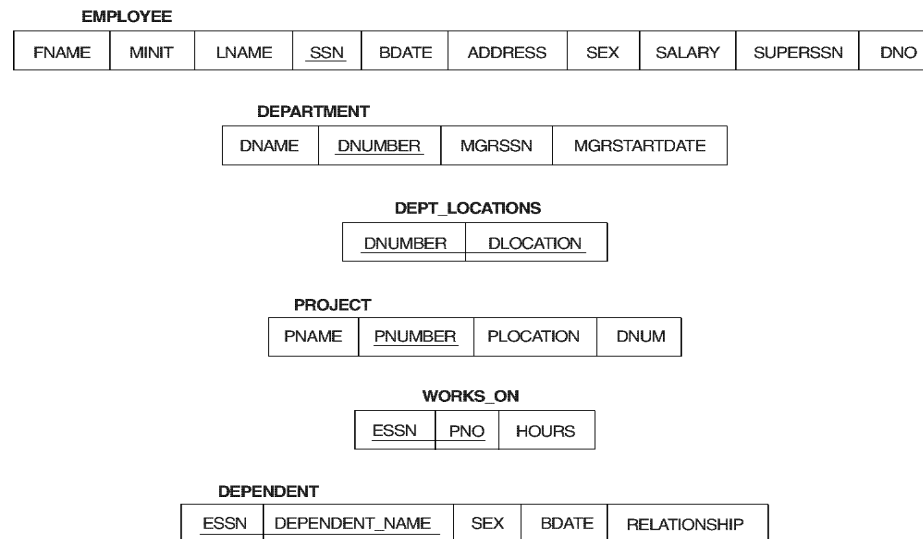
Algebra e Calcolo Relazionale

Sommario

- Esempio di Applicazione con Database (AZIENDA)
- **Algebra Relazionale**
 - Operazioni Relazionali Unarie
 - Operazioni dell'Algebra Relazionale dalla teoria degli insiemi
 - Operazioni Relazionali Binarie
 - Altre Operazioni Relazionali
 - Esempi di Interrogazioni (Query) in Algebra Relazionale
- **Calcolo Relazionale**
 - Calcolo Relazionale su Tuple
 - Calcolo Relazionale su Domini

Stato di Database per AZIENDA

Figure 7.5 Schema diagram for the COMPANY relational database schema; the primary keys are underlined.



© Addison Wesley Longman, Inc. 2000, Elmasri/Navathe, Fundamentals of Database Systems, Third Edition

Tutti gli esempi faranno riferimento a questo Schema di Database

Algebra Relazionale

- L'insieme principale delle operazioni per il modello relazionale è l'**algebra relazionale**. Le operazioni dell'algebra consentono all'utente di specificare le interrogazioni fondamentali.
- Il risultato di un'interrogazione è una **nuova relazione**, che può essere stata formata a partire da una o più relazioni. Le operazioni dell'algebra, quindi, producono nuove relazioni, che possono essere ulteriormente manipolate usando le operazioni della stessa algebra.
- Una sequenza di operazioni dell'algebra relazionale forma un'**espressione dell'algebra relazionale**, il cui risultato sarà ancora una relazione che rappresenta il risultato di un'interrogazione del database.

Operazioni Relazionali Unarie

Operazione di selezione (SELECT)

L'operazione di selezione è usata per selezionare un *sottoinsieme* di tuple di una relazione che soddisfano una condizione di selezione. È una sorta di filtro che trattiene solo quelle tuple che soddisfano una *condizione qualificante*. Può essere vista come una *partizione orizzontale* della relazione in due insiemi di tuple: quelle che soddisfano la condizione e vengono selezionate e quelle che non la soddisfano e vengono scartate.

Esempio: Per selezionare le tuple di DIPENDENTE il cui dipartimento è 4 oppure quelle il cui stipendio supera i 30000 USD, si usa la notazione:

$$\sigma_{DNO = 4}(\text{EMPLOYEE}), \sigma_{SALARY > 30000}(\text{EMPLOYEE})$$

In generale, l'operazione di selezione è indicata con:

$$\sigma_{\langle \text{condizione di selezione} \rangle}(\mathbf{R})$$

dove il simbolo σ (sigma) è usato per denotare l'operatore di selezione e la condizione di selezione è un'espressione *booleana* specificata sugli attributi della relazione R.

