
Fondamenti di Teoria delle Basi di Dati

Riccardo Torlone

Parte 2:
Il modello relazionale

Modelli logici, caratteristiche

- **Gerarchico e reticolare**
 - utilizzano riferimenti espliciti (puntatori) fra record
- **Relazionale "è basato su valori"**
 - anche i riferimenti fra dati in strutture (relazioni) diverse sono rappresentati per mezzo dei valori stessi

Il modello relazionale

- Proposto da E. F. Codd nel 1970 per favorire l'indipendenza dei dati
- Disponibile in DBMS reali nel 1981 (non è facile implementare l'indipendenza con efficienza e affidabilità)
- Si basa sul concetto matematico di relazione (con una variante)
- Le relazioni hanno naturale rappresentazione per mezzo di tabelle

Una tabella relazionale

studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

Una base di dati relazionale

studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

esami

Studente	Voto	Corso
3456	30	04
3456	24	02
9283	28	01
6554	26	01

corsi

Codice	Titolo	Docente
01	Analisi	Mario
02	Chimica	Bruni
04	Chimica	Verdi

Intuitivamente

- Una base di dati è un insieme di tabelle
- Ogni tabella è fatta da:
 - Un insieme di colonne
 - Un insieme di righe:
 - Una “speciale” (intestazione - schema)
 - Le altre “simili” (corpo - istanza)
- In ogni colonna ci sono valori “omogenei”
- Le righe del corpo sono distinte
- Le righe del corpo non si ripetono
- ... cerchiamo di formalizzare

Relazione matematica

- Siano dati n insiemi anche non distinti D_1, \dots, D_n detti **domini** della relazione
- **Prodotto cartesiano** $D_1 \times \dots \times D_n$:
 - l'insieme di tutte le n -uple (d_1, \dots, d_n) tali che $d_1 \in D_1, \dots, d_n \in D_n$
- **Relazione matematica** su D_1, \dots, D_n :
 - un sottoinsieme di $D_1 \times \dots \times D_n$

Relazione matematica, esempio

- $D_1 = \{a, b\}$
- $D_2 = \{x, y, z\}$
- prodotto cartesiano $D_1 \times D_2$

a	x
a	y
a	z
b	x
b	y
b	z

■ una relazione

$$r \subseteq D_1 \times D_2$$

a	x
a	z
b	y

Relazione matematica, proprietà

- una relazione matematica è un insieme di n-uple ordinate:
 - (d_1, \dots, d_n) tali che $d_1 \in D_1, \dots, d_n \in D_n$
- una relazione è un insieme; quindi:
 - non c'è ordinamento fra le n-uple
 - le n-uple sono distinte
 - ciascuna n-upla è ordinata: l' i-esimo valore proviene dall' i-esimo dominio

Relazione matematica, esempio

$Partite \subseteq string \times string \times int \times int$

Juve	Lazio	3	1
Lazio	Milan	2	0
Juve	Roma	0	2
Roma	Milan	0	1

- Ciascuno dei domini ha due **ruoli** diversi, distinguibili attraverso la posizione:
 - La struttura è **posizionale**

Struttura non posizionale

- A ciascun dominio si associa un nome (**attributo**), che ne descrive il "ruolo"

Casa	Fuori	RetiCasa	RetiFuori
Juve	Lazio	3	1
Lazio	Milan	2	0
Juve	Roma	0	2
Roma	Milan	0	1

Definizioni

- Schema di relazione:
un nome R con un insieme di attributi A_1, \dots, A_n :

$$R(A_1, \dots, A_n)$$

- Schema di base di dati:
insieme di schemi di relazione:

$$R = \{R_1(X_1), \dots, R_k(X_k)\}$$

Definizioni, 2

- Una ennupla (o tupla) su un insieme di attributi X è una funzione che associa a ciascun attributo A in X un valore del dominio di A
- $t[A]$ denota il valore della tupla t sull'attributo A

Definizioni, 3

- (Istanza di) relazione su uno schema $R(X)$:
 - insieme r di tuple su X
- (Istanza di) base di dati su uno schema $R = \{R_1(X_1), \dots, R_n(X_n)\}$:
 - insieme di relazioni $r = \{r_1, \dots, r_n\}$ (con r_i relazione su R_i)

Tabelle e relazioni

- Una tabella rappresenta una relazione se
 - i valori di ogni colonna sono fra loro omogenei
 - le righe sono diverse fra loro
 - le intestazioni delle colonne sono diverse tra loro
- In una tabella che rappresenta una relazione
 - l'ordinamento tra le righe è irrilevante
 - l'ordinamento tra le colonne è irrilevante

Il modello è basato su valori

- I riferimenti fra dati in relazioni diverse sono rappresentati per mezzo di valori dei domini che compaiono nelle tuple

studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

esami

Studente	Voto	Corso
3456	30	04
3456	24	02
9283	28	01
6554	26	01

corsi

Codice	Titolo	Docente
01	Analisi	Mario
02	Chimica	Bruni
04	Chimica	Verdi

studenti

Matricola	Cognome	Nome	Data di nascita
-----------	---------	------	-----------------

6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

esami

Studente	Voto	Corso
----------	------	-------

	30	
	24	
	28	
	26	

corsi

Codice	Titolo	Docente
--------	--------	---------

01	Analisi	Mario
02	Chimica	Bruni
04	Chimica	Verdi

Struttura basata su valori: vantaggi

- indipendenza dalle strutture fisiche (si potrebbe avere anche con puntatori di alto livello) che possono cambiare dinamicamente
- si rappresenta solo ciò che è rilevante dal punto di vista dell'applicazione
- l'utente finale vede gli stessi dati dei programmatori
- i dati sono portabili più facilmente da un sistema ad un altro
- i puntatori sono direzionali