

Progettazione di una base di dati relazionale - Problemi e vincoli

Index

- [Introduzione](#)
 - [Ipotesi 1](#)
 - [Problemi](#)
 - [Anomalia di aggiornamento](#)
 - [Anomalia di inserimento](#)
 - [Anomalia di cancellazione](#)
 - [Ipotesi 2](#)
 - [Problemi](#)
 - [Anomalia di aggiornamento](#)
 - [Anomalia di inserimento](#)
 - [Anomalia di cancellazione](#)
 - [Ipotesi 3](#)
 - [Vincoli](#)
 - [Condizioni nella realtà di interesse](#)
 - [Vincoli sulla base di dati](#)
 - [Definizione e verifica dei vincoli nei DBMS](#)
 - [Dipendenze funzionali](#)
-

Introduzione

Supponiamo di voler creare una base di dati contenente i seguenti dati di studenti universitari:

Dati anagrafici e indicativi

- nome e cognome
- data, comune e provincia di nascita
- matricola

- codice fiscale

Dati curriculari

- per ogni esame sostenuto
 - voto
 - data
 - codice
 - titolo
 - docente del corso

Ipotesi 1

La base di dati consiste di una sola relazione con schema

Curriculum(Matn, CF, Cogn, Nome, DataN, ComN, ProvN, C#, TitC, DocC, DataE, Voto)

Curriculum	Matr	CF	Cogn	Nome	DataN	ComN	ProvN	C#	TitC	DocC	DataE	Voto
	01	...	Rossi	Mario	...	Tolfa	Roma	10	Fisica	Pippo
	02	...	Bianchi	Paolo	...	Tolfa	Roma	10	Fisica	Pippo
	01	...	Rossi	Mario	...	Tolfa	Roma	20	Chimica	Pluto

Problemi

I dati anagrafici di uno studente sono memorizzati per ogni esame sostenuto dallo studente e i dati di un corso sono memorizzati per ogni esame sostenuto per quel corso

La **ridondanza** dunque dà luogo a:

- spreco di spazio di memoria
- **anomalie** di aggiornamento, inserimento e cancellazione

Anomalia di aggiornamento

Se cambia il docente del corso il dato deve essere modificato per ogni esame sostenuto per quel corso

Anomalia di inserimento

Non posso inserire i dati anagrafici di uno finché non ha sostenuto almeno un esame a meno che di non usare valori nulli; idem per i corsi

Anomalia di cancellazione

Eliminando i dati anagrafici di uno studente potrebbero essere eliminati i dati di un corso (se lo studente è l'unico ad aver sostenuto l'esame di quel corso); idem quando elimino un corso

Ipotesi 2

La base di dati consiste di tre schemi di relazione:

- Studente(Mat_r, CF, Cogn, Nome, Data, Com, Prov)
- Corso(C#, Tit, Doc)
- Esame(Mat_r, C#, Data, Voto)

Studente	Matr	CF	Cogn	Nome	Data	Com	Prov	Corso	C#	Tit	Doc
	01	...	Rossi	Mario	...	Tolfa	Roma		10	Fisica	Pippo
	02	...	Bianchi	Paolo	...	Tolfa	Roma		20	Chimica	Pluto

Esame	Matr	C#	Data	Voto
	01	10
	02	10
	01	20

Problemi

Si nota della **ridondanza** in quanto il fatto che un comune si trova in una certa provincia è ripetuto per ogni studente nato in quel comune

Anomalia di aggiornamento

Se un comune cambia provincia (in seguito alla creazione di una nuova Provincia) devono essere modificate più tuple

Anomalia di inserimento

Non è possibile memorizzare il fatto che un certo comune si trova in una certa provincia se non c'è almeno uno studente nato in quel comune

Anomalia di cancellazione

Se vengono eliminati i dati anagrafici di uno studente potrebbe perdersi l'informazione che un certo comune si trova in una certa provincia (se è l'unico studente nato in quel comune)

Ipotesi 3

La base di dati consiste di quattro schemi di relazione:

- Studente (Matr, CF, Cogn, Nome, Data, Com)
- Corso (C#, Tit, Doc)
- Esame (Matr, C#, Data, Voto)
- Comune (Com, Prov)

Studente	Matr	CF	Cogn	Nome	Data	Com
	01	...	Rossi	Mario	...	Tolfa
	02	...	Bianchi	Paolo	...	Tolfa

Comune	Com	Prov
	Tolfa	Roma

Esame	Matr	C#	Data	Voto
	01	10
	02	10
	01	20

Corso	C#	Tit	Doc
	10	Fisica	Pippo
	20	Chimica	Pluto

Per progettare uno schema “buono” occorre rappresentare separatamente ogni concetto in una relazione distinta

Vincoli

Condizioni nella realtà di interesse

Nella realtà che si vuole rappresentare in una base di dati sono soddisfatte certe condizioni. Ad esempio:

1. Un voto è un intero compreso tra 18 e 30
2. Il numero di matricola identifica univocamente uno studente
3. Il numero di matricola in un verbale di esame deve essere il numero di matricola di uno studente
4. Lo stipendio di un impiegato non può diminuire
5. Lo straordinario è dato dal numero di ore fatte per la paga oraria

Vincoli sulla base di dati

Quando rappresentiamo una realtà di interesse in una base di dati deve essere possibile rappresentare anche tali condizioni.

Un **vincolo** è la rappresentazione nello schema di una base di dati di una condizione valida nella realtà di interesse.

Un'istanza della base di dati è **legale** se soddisfa tutti i vincoli (cioè se è una rappresentazione fedele della realtà).

Definizione e verifica dei vincoli nei DBMS

Un DBMS ci deve impedire di inserire tuple che violano i vincoli, ma viene fatto solo se questi vincoli sono espliciti (non le dipendenze funzionali). Però se è avvenuta una corretta progettazione il DBMS neanche deve intervenire in quanto questi vengono rispettati automaticamente

Dipendenze funzionali

Vedremo che le dipendenze funzionali definite su uno schema di relazione esprimono particolari vincoli di dipendenza tra sottoinsiemi di attributi dello schema stesso, che devono essere soddisfatti da ogni istanza dello schema