

Troppi valori nulli

studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	NULL
9283	Verdi	Luisa	12/11/1979
NULL	Rossi	Maria	01/02/1978

esami

Studente	Voto	Corso
NULL	30	NULL
NULL	24	02
9283	28	01

corsi

Codice	Titolo
01	Analisi
02	NULL
04	Chimica

Tipi di valore nullo

- Il valore nullo può essere impiegato in tre casi differenti:
 - valore **sconosciuto** (esiste ma non lo conosco)
 - valore **inesistente** (so che non esiste)
 - valore **senza informazione** (non so se esiste o no)
- Tuttavia, i DBMS non distinguono i tipi di valore nullo

Una base di dati "scorretta"

- Alcune istanze di relazioni contengono dati senza senso per il contesto di interesse
- È necessario avere strumenti per **vincolare** i dati a valori sensati

Esami	Studente	Voto	Lode	Corso
	276545	32		01
	276545	30	e lode	02
	787643	27	e lode	03
	739430	24		04

Studenti	Matricola	Cognome	Nome
	276545	Rossi	Mario
	787643	Neri	Piero
	787643	Bianchi	Luca

Vincolo di integrità

- Proprietà che deve essere soddisfatta dalle istanze per rappresentare informazioni corrette relativamente ad un determinato contesto
- Un vincolo è una **funzione booleana** (un predicato) che associa ad ogni istanza il valore vero o falso

Vincoli di integrità

- Ad una base dati possono essere associati un insieme di vincoli
- In questo caso saranno corrette le istanze della base dati che soddisfano TUTTI i vincoli

Tipi di vincoli

- vincoli **intrarelazionali**
 - vincoli di **dominio**
 - vincoli di **ennupla**
 - vincoli di **chiave**
- vincoli **interrelazionali**
 - vincoli di **integrità referenziale**

Vincoli di integrità, perché?

- Descrizione più accurata della realtà
- Contributo alla “qualità dei dati”
- Utili nella progettazione
- Usati dai DBMS nella esecuzione delle interrogazioni

Dati inconsistenti

Esami	Studente	Voto	Lode	Corso
	276545	32		01
	276545	30	e lode	02
	787643	27	e lode	03
	739430	24		04

Studenti	Matricola	Cognome	Nome
	276545	Rossi	Mario
	787643	Neri	Piero
	787643	Bianchi	Luca

Vincoli di ennupla

- Esprimono condizioni sui valori di ciascuna ennupla di una relazione, indipendentemente dalle altre ennuple
- Caso particolare:
 - Vincoli di dominio: coinvolgono un solo attributo

Sintassi ed esempi

- Sintassi:
 - espressione booleana di condizioni atomiche che confrontano valori di attributi o espressioni aritmetiche su di essi

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

$(\text{Voto} = 30) \text{ OR NOT } (\text{Lode} = \text{"e lode"})$

Identificazione delle ennuple

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Ele	5/12/78
78763	Rossi	Mario	Ing Inf	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	3/11/76
67653	Rossi	Piero	Ing Mecc	5/12/78

- non ci sono due ennuple con lo stesso valore sull'attributo Matricola
- non ci sono due ennuple uguali su tutti e tre gli attributi Cognome, Nome e Data di Nascita

Chiave

- insieme di attributi che identificano le ennuple di una relazione

Formalmente:

- un insieme K di attributi è **superchiave** per r se r non contiene due ennuple distinte t_1 e t_2 con $t_1[K] = t_2[K]$
- K è **chiave** per r se è una superchiave minimale per r (cioè non contiene un'altra superchiave)

Chiave

- Una generica relazione può avere più chiavi:
 - anche insiemi di attributi disgiunti

Una chiave

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Inf	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	3/11/76
67653	Rossi	Piero	Ing Mecc	5/12/78

- Matricola è una chiave:
 - è superchiave
 - contiene un solo attributo e quindi è minimale

Un'altra chiave

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Inf	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	3/11/76
67653	Rossi	Piero	Ing Mecc	5/12/78

- Cognome, Nome, Nascita è un'altra chiave:
 - è superchiave
 - minimale

Un'altra chiave

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Civile	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	3/11/76
67653	Rossi	Piero	Ing Mecc	5/12/78

- Non ci sono ennuple uguali su Cognome e Corso:
 - Cognome e Corso formano una chiave
- Ma è sempre vero?

Studenti

Matricola	Cognome	Nome	Corso	Nascita
-----------	---------	------	-------	---------

- chiavi:

Matricola

Cognome, Nome, Nascita

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Civile	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	3/11/76
67653	Rossi	Piero	Ing Mecc	5/12/78

- È corretta: soddisfa i vincoli
- Ne soddisfa anche altri ("per caso"):
 - **Cognome, Corso** è chiave

Esistenza delle chiavi

- Data una relazione, questa può non avere chiavi?

