

Il Modello Relazionale

Atzeni, Ceri, Paraboschi, Torlone
"Basi Di Dati:
Modelli e linguaggi di interrogazione",
McGraw-Hill Italia, Capitolo 2

Il Modello Relazionale

Si basa su due concetti:

- Tabella (concetto primitivo)
- Relazione
 - Matematica
 - Secondo la definizione del modello relazionale
 - Con riferimento al modello concettuale entity-relationship
- La forza del modello è dovuta al suo essere basato su un concetto estremamente intuitivo (tabella) e su un concetto estremamente ben formalizzato dal punto di vista matematico (relazione).

Prodotto Cartesiano

- Dati due insiemi $D1$ e $D2$ si dice prodotto cartesiano di $D1$ e $D2$ ($D1 \times D2$) l'insieme formato dalle coppie ordinate $(v1, v2)$ con $v1 \in D1$ e $v2 \in D2$
- Esempio:
 - $A = \{1, 2, 3\}$ $B = \{a, b, c\}$
 - $A \times B = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, c), (3, a), (3, b), (3, c)\}$

Relazione Matematica

- Una relazione matematica sugli insiemi $D1$ e $D2$ (domini della relazione) è definita come un sottoinsieme di $(D1 \times D2)$.
- Es. $\{(1, a), (2, b), (3, c)\}$ relazione su A e B .
- Gli insiemi e quindi le relazioni possono essere rappresentati in forma tabellare:

1	a
2	b
3	c

Relazione Matematica (continua)

- Le definizioni di prodotto cartesiano e quindi di relazione sono state date con riferimento a coppie di insiemi.
- Esse sono immediatamente generalizzabili rispetto al numero di insiemi coinvolti.
- Il numero n di insiemi coinvolti viene detto grado del prodotto o della relazione.
- Alle coppie si sostituiscono le n -uple.
- Il numero di elementi della relazione viene detto cardinalità della relazione.

Relazioni con Attributi

- In una relazione non è definito alcun ordinamento tra le n -uple che la compongono.
- Due tabelle che differiscono solo nell'ordine delle righe, cioè, rappresentano la stessa relazione.
- Le n -uple di una relazione sono distinte una dall'altra, in quanto tra gli elementi di un insieme non possono essere presenti due elementi uguali.

Relazioni con Attributi (continua)

- Ogni n -upla, al suo interno, è ordinata.

juventus	lazio	0	1
milan	inter	2	0
roma	napoli	2	0
- Se scambiamo la colonna 3 con la 4 certamente alteriamo le informazioni.
- Tale caratteristica è insoddisfacente in quanto si tende, in informatica, a privilegiare notazioni non posizionali.
- La soluzione si ottiene intuitivamente assegnando ad ogni colonna un nome.

Relazioni con Attributi (continua)

SCasa	SOspite	RSC	RSO
juventus	lazio	0	1
milan	inter	2	0
roma	napoli	0	0

- In maniera formale:
 - Una relazione è assimilabile ad un insieme di record omogenei.
 - Nel caso dei record ad ogni campo è associato un nome.
 - Analogamente, si associa ad ogni occorrenza di dominio della relazione un attributo che ne descrive il ruolo.
- Nella rappresentazione tabellare gli attributi vengono utilizzati come intestazione delle colonne.

Relazioni con Attributi (continua)

- Utilizzando la nozione di attributo, l'ordine delle colonne non è più significativo.
- Relazione con Attributi
 - Stabiliamo la corrispondenza fra attributi e domini DOM che associa ad ogni attributo A un dominio.
 - Definiamo tupla su un insieme di attributi la funzione t che associa all'attributo A un valore del dominio DOM(A).

Relazioni con Attributi (continua)

- Relazione con Attributi \leftrightarrow insieme di tuple.

Esempio:

SCasa	SOspite	RSC	RSO
juventus	lazio	0	1
milan	inter	2	0
roma	napoli	0	0

- Con riferimento alla seconda tupla:
 - $t[SCasa] = t.SCasa = milan$

Relazioni e Basi di Dati

- Nel modello relazionale, una Base di Dati è costituita da più relazioni collegate, ove necessario e/o opportuno, attraverso valori comuni sulle tuple.

