

Il modello relazionale dei dati

Prof. Alfredo Pulvirenti (A-L)

Prof. Salvatore Alaimo (M-Z)

(Atzeni-Ceri Capitolo 2)

Definizione

- Un **modello dei dati** è un insieme di *meccanismi di astrazione* per definire una base di dati, con associato un insieme predefinito di *operatori* e di *vincoli di integrità*.
- Questi meccanismi di astrazione rivestono un ruolo equivalente a quello delle strutture dati nella programmazione.

Il modello relazionale dei dati

- Introdotto da E.F. Codd nel 1970
- Implementa bene il concetto di indipendenza dei dati
- Oggi è di gran lunga il modello dei dati più diffuso
- Implementato nei DBMS commerciali dal 1981
- E' basato sul concetto matematico di **RELAZIONE**

Relazione: tre accezioni

- **relazione matematica**: come nella teoria degli insiemi
- **relazione** secondo il modello relazionale dei dati
- **relazione** (dall'inglese **relationship**) che rappresenta una classe di fatti, nel modello Entità-Relazione; tradotto anche con **associazione** o **correlazione**

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

- D_1, \dots, D_n (n insiemi anche non distinti)
- **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 $D_1 \times \dots \times D_n$.
- D_1, \dots, D_n sono i **domini** della relazione

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 dove:
 - 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

$Partite \subseteq \text{stringa} \times \text{stringa} \times \text{intero} \times \text{intero}$

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

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

Struttura non posizionale

- A ciascun dominio si associa un nome unico nella tabella (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

Struttura non posizionale (2)

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

Schema di relazione

- Uno **schema di relazione** $R : \{T\}$ è una coppia formata dal *nome di una relazione* R e da un *tipo relazione* T

Definizione Tipo Relazione

- Interi, Reali, Booleani, Stringhe sono **tipi primitivi**
- Siano T_1, T_2, \dots, T_n tipi primitivi e A_1, A_2, \dots, A_n etichette, dette attributi, allora $(A_1:T_1, A_2:T_2, \dots, A_n:T_n)$ è un ***tipo n-upla di grado n***
- Se T è un tipo n-upla allora $\{T\}$ è un ***tipo relazione (tipo insieme di n-uple)***

Schema Relazionale

- Uno **Schema Relazionale** è costituito da:
 - un insieme di *schemi di relazione*
 $R_i : \{T_i\}, i=1, 2, \dots, k$
 - un insieme di *vincoli di integrità* relativi a tali schemi.
- Lo schema relazionale costituisce l'aspetto ***intensionale***

Aspetto *estensionale*

- Una n-upla $t = (A_1 := v_1, A_2 := v_2, \dots, A_n := v_n)$ di tipo $(A_1 : T_1, A_2 : T_2, \dots, A_n : T_n)$ è un insieme di coppie (A_i, v_i) con v_i di tipo T_i .
- Una **Relazione** R di tipo $(A_1 : T_1, A_2 : T_2, \dots, A_n : T_n)$ è un insieme finito di n-uple di tipo $(A_1 : T_1, A_2 : T_2, \dots, A_n : T_n)$.
- La **cardinalità** di una relazione è il numero di n-uple

Esempio di relazione

Nome	Matricola	Indirizzo	Telefono
Mario Rossi	123456	Via Etnea 18	777777
Maria Bianchi	234567	Via Roma 2	888888
Giovanni Verdi	345678	Via Etnea 18	999999
Enzo Gialli	456789	Via Catania 3	444444

- E' una relazione di tipo
 $\{(\textit{Nome} : \textit{Stringa}, \textit{Matricola} : \textit{intero},$
 $\textit{Indirizzo} : \textit{Stringa}, \textit{Telefono} : \textit{intero})\}$
- Ciascuna riga è una n-upla della relazione.