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:
 - modello basato su valori

Importanza delle chiavi

- **Attenzione:** negli esempi precedenti abbiamo verificato il vincolo di chiave e superchiave facendo riferimento all'istanza di una tabella
- In realtà i vincoli di chiave e superchiave esprimono una **regola che deve valere a livello di schema**, sulla base della semantica degli attributi coinvolti

Chiavi e valori nulli

- La presenza di valori nulli nei valori della chiave rende critica
 - Identificazione delle ennuple
 - Definizione di corrispondenze tra ennuple in relazioni diverse

Matricola	Cognome	Nome	Corso	Nascita
NULL	NULL	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Civile	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	NULL
NULL	Neri	Mario	NULL	5/12/78

- La presenza di valori nulli nelle chiavi deve essere limitata

Chiave primaria

- Chiave su cui non sono ammessi valori nulli
- Notazione: sottolineatura

<u>Matricola</u>	Cognome	Nome	Corso	Nascita
86765	NULL	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Civile	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	NULL
43289	Neri	Mario	NULL	5/12/78

Integrità referenziale

- informazioni in relazioni diverse sono correlate attraverso valori comuni
- in particolare, valori delle chiavi (primarie)
- le correlazioni debbono essere "coerenti"

Infrazioni

<u>Codice</u>	Data	Vigile	Prov	Numero
34321	1/2/95	3987	MI	39548K
53524	4/3/95	3295	TO	E39548
64521	5/4/96	3295	PR	839548
73321	5/2/98	9345	PR	839548

Vigili

<u>Matricola</u>	Cognome	Nome
3987	Rossi	Luca
3295	Neri	Piero
9345	Neri	Mario
7543	Mori	Gino

Infrazioni

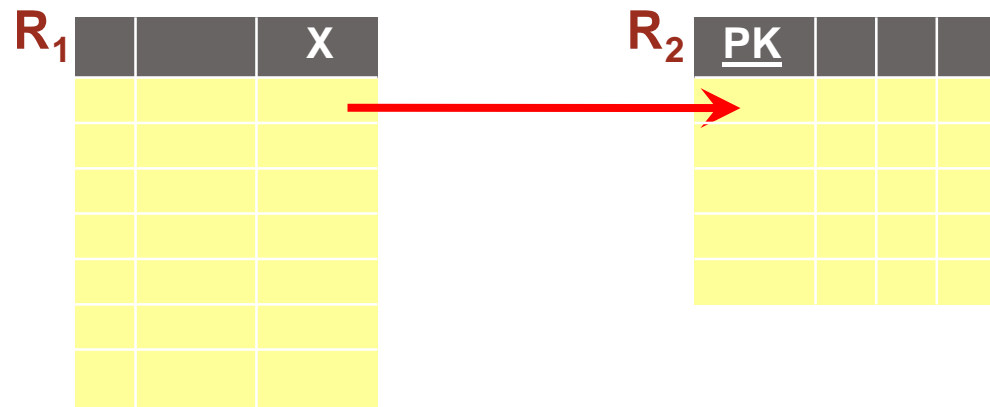
<u>Codice</u>	Data	Vigile	Prov	Numero
34321	1/2/95	3987	MI	39548K
53524	4/3/95	3295	TO	E39548
64521	5/4/96	3295	PR	839548
73321	5/2/98	9345	PR	839548

Auto

<u>Prov</u>	<u>Numero</u>	Cognome	Nome
MI	39548K	Rossi	Mario
TO	E39548	Rossi	Mario
PR	839548	Neri	Luca

Vincolo di integrità referenziale

- Un vincolo di **integrità referenziale** (**Foreign Key**) fra gli attributi X di una relazione R_1 e un'altra relazione R_2 impone ai valori su X in R_1 di comparire come valori della Primary Key (PK) di R_2 (oppure di avere valore NULL)
 - Può verificarsi che alcuni valori presenti nella PK di R_2 non siano presenti in X ?
 - Può verificarsi che valori presenti in X non siano presenti nella PK di R_2 ?