Operazioni Relazionali Unarie

Proprietà dell'operazione SELECT

- L'operazione SELECT $\sigma_{\langle \text{condizione di selezione} \rangle}(R)$ produce una relazione S che ha lo **stesso** schema della relazione R .
- L'operazione SELECT è **commutativa**:

$$\sigma_{\langle \text{cond1} \rangle}(\sigma_{\langle \text{cond2} \rangle}(R)) = \sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond1} \rangle}(R))$$

- Operazioni di SELECT in **cascata** possono quindi essere eseguite in **qualunque ordine**:

$$\sigma_{\langle \text{cond1} \rangle}(\sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond3} \rangle}(R))) = \sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond3} \rangle}(\sigma_{\langle \text{cond1} \rangle}(R)))$$

- Operazioni di SELECT in cascata possono essere sostituite da una singola operazione di SELECT avente come condizione la **coniunzione** di tutte le condizioni precedenti:

$$\sigma_{\langle \text{cond1} \rangle}(\sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond3} \rangle}(R))) = \sigma_{\langle \text{cond1} \rangle \text{ AND } \langle \text{cond2} \rangle \text{ AND } \langle \text{cond3} \rangle}(R)$$

Operazioni Relazionali Unarie

Figure 7.8 Results of SELECT and PROJECT operations.

(a) $\sigma_{(DNO=4 \text{ AND } SALARY > 25000) \text{ OR } (DNO=5 \text{ AND } SALARY > 30000)}(EMPLOYEE)$.

(b) $\pi_{LNAME, FNAME, SALARY}(EMPLOYEE)$. (c) $\pi_{SEX, SALARY}(EMPLOYEE)$

(a)

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
Franklin	T	Wong	333445555	1955-12-08	638 Voss,Houston,TX	M	40000	888665555	5
Jennifer		Wallace	987654321	1941-06-20	291 Berry,Bellaire,TX	F	43000	888665555	4
Ramesh		Narayan	666884444	1962-09-15	975 FireOak,Humble,TX	M	38000	333445555	5

(b)

LNAME	FNAME	SALARY
Smith	John	30000
Wong	Franklin	40000
Zelaya	Alicia	25000
Wallace	Jennifer	43000
Narayan	Ramesh	38000
English	Joyce	25000
Jabbar	Ahmad	25000
Borg	James	55000

(c)

SEX	SALARY
M	30000
M	40000
F	25000
F	43000
M	38000
M	25000
M	55000

Operazioni Relazionali Unarie

Operazione di proiezione (PROJECT)

L'operazione di proiezione seleziona alcuni **attributi** (colonne) da una relazione (tabella) e scarta gli altri. L'operazione PROJECT può essere vista come una *partizione verticale* della relazione in due relazioni: una con gli attributi richiesti, contenente il risultato dell'operazione, e l'altra con quelli non richiesti.

Esempio: Per elencare solo il nome, il cognome e lo stipendio dei dipendenti, si usa la notazione:

$$\pi_{\text{LNAME, FNAME, SALARY}}(\text{EMPLOYEE})$$

In generale, l'operazione di proiezione è indicata con:

$$\pi_{\langle \text{elenco attributi} \rangle}(\mathbf{R})$$

dove π (pi-greco) è il simbolo utilizzato per denotare l'operazione e $\langle \text{elenco attributi} \rangle$ è l'elenco degli attributi *desiderati* tra gli attributi della relazione R .

L'operazione di proiezione **rimuove** eventuali **duplicati** delle tuple, in modo che il risultato dell'operazione sia un'insieme di tuple, e quindi una relazione, **valida**.

Operazioni Relazionali Unarie

Proprietà dell'operazione PROJECT

- Il numero di tuple nel risultato di un'operazione di proiezione $\pi_{\langle \text{attributi} \rangle}(R)$ è sempre **minore o uguale** al numero di tuple in R .
- Se l'elenco degli attributi include una **chiave** di R , allora il numero di tuple nel risultato della proiezione è **uguale** al numero di tuple in R .
- Vale la seguente uguaglianza:

$$\pi_{\langle \text{attributi}_1 \rangle}(\pi_{\langle \text{attributi}_2 \rangle}(R)) = \pi_{\langle \text{attributi}_1 \rangle}(R)$$

se $\langle \text{attributi}_2 \rangle$ contiene gli attributi elencati in $\langle \text{attributi}_1 \rangle$

Operazioni Relazionali Unarie

Figure 7.8 Results of SELECT and PROJECT operations.

(a) $\sigma_{(DNO=4 \text{ AND } SALARY > 25000) \text{ OR } (DNO=5 \text{ AND } SALARY > 30000)}(EMPLOYEE)$.

(b) $\pi_{LNAME, FNAME, SALARY}(EMPLOYEE)$. (c) $\pi_{SEX, SALARY}(EMPLOYEE)$

(a)

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
Franklin	T	Wong	333445555	1955-12-08	638 Voss,Houston,TX	M	40000	888665555	5
Jennifer		Wallace	987654321	1941-06-20	291 Berry,Bellaire,TX	F	43000	888665555	4
Ramesh		Narayan	666884444	1962-09-15	975 FireOak,Humble,TX	M	38000	333445555	5

(b)

LNAME	FNAME	SALARY
Smith	John	30000
Wong	Franklin	40000
Zelaya	Alicia	25000
Wallace	Jennifer	43000
Narayan	Ramesh	38000
English	Joyce	25000
Jabbar	Ahmad	25000
Borg	James	55000

(c)

SEX	SALARY
M	30000
M	40000
F	25000
F	43000
M	38000
M	25000
M	55000

Operazioni Relazionali Unarie

Operazione di ridenominazione (RENAME)

È probabile che si vogliano eseguire più operazioni di algebra relazionale una di seguito all'altra. È possibile scrivere le operazioni desiderate come una singola espressione dell'algebra relazionale nidificando le operazioni, oppure applicare un'operazione alla volta e creare relazioni contenenti i risultati intermedi. In quest'ultimo caso occorre dare un nome alle relazioni intermedie.