Esempio di schema relazionale

```
{  
  Studenti : { (Nome : stringa, Matricola : intero(6),  
               Indirizzo : stringa, Telefono : intero(6)) }  
  
  Esami : { (Corso : stringa, Matricola : intero(6), Voto : {18,19,...,30}) }  
  
  Corsi : { (Corso : stringa, Professore : stringa) }  
}
```

- Si abbrevia (se non interessa riportare il tipo degli attributi) in:
 - Studenti (Nome, Matricola, Indirizzo, Telefono);
 - Esami (Corso, Matricola, Voto);
 - Corsi (Corso, Professore);

Dominio di un attributo

- $\text{dom}(A_i)$ è l'insieme dei possibili valori dell'attributo A_i
- Esempio: nella tabella

Esami(Corso, Matricola, Voto)

$\text{dom}(\text{Voto}) = \{18, 19, \dots, 30\}$

Vincoli di integrità

- I **vincoli di integrità** servono a migliorare la *qualità* delle informazioni contenute nella base di dati
- Un vincolo è un predicato che deve essere soddisfatto da ogni n-upla nella base di dati
- Un'**istanza valida** di uno schema di relazione è una relazione dello schema che soddisfa tutti i vincoli di integrità.

Esempio di vincoli d'integrità

- Il voto deve essere compreso tra 18 e 30
- La lode può apparire solo se voto =30
- Ogni studente deve avere un numero di matricola
- Il numero di matricola di uno studente deve essere univoco
- Esami dati devono fare riferimento solo a corsi offerti

Vincoli d'integrità

- I tre tipi più importanti specificano:

1. Quali attributi non possono assumere il valore **NULL**
2. Quali attributi sono **chiave**

Vincoli
Intrarelazionali

1. Quali attributi sono **chiavi esterne**

Vincolo
Interrelazionale

Esempio di valori nulli

Nome	Matricola	Indirizzo	Telefono
Mario Rossi	123456	Via Etnea 18	777777
Maria Bianchi	234567	Via Roma 2	888888
Giovanni Verdi	345678	NULL	999999
Enzo Gialli	456789	Via Catania 3	NULL

- Un insieme X di attributi di uno schema di relazione R è una **superchiave** dello schema se ogni istanza valida dello schema non contiene due n -uple distinte t_1 e t_2 con $t_1[X] = t_2[X]$
- Una **chiave** è una superchiave minimale rispetto alla relazione \subseteq (non contiene un'altra superchiave).

Esempi di chiavi

- La coppia {Nome, Matricola} è una superchiave ma non è una chiave, infatti
- {Matricola} è una chiave

Nome	Matricola	Indirizzo	Telefono
Mario Rossi	123456	Via Etnea 18	777777
Maria Bianchi	234567	Via Roma 2	888888
Giovanni Verdi	345678	Via Etnea 18	999999
Enzo Gialli	456789	Via Catania 3	444444

Un'altra chiave

- L'attributo {Nome} è una chiave?
- Sì, in questa istanza..
- Ma è sempre vero?

Nome	Matricola	Indirizzo	Telefono
Mario Rossi	123456	Via Etnea 18	777777
Maria Bianchi	234567	Via Roma 2	888888
Giovanni Verdi	345678	Via Etnea 18	999999
Enzo Gialli	456789	Via Catania 3	444444

Esistenza delle chiavi

- Una relazione non può contenere n-uple distinte ma uguali
- Ogni relazione ha come superchiave l'insieme degli attributi su cui è definita e quindi ha (almeno) una chiave

Importanza delle chiavi

- L'esistenza delle chiavi garantisce l'accessibilità a ciascun dato della base di dati
- Le chiavi permettono di correlare i dati in relazioni diverse:
 - Il modello relazionale è basato su valori

Chiavi e valori nulli

- In presenza di valori nulli, i valori della chiave non permettono:
 - di identificare le n-uple
 - di realizzare facilmente i riferimenti da altre relazioni

Chiave primaria

- La **Chiave Primaria** è una delle chiavi scelta per un dato schema.
- Non ammette valori NULL
- Notazione: **Sottolineatura**