- R_1 : **tabella interna**
- R_2 : **tabella esterna**

Vincolo di integrità referenziale

- vincoli di integrità referenziale fra:
 - l'attributo Vigile della relazione INFRAZIONI e la relazione VIGILI
 - gli attributi Prov e Numero di INFRAZIONI e la relazione AUTO

Vigili (matricola, cognome, nome)

Infrazioni (codice, data, vigile, prov, numero)

Auto (prov, numero, cognome, nome)



Vincolo di integrità referenziale

Infrazioni

<u>Codice</u>	Data	Vigile	Prov	Numero
34321	1/2/95	3987	MI	39548K
53524	4/3/95	3295	TO	E39548
64521	5/4/96	3295	PR	839548
73321	5/2/98	9345	PR	839548

Auto

<u>Prov</u>	<u>Numero</u>	Cognome	Nome
MI	E39548	Rossi	Mario
TO	F34268	Rossi	Mario
PR	839548	Neri	Luca

Vincoli multipli su più attributi

Incidenti

<u>Codice</u>	Data	ProvA	NumeroA	ProvB	NumeroB
34321	1/2/95	TO	E39548	MI	39548K
64521	5/4/96	PR	839548	TO	E39548

Auto

<u>Prov</u>	<u>Numero</u>	Cognome	Nome
MI	39548K	Rossi	Mario
TO	E39548	Rossi	Mario
PR	839548	Neri	Luca

- vincoli di integrità referenziale fra:
 - gli attributi ProvA e NumeroA di INCIDENTI e la relazione AUTO
 - gli attributi ProvB e NumeroB di INCIDENTI e la relazione AUTO
- L'ordine degli attributi è significativo

Riassunto

- Valori nulli e nuova definizione di ennupla
- I vincoli di integrità
 - Vincoli intra-relazionali: di dominio, ennupla e chiave
 - Esistenza della chiave
 - Chiave primaria
 - Vincolo di integrità referenziale (tabella interna ed esterna)

Interna

		X
		NULL

Esterna

<u>PK</u>			



Operazioni di aggiornamento e violazione dei vincoli

- Un DBMS relazionale prevede due classi

Le operazioni di ricerca non alterano l'istanza della base dati mentre le operazioni di aggiornamento sì: possono portare a violazione

- Operazioni di ricerca (querying)

Istruzioni per l'estrazione di dati che rispettano determinati requisiti (voto=30)

- Operazioni di aggiornamento (update)

Istruzioni per l'inserimento, la cancellazione e modifica di tuple

Operazioni di aggiornamento e violazione dei vincoli

- Inserimento:
 - È usato per inserire nuove tuple in una relazione
- Cancellazione:
 - È usata per cancellare tuple da una relazione
- Modifica:
 - È usato per cambiare il valore di alcuni attributi nelle tuple di una relazione
- Le operazioni di aggiornamento devono rispettare i vincoli di integrità specificati nello schema della base dati...

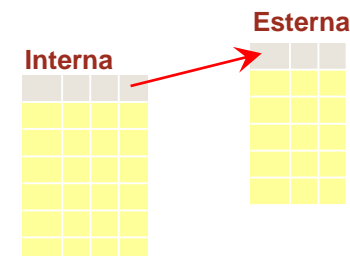
Operazioni di aggiornamento e violazione dei vincoli

- Inserimento:
 - È usato per inserire nuove tuple in una relazione
- Cancellazione:
 - È usata per cancellare tuple da una relazione
- Modifica:
 - È usato per cambiare il valore di alcuni attributi nelle tuple di una relazione
- Le operazioni di aggiornamento devono rispettare i vincoli di integrità specificati nello schema della base dati. E se questo non succede?

```
INSERT INTO voti(matricola, voto, esame)  
VALUES('m0328', 31, 'fisica1')
```

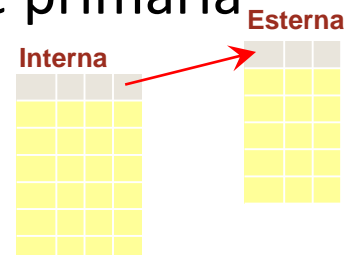
Operazioni di inserimento e violazione dei vincoli

- Nella operazione di **inserimento** vengono elencati i valori degli attributi di una nuova tupla t che deve essere inserita in una relazione R .
- L'operazione di inserimento può violare tutti i vincoli di integrità che abbiamo visto:
 - dominio/ennupla
 - chiave
 - chiave primaria
 - integrità referenziale



Operazioni di inserimento e violazione dei vincoli

- dominio/ennupla
 - Se viene inserito un valore di attributo che non è compatibile con il requisito del vincolo
- chiave
 - Se viene inserita una tupla che ha nella chiave un valore già esistente in una delle tuple della relazione
- chiave primaria
 - Se viene inserita una tupla che ha nella chiave un valore nullo
- integrità referenziale
 - Se viene inserita una tupla che ha negli attributi vincolati un valore che non compare tra quelli della chiave primaria nella tabella esterna

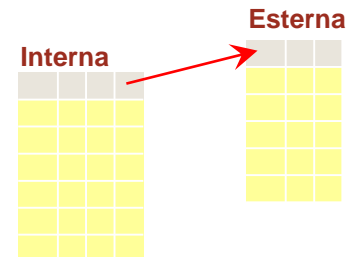


Operazioni di cancellazione e violazione dei vincoli

- Nella operazione di **cancellazione** ...