Esempio: Per trovare il nome, cognome e lo stipendio dei dipendenti che lavorano nel dipartimento numero 5, occorre utilizzare un'operazione di selezione ed una di proiezione.

Possiamo scrivere una singola espressione dell'algebra relazione in questo modo:

$$\pi_{\text{FNAME, LNAME, SALARY}}(\sigma_{\text{DNO}=5}(\text{EMPLOYEE}))$$

OPPURE possiamo mostrare esplicitamente la sequenza delle singole operazioni algebriche, dando un nome ad ogni relazione intermedia:

$$\begin{aligned}\text{DEP5_EMPS} &\leftarrow \sigma_{\text{DNO}=5}(\text{EMPLOYEE}) \\ \text{RESULT} &\leftarrow \pi_{\text{FNAME, LNAME, SALARY}}(\text{DEP5_EMPS})\end{aligned}$$

Operazioni Relazionali Unarie

Operazione di ridenominazione

L'operatore di ridenominazione è ρ (rho)

In generale un'operazione RENAME può essere espressa in una delle seguenti forme:

- $\rho_{S(B_1, B_2, \dots, B_n)}(R)$
è una relazione ridenominata S , basata su $R(A_1, A_2, \dots, A_n)$, con gli attributi ridenominati in B_1, B_2, \dots, B_n .
- $\rho_S(R)$
è una relazione ridenominata S , basata su R , che non specifica i nomi degli attributi.
- $\rho_{(B_1, B_2, \dots, B_n)}(R)$
è una relazione con attributi ridenominati B_1, B_2, \dots, B_n , che non specifica il nuovo nome della relazione.

Operazioni Relazionali Unarie

Figure 7.9 Results of relational algebra expressions.

(a) $\pi_{\text{LNAME, FNAME, SALARY}}(\sigma_{\text{DNO}=5}(\text{EMPLOYEE}))$. (b) The same expression using intermediate relations and renaming of attributes.

(a)

FNAME	LNAME	SALARY
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000
Joyce	English	25000

(b)

TEMP	FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
	John	B	Smith	123456789	1965-01-09	731 Fondren,Houston,TX	M	30000	333445555	5
	Franklin	T	Wong	333445555	1955-12-08	638 Voss,Houston,TX	M	40000	888665555	5
	Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak,Humble,TX	M	38000	333445555	5
	Joyce	A	English	453453453	1972-07-31	5631 Rice,Houston,TX	F	25000	333445555	5

	FIRSTNAME	LASTNAME	SALARY
	John	Smith	30000
	Franklin	Wong	40000
	Ramesh	Narayan	38000
	Joyce	English	25000

Operazioni “Insiemistiche”

Operazione di unione (UNION)

Il risultato di questa operazione, indicata con $\mathbf{R} \cup \mathbf{S}$, è una relazione che include tutte le tuple presenti in R , oppure in S , oppure in entrambe le relazioni. Eventuali duplicati vengono eliminati.

Esempio: Per trovare il codice fiscale di tutti gli impiegati che lavorano nel dipartimento numero 5 o che sono supervisionati da un impiegato che lavora nel dipartimento numero 5, si può usare un'operazione di unione in questo modo:

$$\text{DEP5_EMPS} \leftarrow \sigma_{\text{DNO}=5}(\text{EMPLOYEE})$$

$$\text{RESULT}_1 \leftarrow \pi_{\text{SSN}}(\text{DEP5_EMPS})$$

$$\text{RESULT}_2(\text{SSN}) \leftarrow \pi_{\text{SUPERSSN}}(\text{DEP5_EMPS})$$

$$\text{RESULT} \leftarrow \text{RESULT}_1 \cup \text{RESULT}_2$$

L'operazione di unione produce come risultato le tuple presenti in RESULT_1 , o in RESULT_2 , o in entrambi.

I due operandi devono essere **compatibili all'unione**, cioè devono avere lo stesso **tipo di tuple**.