Nome	<u>Matricola</u>	Indirizzo	Telefono
Mario Rossi	123456	Via Etnea 18	777777
Maria Bianchi	234567	Via Roma 2	888888
Giovanni Verdi	345678	Via Etnea 18	999999
Enzo Gialli	456789	Via Catania 3	444444

Chiavi esterne

- Un insieme di attributi $\{A_1, A_2, \dots, A_n\}$ di uno schema di relazione R è una **chiave esterna** che riferisce una *chiave primaria* $\{B_1, B_2, \dots, B_n\}$ di un altro schema S se in ogni istanza valida della base di dati, per ogni n -upla r dell'istanza di R esiste una n -upla s (*riferita da r*) dell'istanza di S tale che $r.A_i = s.B_i$ per $i=1, \dots, n$.

Esempio di chiave esterna

Notazione chiave esterna:
Sottolineatura tratteggiata

Esami

Corso	<u>Matricola</u>	Voto
Programmazione 1	345678	27
Architettura	123456	30
Matematica discreta	234567	19
Basi di Dati	345678	28

Studenti

Nome	<u>Matricola</u>	Indirizzo	Telefono
Mario Rossi	123456	Via Etnea 18	777777
Maria Bianchi	234567	Via Roma 2	888888
Giovanni Verdi	345678	Via Etnea 18	999999
Enzo Gialli	456789	Via Catania 3	444444

Esempio di chiave esterna

Esami

<u>Corso</u>	<u>Matricola</u>	Voto
Programmazione 1	345678	27
Architettura	123456	30
Matematica discreta	234567	19
Algoritmi	345678	28

Corsi

<u>Corso</u>	Professore
Architettura	Barbanera
Programmaizione 1	Cincotti
Matematica discreta	Milici
Algoritmi	Cantone

Esercizi

1. Individuare le chiavi e i vincoli di integrità referenziale che sussistono nella seguente base di dati. Individuare anche quali attributi possano ammettere valori nulli.

Pazienti

<u>Cod.</u>	Cognome	Nome
A102	Necchi	Luca
B372	Rossini	Piero
B543	Missoni	Nadia
B444	Missoni	Luigi
S555	Rossetti	Gino

Ricoveri

<u>Paz.</u>	<u>Inizio</u>	Fine	<u>Rep.</u>
A102	2/5/11	9/5/11	A
A102	2/12/11	2/1/12	A
S555	25/4/11	3/5/11	B
B444	1/12/11	2/1/12	B
S555	5/10/11	1/11/11	A

Medici

<u>Matr.</u>	<u>Cogn.</u>	<u>Nome</u>	<u>Rep.</u>
203	Neri	Piero	A
574	Bisi	Mario	B
461	Bargio	Sergio	B
530	Belli	Nicola	C
405	Mizzi	Nicola	A
501	Monti	Mario	A

Reparti

<u>Cod.</u>	<u>Nome</u>	<u>Primario</u>
A	Chirurgia	203
B	Pediatria	574
C	Medicina	530

1. Definire uno schema relazionale per organizzare le informazioni di un'azienda che ha impiegati (ognuno con codice fiscale, cognome, nome e data di nascita), filiali (con codice, sede e direttore, che è un impiegato). Ogni lavoratore lavora presso una filiale. Indicare le chiavi e i vincoli di integrità referenziale dello schema. Mostrare un'istanza della base di dati e verificare che soddisfa i vincoli.

1. Un albero genealogico rappresenta, in forma grafica, la struttura di una famiglia (o più famiglie, quando è ben articolato). Mostrare come si possa rappresentare, in una base di dati relazionale, un albero genealogico, cominciando eventualmente da una struttura semplificata, in cui si rappresentano solo le discendenze in linea maschile (cioè i figli vengono rappresentati solo per i componenti di sesso maschile) oppure solo quelle in linea femminile.