Studente:		Corsi:		
Matr.	Nome	Cod.	Titolo	Doc.
1555	Pippo	01	DB	Topolino
1578	Pluto	02	SI	Paperino

Esami:		
Stud.	Corso	Voto
1555	01	28
1578	02	27

Relazioni e Basi di Dati (continua)

- L'esempio mostra una caratteristica fondamentale del modello relazionale:
 - Il modello relazionale è basato su valori
 - I riferimenti fra dati in relazioni diverse sono cioè rappresentati attraverso valori uguali dei domini che compaiono nelle tuple.
 - I modelli gerarchico e reticolare sono basati su "record e puntatori".

Vantaggi del modello basato su valori.

- Rappresenta solo ciò che è rilevante.
- Tutta l'informazione è contenuta nei valori \Rightarrow è facile trasferire informazioni tra vari contesti.
- Non vi è alcun legame tra rappresentazione logica e fisica.
- A livello fisico anche un modello relazionale può essere (e, ovviamente, è) basato su puntatori.
- Il punto è che essi non sono né visibili né necessari a livello logico.

Nomenclatura

- Uno schema di relazione è costituito da un simbolo (nome della relazione) e da un insieme di attributi
Studente (Matr, Nome)
- Uno schema di basi di dati è un insieme di schemi di relazione
 $R = \{ \text{Studente (Matr, Nome)}, \text{Corsi (codice, Titolo, Doc)}, \text{Esami(Stud, Corso, Voto)} \}$
- Una istanza di relazione (o relazione) è un insieme di tuple.
- Una istanza di base di dati sullo schema $R\{R_1, \dots, R_N\}$ è un insieme di relazioni $r\{r_1, \dots, r_n\}$ dove ogni r_i è una istanza di relazione sullo schema R_i .

Valori Nulli

- Il modello relazionale è semplice e potente. Esso è però rigido.
- Es. Persone (Nome, Cognome, Indirizzo, Telefono)
 - Cosa succede se una persona non ha il telefono, o se non se ne conosce il numero, o se non si sa se ha il telefono?
 - **Non è** concettualmente corretto usare un valore del dominio per rappresentare l'assenza di informazione

Valori Nulli (continua)

- Il problema viene risolto estendendo il concetto di relazione e prevedendo cioè che una tupla possa assumere o un valore del dominio o un valore speciale, detto valore nullo (NULL).
- Questa soluzione è semplice ed efficace ma, ovviamente, non risolve tutti i problemi.
- Il valore NULL deve essere usato con estrema attenzione sia per quegli attributi usati per stabilire relazioni fra tabelle sia perché si possono generare dubbi sulla identità delle tuple.

Vincoli di Integrità

- Esempio:

Studente:	Matr	Nome
	1555	Pippo
	1578	Pluto
	1555	Gastone

Corsi:	Cod.	Titolo	Doc.
	01	DB	Topolino
	02	SI	Paperino

Esami:	Stud.	Corso	Voto	Lode
	1556	01	28	Si
	1578	02	36	No

Vincoli di Integrità (continua)

- Si è introdotto, per evitare le situazioni appena viste, il concetto di Vincolo di Integrità, inteso come una proprietà che deve essere soddisfatta dalle istanze (che si dicono allora *corrette*).
- Un vincolo è un predicato che associa alla generica istanza il valore vero/falso. Se il predicato è vero, il vincolo è soddisfatto.
- Ad uno schema di base di dati sono assegnati un certo numero di vincoli.
- Una istanza è lecita (o corretta, o consistente) se soddisfa tutti i vincoli.

Vincoli di Integrità (continua)

- Vincolo intrarelazionale** \Leftrightarrow il suo soddisfacimento è definito relativamente alla singola relazione.
 - Vincolo di tupla** \Leftrightarrow il suo soddisfacimento è definito a livello di tupla.
 - Vincolo su valori \Leftrightarrow il suo soddisfacimento riguarda singoli valori.
- Vincolo interrelazionale** \Leftrightarrow coinvolge due o più relazioni.