Operazioni “Insiemistiche”

Compatibilità all'unione (di tipo)

- Le relazioni operande $R_1(A_1, A_2, \dots, A_n)$ e $R_2(B_1, B_2, \dots, B_n)$ devono avere lo stesso numero di attributi, ed il dominio degli attributi corrispondenti deve essere corrispondente; si deve cioè avere che $dom(A_i) = dom(B_i)$ per ogni $i=1, 2, \dots, n$.
- La relazione risultante dalle operazioni $R_1 \cup R_2$, $R_1 \cap R_2$, o $R_1 - R_2$ avrà gli stessi nomi di attributo della prima relazione operanda R_1 (per convenzione).

Operazioni “Insiemistiche”

Figure 7.11 Illustrating the set operations union, intersection, and difference. (a) Two union compatible relations. (b) $\text{STUDENT} \cup \text{INSTRUCTOR}$. (c) $\text{STUDENT} \cap \text{INSTRUCTOR}$. (d) $\text{STUDENT} - \text{INSTRUCTOR}$. (e) $\text{INSTRUCTOR} - \text{STUDENT}$.

(a)

STUDENT	FN	LN
	Susan	Yao
	Ramesh	Shah
	Johnny	Kohler
	Barbara	Jones
	Amy	Ford
	Jimmy	Wang
	Ernest	Gilbert

INSTRUCTOR	FNAME	LNAME
	John	Smith
	Ricardo	Browne
	Susan	Yao
	Francis	Johnson
	Ramesh	Shah

(b)

FN	LN
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert
John	Smith
Ricardo	Browne
Francis	Johnson

(c)

FN	LN
Susan	Yao
Ramesh	Shah

(d)

FN	LN
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

(e)

FNAME	LNAME
John	Smith
Ricardo	Browne
Francis	Johnson

Operazioni “Insiemistiche”

Operazione di intersezione (INTERSECTION)

Il risultato di questa operazione, indicata con $\mathbf{R} \cap \mathbf{S}$, è una relazione che include tutte le tuple che sono presenti sia nella relazione R , sia nella relazione S . I due operandi devono essere *compatibili all'unione*.

Esempio: Il risultato dell'operazione di intersezione tra STUDENTI ed INSEGNANTI, include solo coloro che sono sia studenti che insegnanti.

FN	LN
Susan	Yao
Ramesh	Shah

STUDENT \cap INSTRUCTOR

Operazioni “Insiemistiche”

Operazione di differenza (MINUS)

Il risultato di questa operazione, indicata con $\mathbf{R} - \mathbf{S}$, è una relazione che include le tuple che sono presenti in R ma non in S . I due operandi devono essere *compatibili all'unione*.

Esempio: La figura mostra i nomi degli studenti che non sono insegnanti ed i nomi degli insegnanti che non sono studenti.

STUDENT	FN	LN
	Susan	Yao
	Ramesh	Shah
	Johnny	Kohler
	Barbara	Jones
	Amy	Ford
	Jimmy	Wang
	Ernest	Gilbert

FN	LN
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

STUDENT – INSTRUCTOR

FNAME	LNAME
John	Smith
Ricardo	Browne
Francis	Johnson

INSTRUCTOR – STUDENT

Operazioni “Insiemistiche”

Proprietà delle operazioni UNION, INTERSECTION e MINUS

- Sia l'operazione di unione, sia quella di intersezione sono operazioni commutative; quindi:

$$R \cup S = S \cup R, \text{ e } R \cap S = S \cap R$$

- Sia le operazioni di unione, sia quelle di intersezione possono essere trattate come operazioni n-arie, applicabili a qualunque numero di relazioni. Sono entrambe operazioni associative; quindi:

$$R \cup (S \cup T) = (R \cup S) \cup T, \text{ e } (R \cap S) \cap T = R \cap (S \cap T)$$

- L'operazione di differenza **non** è commutativa; in generale quindi:

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

Operazioni “Insiemistiche”

Operazione di prodotto cartesiano (prodotto incrociato)

Questa operazione viene utilizzata per unire tuple di due relazioni in modo combinatorio. In generale, il risultato dell'operazione $R(A_1, A_2, \dots, A_n) \times S(B_1, B_2, \dots, B_m)$ è una relazione \mathcal{Q} di grado $n+m$ attributi $\mathcal{Q}(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$, nell'ordine indicato. La relazione risultante \mathcal{Q} ha una tupla per ogni possibile combinazione tra una tupla di R ed una di S .

Quindi, se R ha n_R tuple (indicate come $|R| = n_R$), e S ha n_S tuple, allora

$$|R \times S| \text{ avrà } n_R \cdot n_S \text{ tuple.}$$

I due operandi **non** devono essere compatibili all'unione.

Esempio:

```
FEMALE_EMPS ← σSEX='F'(EMPLOYEE)
EMPNAMES ← πFNAME, LNAME, SSN(FEMALE_EMPS)
EMP_DEPENDENTS ← EMPNAMES x DEPENDENT
```

Operazioni “Insiemistiche”

Figure 7.12 An illustration of the CARTESIAN PRODUCT operation.

