

Vincoli di integrità

Non tutte le istanze di basi di dati sintatticamente corrette rappresentano informazioni plausibili per l'applicazione di interesse

Studenti	Matricola	Nome	Nascita
	276545	Rossi	23-04-72
	276545	Bianchi	23-04-74
	784563	Verdi	30-07-74

Esami	Studente	Voto	Lode	Corso
	276545	28	SI	01
	276545	32	NO	01
	784563	23	NO	03
	200768	30	SI	03

Corso	Codice	Nome	Docente
	01	Analisi	Belli
	02	NULL	NULL
	03	Chimica	Viola

Definizione di vincolo di integrità

una proprietà che deve essere
soddisfatta dalla istanza della base
di dati

Ogni vincolo può essere visto come un *predicato* (funzione booleana) che associa ad ogni istanza il valore vero o falso. Se il predicato assume il valore vero diciamo che l'istanza *soddisfa* il vincolo

Ad uno schema di base di dati associamo un insieme di vincoli di integrità. Se l'istanza di base di dati soddisfa tutti i vincoli, allora essa è *corretta* (*lecita, valida, ammissibile*)

Tipi di vincoli

definito rispetto ad
una sola relazione?

coinvolge più
relazioni?

intrarelazionale

interrelazionale

può essere valutato
su una singola tupla
indipendentemente
dalle altre?

di tupla

può essere valutato
su un singolo
attributo?

di dominio
(o *su valori*)

Vincoli di tupla

Possibile sintassi: espressioni booleane, cioè con connettivi \wedge , \vee , e \neg (AND, OR e NOT), i cui atomi sono confronti (con operatori di uguaglianza, disuguaglianza e ordinamento) tra attributi, valori di attributo, e espressioni aritmetiche su valori di attributo

Esempio: nella relazione Esami

$$(\text{Voto} \geq 18) \wedge (\text{Voto} \leq 30)$$

$$(\neg (\text{Lode} = \text{SI})) \vee (\text{Voto} = 30)$$

Altro esempio: nella relazione Pagamenti(Data, Importo, Ritenute, Netto)

$$\text{Netto} = \text{Importo} - \text{Ritenute}$$

Valori che identificano le ennuple

Studenti	Matricola	Nome	Nascita
	276545	Rossi	23-04-72
	045683	Bianchi	23-04-74
	784563	Verdi	30-07-74

- il numero di matricola identifica gli studenti \Rightarrow non ci possono essere due tuple con lo stesso valore del numero di matricola
- il nome e la data di nascita identificano gli studenti \Rightarrow non ci possono essere due tuple con lo stessi valori sui due attributi

Attenzione: anche numero di matricola e nome identificano gli studenti, e non ci possono essere due tuple con lo stessi valori sui due attributi

Vincoli di chiave

Sono vincoli intrarelazionali ma non di tupla

Superchiave: un insieme K di attributi che non può contenere due tuple t_1 e t_2 con $t_1[K] = t_2[K]$

Chiave: una superchiave che non contiene propriamente un'altra superchiave (cioè: una superchiave minimale)

Esempio

Studenti	Matricola	Cognome	Nome	AA	Nascita
	0345623	Rossi	Luigi	1°	24-11-73
	0435234	Rossi	Maria	2°	10-10-71
	6834623	Verdi	Paolo	1°	30-02-72
	9872341	Bianchi	Maria	2°	02-12-74

Matricola è una chiave:

- Matricola è superchiave
- contiene un solo attributo quindi è minimale

Cognome, Nome e Nascita è una chiave:

- è superchiave perché identifica la tupla (non ci possono essere due studenti con gli stessi valori)
- nessuno dei suoi sottoinsiemi è superchiave

Individuazione delle chiavi

Consiste in un **trasferimento di vincoli** dalla realtà di interesse allo schema di base di dati che vuole rappresentarla

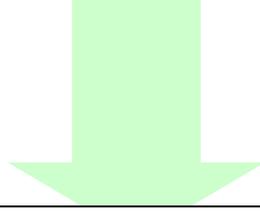
Ad uno schema associamo un insieme di vincoli minimali e consideriamo **corrette** (lecite, valide, ammissibili) solo le istanze che soddisfano tali vincoli

*Nota: i vincoli sono **definiti** sullo schema e **soddisfatti** dall'istanza*

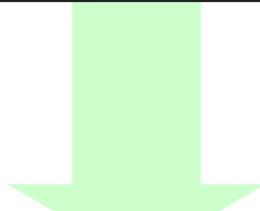
Attenzione: singole istanze di basi di dati possono soddisfare ulteriori vincoli per pura coincidenza