Operazioni di cancellazione e violazione dei vincoli

- Nella operazione di **cancellazione** si può violare solo il vincolo di integrità referenziale:



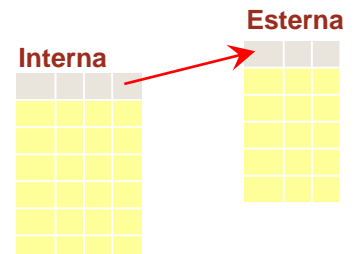
Se da una tabella esterna si cancella una tupla il cui valore della chiave primaria compare in almeno una delle tuple della tabella interna
Nota: la cancellazione di una tupla della tabella interna non viola alcun vincolo

Operazioni di modifica e violazione dei vincoli

- Nella operazione di **modifica** vengono elencati i nuovi valori degli attributi di un insieme di tuple:
 - Riduci del 10% il costo di tutti i televisori al plasma
- L'operazione di modifica può violare tutti i vincoli di integrità che abbiamo visto:
 - dominio/ennupla
 - chiave
 - chiave primaria
 - integrità referenziale

Operazioni di modifica e violazione dei vincoli

- dominio/ennupla
 - Se il nuovo valore di un attributo che non è compatibile con il requisito del vincolo
- chiave
 - Se il nuovo valore di un attributo chiave è già esistente in una delle tuple della relazione
- chiave primaria
 - Se il nuovo valore di un attributo chiave primaria è nullo
- integrità referenziale
 - Se in una tabella interna viene modificata una tupla mettendo, agli attributi vincolati, dei nuovi valori che non compaiono tra quelli della chiave primaria nella tabella esterna...
 - Se in una tabella esterna si modifica il valore della chiave primaria di una tupla era presente in almeno una delle tuple della tabella interna



Operazioni e violazione dei vincoli

- Quale è la risposta del DBMS a fronte della richiesta di esecuzione di una operazione che viola un vincolo di integrità ?

Operazioni e violazione dei vincoli

- Quale è la risposta del DBMS a fronte della richiesta di esecuzione di una operazione che viola un vincolo di integrità ?
 - Bloccare l'esecuzione dell'operazione notificando un messaggio d'errore
 - Eseguire l'operazione gestendo la violazione del vincolo con opportune azioni compensative...

Azioni compensative

- Esempio:
 - L'eliminazione di una ennupla dalla tabella esterna può causare una violazione del vincolo di integrità referenziale
- Azioni
 - Rifiuto dell'operazione
 - Eliminazione in cascata
 - Introduzione di valori nulli

Eliminazione in cascata

Quale delle due è la tabella esterna?

Impiegati

<u>Matricola</u>	Cognome	Progetto
34321	Rossi	IDEA
53524	Neri	XYZ
64521	Verdi	NULL
73032	Bianchi	IDEA

Progetti

<u>Codice</u>	Inizio	Durata	Costo
IDEA	01/2000	36	200
XYZ	07/2001	24	120
BOH	09/2001	24	150

Introduzione di valori nulli

Impiegati

<u>Matricola</u>	Cognome	Progetto
34321	Rossi	IDEA
53524	Neri	NULL
64521	Verdi	NULL
73032	Bianchi	IDEA

Progetti

<u>Codice</u>	Inizio	Durata	Costo
IDEA	01/2000	36	200
XYZ	07/2001	24	120
BOH	09/2001	24	150

Esercizio

- Definire uno schema di base dati per organizzare le informazioni di una azienda che ha impiegati (ognuno con codice fiscale, nome, cognome e data di nascita), filiali (con codice, nome sede e direttore, che è un impiegato). Ogni impiegato lavora presso una o più filiali. Ogni filiale ha un solo direttore. Indicare le chiavi ed i vincoli di integrità referenziale dello schema.

Bastano 2 tabelle?

Esercizio 1

- Rappresentare per mezzo di una o più relazioni le informazioni contenute nell'orario delle partenze dei treni da una stazione ferroviaria: numero treno, categoria, stazione di partenza, di destinazione finale, fermate intermedie, nome di ciascuna fermata ed orario di arrivo e partenza.
Ciascun treno (num. treno) non può passare lo stesso giorno due volte dalla stessa fermata.

Soluzione ?

Treno(numero, categoria, part, orP, s1, or1, s2, or2, ...,dest)

Stazione(codice, descrizione)