FEMALE EMPS	FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
	Alicia	J	Zelaya	999897777	1968-07-19	3321 Castle Spring,TX	F	25000	987654321	4
	Jennifer	S	Wallace	987654321	1941-06-20	291 Berry,Bellarm,TX	F	43000	888665555	4
	Joyce	A	English	453453453	1972-07-31	5631 Rice,Houston,TX	F	25000	333445555	6

EMPINAMES	FNAME	LNAME	SSN
	Alicia	Zelaya	999897777
	Jennifer	Wallace	987654321
	Joyce	English	453453453

EMP_DEPENDENTS	FNAME	LNAME	SSN	ESSN	DEPENDENT_NAME	SEX	BDATE	...
	Alicia	Zelaya	999897777	333445555	Alice	F	1966-04-05	...
	Alicia	Zelaya	999897777	333445555	Theodore	M	1983-10-25	...
	Alicia	Zelaya	999897777	333445555	Joy	F	1958-05-03	...
	Alicia	Zelaya	999897777	987654321	Abner	M	1942-02-28	...
	Alicia	Zelaya	999897777	123456789	Michael	M	1988-01-04	...
	Alicia	Zelaya	999897777	123456789	Alice	F	1988-12-30	...
	Alicia	Zelaya	999897777	123456789	Elizabeth	F	1967-05-05	...
	Jennifer	Wallace	987654321	333445555	Alice	F	1966-04-05	...
	Jennifer	Wallace	987654321	333445555	Theodore	M	1983-10-25	...
	Jennifer	Wallace	987654321	333445555	Joy	F	1958-05-03	...
	Jennifer	Wallace	987654321	987654321	Abner	M	1942-02-28	...
	Jennifer	Wallace	987654321	123456789	Michael	M	1988-01-04	...
	Jennifer	Wallace	987654321	123456789	Alice	F	1988-12-30	...
	Jennifer	Wallace	987654321	123456789	Elizabeth	F	1967-05-05	...
	Joyce	English	453453453	333445555	Alice	F	1966-04-05	...
	Joyce	English	453453453	333445555	Theodore	M	1983-10-25	...
	Joyce	English	453453453	333445555	Joy	F	1958-05-03	...
	Joyce	English	453453453	987654321	Abner	M	1942-02-28	...
	Joyce	English	453453453	123456789	Michael	M	1988-01-04	...
	Joyce	English	453453453	123456789	Alice	F	1988-12-30	...
	Joyce	English	453453453	123456789	Elizabeth	F	1967-05-05	...

ACTUAL_DEPENDENTS	FNAME	LNAME	SSN	ESSN	DEPENDENT_NAME	SEX	BDATE
	Jennifer	Wallace	987654321	987654321	Abner	M	1942-02-28

RESULT	FNAME	LNAME	DEPENDENT_NAME
	Jennifer	Wallace	Abner

Operazioni Relazionali Binarie

Operazione di JOIN

La sequenza di un'operazione di prodotto cartesiano seguita da una di selezione, è usata molto comunemente per identificare e selezionare tuple correlate da due relazioni. Si tratta di un'operazione speciale, chiamata JOIN, che viene indicata con il simbolo \bowtie .

Questa operazione è molto importante per tutti i database relazionali con più di una relazione perché permette di eseguire associazioni tra relazioni.

In generale l'operazione di join tra due relazioni $R(A_1, A_2, \dots, A_n)$ e $S(B_1, B_2, \dots, B_m)$ è indicata:

$$R \bowtie_{\langle \text{condizione di join} \rangle} S$$

dove R e S possono essere qualunque relazione che risulti da un'espressione dell'algebra relazionale generale.

Operazioni Relazionali Binarie

- *Esempio:* Trovare i nomi dei manager di ciascun dipartimento.
Per trovare i nomi dei manager occorre combinare ciascuna tupla di DIPARTIMENTO con la tupla di DIPENDENTE in cui il valore di SSN corrisponde al valore di MGRSSN della tupla del dipartimento. Questo si ottiene mediante l'operazione di join:

$\text{DIP_MGR} \leftarrow \text{DIPARTIMENTO} \bowtie_{\text{MGRSSN}=\text{SSN}} \text{DIPENDENTE}$

DEPT_MGR	DNAME	DNUMBER	MGRSSN	...	FNAME	MINIT	LNAME	SSN	...
	Research	5	333445555	...	Franklin	T	Wong	333445555	...
	Administration	4	987654321	...	Jennifer	S	Wallace	987654321	...
	Headquarters	1	888665555	...	James	E	Borg	888665555	...

FIGURE 6.6 Result of the JOIN operation $\text{DEPT_MGR} \leftarrow \text{DEPARTMENT} \bowtie_{\text{MGRSSN}=\text{SSN}} \text{EMPLOYEE}$.