Vincoli di Integrità (continua)

- Vincoli di Tupla**
 - Esprimono condizioni sui valori della singola tupla.
 - $(voto \leq 30) \text{ and } (voto \geq 18)$
 - $(\text{not } (lode = 'si')) \text{ or } (voto = 30)$
 - $\text{netto} = \text{importo} - \text{ritenuta}$
- Vincoli di Chiave**
 - Sono i più importanti vincoli intrarelazionali e sono la base stessa del modello.
Si consideri il campo *Matricola*: non esistono due studenti con lo stesso numero di matricola

Chiave

Una chiave è un insieme di attributi che identificano univocamente una tupla di una relazione.

- Un insieme di attributi K si dice superchiave per la relazione r se in r non esistono due tuple t1 e t2 tali che $t1[K]=t2[K]$.
- Un insieme di attributi si dice chiave se è una chiave minimale \Leftrightarrow se non esiste una superchiave K1 contenuta propriamente in K.

Chiave (continua)

• Esempio:

Matricola	Cognome	Nome	Nascita	Corso
0001	Rossi	Mario	1963	Inf
0002	Rossi	Luca	1963	DB
0003	Neri	Luca	1962	Analisi
0004	Neri	Luca	1960	DB

{Corso, Cognome} è una superchiave minimale (chiave).

- Questo fatto è vero in generale?
- In altri termini, si tratta di una chiave reale o di una chiave "casuale" legata alla particolare istanza della relazione?
 - Si tratta di una chiave casuale legata al fatto, ad esempio, che Neri Luca del 1960 non segue il corso di Analisi.

Chiave (continua)

- Ogni relazione ha sempre una chiave.
 - Una relazione è, per definizione, un insieme \Rightarrow
 - Non possono esistere due tuple uguali \Rightarrow
 - L'insieme di tutti gli attributi è una superchiave \Rightarrow
 - esso è chiave oppure
 - esiste un'altra superchiave in esso contenuta.
 - Ragionando ricorsivamente, poiché il numero di attributi è finito, il processo termina in numero finito di passi con una superchiave minimale.

Chiavi e Valori Nulli

• Studente:

Matr	Cognome	Nome	Nascita
Null	Rossi	Mario	Null
0001	Rossi	Luca	1961
0002	Neri	Luca	Null
Null	Neri	Luca	1962

- Ipotesi: esistono 2 chiavi
 - {matr}
 - {cognome, nome, nascita}

Chiave Primaria

- Esiste allora la necessità di limitare drasticamente la proliferazione di valori nulli sulle chiavi.
- In pratica, su una delle chiavi, si vieta la presenza di valori nulli
 - (chiave primaria)

Vincoli Di Integrità Referenziale

- La più importante classe di vincoli interrelazionali.

Agente: Matr	Cognome
001	Esosito
002	Gargiulo

Auto:	Prov.	Num	Prop
TP	266	Pippo	
PA	113	Paolino	

Infrazioni:	Cod	Ag.	Prov.	Num.
	01	001	TP	266
	02	002	PA	113

Vincoli Di Integrità Referenziale (continua)

- Un vincolo di integrità referenziale (foreign key o referential integrity constraint) fra un insieme di attributi X di R1 e uno Y di R2 è vero se i valori su X di ogni tupla dell' istanza di R1 compaiono come valori di Y dell'istanza di R2
- In genere Y è la chiave (primaria) dell' istanza di R2.
- Caso 1:
 - X costituito da un solo valore (A).
 - chiave di R2 B (singolo attributo);
 - per ogni tupla t1 in R1 esiste una tupla t2 in R2 tale che: $t1[A]=t2[B]$.

Vincoli Di Integrità Referenziale (continua)

- Caso 2: X composto da più attributi.
 - In questo caso, è necessario far corrispondere ad ogni attributo di X un attributo della chiave K di R2.
- In maniera formale:
 - si stabilisca un ordinamento sugli attributi di $X=A1,A2,...AP$ e di $K=B1,B2,...BP$
- Il vincolo è vero se per ogni tupla t1 di R1 esiste una tupla t2 in R2 tale che:
 - $t1[Ai] = t2[Bi]$ per $1 \leq i \leq p$