Esistenza delle chiavi

una relazione non può avere due tuple uguali tra loro



l'insieme di tutti gli attributi è per definizione una superchiave per lo schema di relazione



poiché il numero degli attributi è finito, ogni schema di relazione ha (almeno) una chiave

Uso di codici identificativi

Talvolta sarebbe necessario contemplare un eccessivo numero di attributi per garantire il fatto che non ci siano due tuple uguali in una relazione

Una soluzione efficace all'atto della definizione dello schema di base di dati è l'inserimento di un attributo aggiuntivo unico, che funga da chiave.

D'altronde molti codici identificativi (numero di matricola, codice fiscale, numero di targa, ecc.) sono stati introdotti appositamente per garantire l'identificazione dei soggetti del dominio (degli studenti, dei contribuenti, degli autoveicoli, ecc.)

Importanza delle chiavi

Garantiscono l'accessibilità a ciascun dato della base di dati.

Infatti ogni singolo valore è univocamente accessibile tramite:

- il nome della relazione
- il valore della chiave
- il nome dell'attributo

Le chiavi sono lo strumento principale attraverso il quale vengono correlati i dati in relazioni diverse (il modello relazionale è basato su valori)

Chiavi e valori nulli

In presenza di valori nulli, i valori degli attributi che formano la chiave

- non permettono di identificare le tuple come desiderato
- non consentono di realizzare facilmente i riferimenti da altre relazioni

Studenti	Matricola	Nome	Nascita
	NULL	Rossi	23-04-72
	045683	Bianchi	NULL
	NULL	Verdi	30-07-74

Chiave primaria

La presenza di valori nulli nelle chiavi deve essere limitata

Soluzione pratica: per ogni relazione scegliamo una chiave (*chiave primaria*) su cui i valori nulli non sono ammessi

Notazione: gli attributi della chiave primaria vengono sottolineati

Studenti	<u>Matricola</u>	Nome	Nascita
	054637	Rossi	23-04-72
	045683	Bianchi	NULL
	987326	Verdi	30-07-74

Vincoli di integrità referenziale (*foreign keys*)

Sono vincoli di integrità interrelazionali

Informazioni in relazioni diverse sono correlate attraverso valori comuni

Un **vincolo di integrità referenziale** fra un insieme di attributi X di uno schema di relazione R_1 e uno schema di relazione R_2 impone ai valori su X di ciascuna tupla di R_1 di comparire come valori della chiave primaria dell'istanza di R_2

*Sono definiti sugli schemi e
soddisfatti dalle istanze*

Foreign keys (considerazioni)

Nel caso di un solo attributo è un predicato completamente definito una volta specificato l'attributo della relazione R_1 e la relazione R_2

Nel caso generale di un insieme di attributi X occorrerebbe specificare anche:

- un ordinamento sugli attributi nell'insieme X
- un ordinamento sugli attributi della chiave primaria di R_2

Il vincolo è soddisfatto se per ogni tupla t_1 di R_1 esiste una tupla t_2 di R_2 tale che l' i -esimo valore della chiave primaria di R_2 è uguale all' i -esimo valore di $t_1[X]$

Esempio di chiave esterna

incidenti

<u>codice</u>	prov1	num1	prov2	num2
00001	RM	034524	PG	982453
00002	BO	823507	MI	827283

autoveicoli

<u>prov</u>	<u>num</u>	proprietario	...
RM	034524	Bianchi Mario	...
PG	982453	Rossi Ugo	...
BO	823507	Verdi Paola	...
MI	827283	Gialli Antonio	...

*non si può contare sulla
coincidenza dei nomi degli
attributi*

Infrazioni

<u>Codice</u>	Data	Vigile	Prov	Numero
0345623	04-10-00	230	PG	2353465
0435234	05-10-00	230	RM	4394936
6834623	05-10-00	134	PG	2989493
9872341	06-10-00	230	PG	3542647

Vigili

<u>Matricola</u>	Cogn.	Nome
230	Rossi	Mario
134	Bianchi	Luigi
219	Verdi	Antonio

Automobili

<u>Numero</u>	<u>Prov</u>	CC	...
2353465	PG	1000	...
4394936	RM	1100	...
2989493	PG	2000	...
3542647	PG	1000	...