Operazioni Relazionali Binarie

THETA-JOIN: $R \bowtie_{\langle A_i \theta B_j \dots A_k \theta B_l \rangle} S$

Il risultato di un'operazione di join sulle relazioni $R(A_1, \dots, A_n)$ e $S(B_1, \dots, B_m)$ è una relazione \mathcal{Q} con $n+m$ attributi $\mathcal{Q}(A_1, \dots, A_n, B_1, \dots, B_m)$.

Le tuple di \mathcal{Q} sono formate dalle combinazioni delle tuple di R ed S per le quali viene soddisfatta la condizione di join.

Una condizione di join assume la forma generale:

$\langle \text{condizione} \rangle \text{ AND } \langle \text{condizione} \rangle \text{ AND } \dots \text{ AND } \langle \text{condizione} \rangle$

in cui ogni condizione assume la forma $A_i \theta B_j$ con θ uno degli operatori di confronto $\{=, <, \leq, >, \geq, \neq\}$.

Le tuple i cui attributi di join sono NULL **non** compaiono nel risultato. Per questo l'operazione di join **non** conserva necessariamente tutte le informazioni presenti nelle relazioni partecipanti.

Operazioni Relazionali Binarie

EQUI-JOIN

L'uso più comune delle operazioni di join utilizza condizioni di sola **uguaglianza**. Questo tipo particolare di theta-join, in cui l'unico operatore di comparazione usato è =, viene chiamata **equi-join**. Nel risultato di un'equi-join si avranno sempre una o più coppie di attributi con valori identici in ciascuna tupla.

JOIN NATURALE

Poiché uno degli attributi nelle coppie con valori identici è superfluo, è stata creata una nuova operazione, chiamata **join naturale** ed indicata con *, che non include il secondo attributo ripetuto in una condizione di uguaglianza.

La definizione comune di una join naturale richiede che i due attributi di join, o ciascuna coppia di attributi di join, abbiano lo **stesso nome** in entrambe le relazioni. Altrimenti viene applicata una ridenominazione.

Operazioni Relazionali Binarie

Esempio: Per applicare una join naturale sugli attributi DNUMBER di DIPARTIMENTO e DIP_LUOGO, è sufficiente scrivere:

DEPT_LOCS \leftarrow DEPARTMENT * DEPT_LOCATIONS

(a)	PROJ_DEPT	PNAME	PNUMBER	PLOCATION	DNUM	DNAME	MGRSSN	MGRSTARTDATE
		ProductX	1	Bellaire	5	Research	333445555	1988-05-22
		ProductY	2	Sugarland	5	Research	333445555	1988-05-22
		ProductZ	3	Houston	5	Research	333445555	1988-05-22
		Computerization	10	Stafford	4	Administration	987654321	1995-01-01
		Reorganization	20	Houston	1	Headquarters	888665555	1981-06-19
		Newbenefits	30	Stafford	4	Administration	987654321	1995-01-01

(b)	DEPT_LOCS	DNAME	DNUMBER	MGRSSN	MGRSTARTDATE	LOCATION
		Headquarters	1	888665555	1981-06-19	Houston
		Administration	4	987654321	1995-01-01	Stafford
		Research	5	333445555	1988-05-22	Bellaire
		Research	5	333445555	1988-05-22	Sugarland
		Research	5	333445555	1988-05-22	Houston

FIGURE 6.7 Results of two NATURAL JOIN operations. (a) PROJ_DEPT \leftarrow PROJECT * DEPT. (b) DEPT_LOCS \leftarrow DEPARTMENT * DEPT_LOCATIONS.

Insieme completo di operazioni relazionali

- L'insieme delle operazioni dell'algebra relazionale formato da selezione σ , proiezione π , unione \cup , differenza $-$, e prodotto cartesiano \times è detto insieme completo perché ogni altra espressione dell'algebra relazionale può essere espressa come combinazione di queste cinque operazioni.

-

- Ad esempio:

$$R \cap S = (R \cup S) - ((R - S) \cup (S - R))$$

$$R \bowtie_{\langle \text{condizione di join} \rangle} S = \sigma_{\langle \text{condizione di join} \rangle} (R \times S)$$

Operazioni Relazionali Binarie

Operazione di divisione (DIVISION)

Un'operazione di divisione si applica a due relazioni $\mathbf{R(Z)} \div \mathbf{S(X)}$, in cui $X \subseteq Z$. Sia $\gamma = Z - X$ (e quindi $Z = X \cup \gamma$); sia, cioè, γ l'insieme degli attributi di R che non sono attributi di S .

Il risultato della divisione è una relazione $T(\gamma)$ che include una tupla t se in R sono presenti tuple con $t_R[\gamma] = t$, e con $t_R[X] = t_S$ per ogni tupla t_S di S .

Ciò significa che, affinché una tupla t appaia nel risultato T della divisione, in R devono comparire i valori di t in combinazione con ogni tupla di S .

TABLE 6.1 OPERATIONS OF RELATIONAL ALGEBRA

Operation	Purpose	Notation
SELECT	Selects all tuples that satisfy the selection condition from a relation R .	$\sigma_{\langle \text{SELECTION CONDITION} \rangle}(R)$
PROJECT	Produces a new relation with only some of the attributes of R , and removes duplicate tuples.	$\pi_{\langle \text{ATTRIBUTE LIST} \rangle}(R)$
THETA JOIN	Produces all combinations of tuples from R_1 and R_2 that satisfy the join condition.	$R_1 \bowtie_{\langle \text{JOIN CONDITION} \rangle} R_2$
EQUIJOIN	Produces all the combinations of tuples from R_1 and R_2 that satisfy a join condition with only equality comparisons.	$R_1 \bowtie_{\langle \text{JOIN CONDITION} \rangle} R_2$, OR $R_1 \bowtie_{(\langle \text{JOIN ATTRIBUTES 1} \rangle, \langle \text{JOIN ATTRIBUTES 2} \rangle)} R_2$
NATURAL JOIN	Same as EQUIJOIN except that the join attributes of R_2 are not included in the resulting relation; if the join attributes have the same names, they do not have to be specified at all.	$R_1^*_{\langle \text{JOIN CONDITION} \rangle} R_2$, OR $R_1^*_{(\langle \text{JOIN ATTRIBUTES 1} \rangle, \langle \text{JOIN ATTRIBUTES 2} \rangle)} R_2$ OR $R_1 \star R_2$
UNION	Produces a relation that includes all the tuples in R_1 or R_2 or both R_1 and R_2 ; R_1 and R_2 must be union compatible.	$R_1 \cup R_2$
INTERSECTION	Produces a relation that includes all the tuples in both R_1 and R_2 ; R_1 and R_2 must be union compatible.	$R_1 \cap R_2$
DIFFERENCE	Produces a relation that includes all the tuples in R_1 that are not in R_2 ; R_1 and R_2 must be union compatible.	$R_1 - R_2$
CARTESIAN PRODUCT	Produces a relation that has the attributes of R_1 and R_2 and includes as tuples all possible combinations of tuples from R_1 and R_2 .	$R_1 \times R_2$
DIVISION	Produces a relation $R(X)$ that includes all tuples $t[X]$ in $R_1(Z)$ that appear in R_1 in combination with every tuple from $R_2(Y)$, where $Z = X \cup Y$.	$R_1(Z) \div R_2(Y)$

Altre Operazioni Relazionali

Funzioni aggregate e raggruppamento

Un tipo di interrogazione che non può essere espressa in termini di algebra relazionale di base consiste nello specificare **funzioni aggregate** matematiche su collezioni di valori del database.

Esempi di queste funzioni prevedono il recupero dello stipendio medio o complessivo di tutti gli impiegati, oppure il numero totale delle tuple di impiegato. Queste funzioni vengono usate in semplici interrogazioni statistiche che riassumono le informazioni provenienti dalla tuple del database.

Funzioni comuni applicate a collezioni di valori numerici sono SUM, AVERAGE, MAXIMUM, e MINIMUM. La funzione COUNT viene utilizzata per contare tuple o valori.

Altre Operazioni Relazionali

Uso dell'operatore funzionale \mathcal{F}

Operatore di funzione aggregata, indicato con \mathcal{F} , nella forma:

<attributi raggruppamento> \mathcal{F} <lista funzioni> (R)

dove <attributi raggruppamento> è una lista di attributi della relazione R e <lista funzioni> è una lista di coppie (<funzione>, <attributo>). La relazione risultante presenta gli attributi di raggruppamento più un attributo per ogni elemento nella lista di funzioni.

$\mathcal{F}\text{MAX}_{\text{Stipendio}}$ (DIPENDENTE) ricava il valore dello stipendio massimo dalla relazione
DIPENDENTE

$\mathcal{F}\text{MIN}_{\text{Stipendio}}$ (DIPENDENTE) ricava il valore dello stipendio minimo dalla relazione
DIPENDENTE

$\mathcal{F}\text{SUM}_{\text{Stipendio}}$ (DIPENDENTE) ricava la somma degli stipendi dalle relazione DIPENDENTE

DNO $\mathcal{F}\text{COUNT}_{\text{SSN}}$, $\text{AVERAGE}_{\text{Stipendio}}$ (DIPENDENTE) raggruppa i dipendenti per DNO (numero di dipartimento) e calcola il numero di dipendenti e lo stipendio medio per ciascun dipartimento. [Nota: COUNT conta semplicemente il numero di righe, senza rimozione dei duplicati]

Altre Operazioni Relazionali

(a)

R	DNO	NO_OF_EMPLOYEES	AVERAGE_SAL
	5	4	33250
	4	3	31000
	1	1	55000

(b)

DNO	COUNT_SSN	AVERAGE_SALARY
5	4	33250
4	3	31000
1	1	55000

(c)

COUNT_SSN	AVERAGE_SALARY
8	35125

- (a) $\rho_{R(DNO, NO_OF_EMPLOYEES, AVERAGE_SAL)}(DNO \bowtie COUNT_{SSN}, AVERAGE_{Salary})(Employee)$
- (b) $DNO \bowtie COUNT_{SSN}, AVERAGE_{Salary}(Employee)$
- (c) $\bowtie COUNT_{SSN}, AVERAGE_{Salary}(Employee)$

Altre Operazioni Relazionali

Operazioni di Chiusura Ricorsiva

- Un altro tipo di operazione che, in generale, non può essere specificato in algebra relazionale di base è la chiusura ricorsiva. Questa operazione si applica ad un'associazione ricorsiva.
- Un esempio di operazione ricorsiva è la ricerca di tutti i supervisori di un dipendente e a tutti i livelli: cioè di tutti i dipendenti e' direttamente sottoposti ad e ; di tutti i dipendenti e'' direttamente sottoposti agli e' ; di tutti i dipendenti e''' direttamente sottoposti agli e'' , etc...
- Anche se è possibile trovare gli impiegati a ciascun livello e poi utilizzare un'unione, non è possibile, in generale, specificare un'interrogazione di questo tipo senza utilizzare un ciclo.
- Lo standard SQL₃ include una sintassi per la chiusura ricorsiva.

Altre Operazioni Relazionali

(Borg's SSN is 888665555)

(SSN) (SUPERSSN)

SUPERVISION	SSN1	SSN2
	123456789	333445555
	333445555	888665555
	999887777	987654321
	987654321	888665555
	666884444	333445555
	453453453	333445555
	987987987	987654321

RESULT 1	SSN
	333445555
	987654321

(Supervised by Borg)

RESULT 2	SSN
	123456789
	999887777
	666884444
	453453453
	987987987

(Supervised by Borg's subordinates)

RESULT	SSN
	123456789
	999887777
	666884444
	453453453
	987987987
	333445555
	987654321

(RESULT1 \cup RESULT2)

$BORG_SSN \leftarrow \pi_{SSN}(\sigma_{NOME='James' \text{ AND } COGNOME='BORG'}(DIPENDENTE))$

$SUPERVISION(SSN1, SSN2) \leftarrow \pi_{SSN, SUPERSSN}(DIPENDENTE)$

$RESULT1(SSN) \leftarrow \pi_{SSN1}(SUPERVISION \bowtie_{SSN2=SSN} BORG_SSN)$

$RESULT2(SSN) \leftarrow \pi_{SSN1}(SUPERVISION \bowtie_{SSN2=SSN} RESULT1)$

$RESULT \leftarrow RESULT2 \cup RESULT1$

Altre Operazioni Relazionali

L'operazione di join esterna (OUTER JOIN)

Nelle join naturali, le tuple senza corrispondenze sono eliminate dal risultato dell'operazione. Anche le tuple con valori NULL negli attributi di join sono eliminate. Ciò causa perdita di informazione.

Un insieme di operazioni, chiamate **join esterne**, possono essere usate quando si vuole tenere nel risultato di una join tutte le tuple di R , oppure tutte le tuple di S , oppure quelle di entrambe le relazioni, anche nel caso non si abbiano corrispondenze negli attributi di join.

- L'operazione di join esterna sinistra (LEFT OUTER JOIN) mantiene tutte le tuple della prima relazione (di sinistra) R in $\mathbf{R} \bowtie_{\text{LEFT}} \mathbf{S}$. Se non c'è una corrispondenza con una tupla di S , gli attributi di S del risultato di join vengono riempiti con valori NULL.
- Un'operazione analoga, join esterna destra (RIGHT OUTER JOIN), mantiene tutte le tuple della seconda relazione (di destra) S nel risultato di $\mathbf{R} \bowtie_{\text{RIGHT}} \mathbf{S}$.
- Una terza operazione, join esterna totale (FULL OUTER JOIN), mantiene tutte le tuple di entrambe le relazioni in $\mathbf{R} \bowtie_{\text{FULL}} \mathbf{S}$.

Altre Operazioni Relazionali

RESULT	FNAME	MINIT	LNAME	DNAME
	John	B	Smith	null
	Franklin	T	Wong	Research
	Alicia	J	Zelaya	null
	Jennifer	S	Wallace	Administration
	Ramesh	K	Narayan	null
	Joyce	A	English	null
	Ahmad	V	Jabbar	null
	James	E	Borg	Headquarters

LEFT OUTER JOIN

$TEMP \leftarrow (IMPIEGATO \bowtie_{LEFTSSN=SSN_DIR} DIPARTIMENTO)$

$RESULT \leftarrow \pi_{FNAME, MINIT, LNAME, DNAME}(TEMP)$