



# Atzeni, Ceri, Fraternali, Paraboschi, Torlone Basi di dati Quarta edizione McGraw-Hill, 2013

Capitolo 2:

IL MODELLO RELAZIONALE



## I modelli logici dei dati



- Tre modelli logici tradizionali
  - gerarchico
  - reticolare
  - relazionale
- Più recenti
  - a oggetti (poco diffuso)
  - basato su XML ("complementare" a quello relazionale)



## Modelli logici, caratteristiche



- Gerarchico e reticolare
  - utilizzano riferimenti espliciti (puntatori) fra record
- Relazionale "è basato su valori"
  - anche i riferimenti fra dati in strutture (relazioni) diverse sono rappresentati per mezzo dei valori stessi



studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978



esami	Studente	Voto	Corso
	3456	30	04
	3456	24	02
	9283	28	01
	6554	26	01

corsi	Codice	Titolo	Docente
	01	Analisi	Mario
	02	Chimica	Bruni
	04	Chimica	Verdi



#### Il modello relazionale



- Proposto da E. F. Codd nel 1970 per favorire l'indipendenza dei dati
- Disponibile in DBMS reali nel 1981 (non è facile implementare l'indipendenza con efficienza e affidabilità!)
- Si basa sul concetto matematico di relazione (con una variante)
- Le relazioni hanno naturale rappresentazione per mezzo di tabelle



#### 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 Entity-Relationship; tradotto anche con associazione o correlazione



## Relazione matematica, esempio



• 
$$D_1 = \{a,b\}$$

• 
$$D_2 = \{x, y, z\}$$

prodotto cartesiano D<sub>1</sub> × D<sub>2</sub>

a x y a z b x b y b z

una relazione

$$r \subseteq D_1 \times D_2$$

a xa zb y



#### Relazione matematica



- D<sub>1</sub>, ..., D<sub>n</sub> (n insiemi anche non distinti)
- prodotto cartesiano D<sub>1</sub>×...×D<sub>n</sub>:
  - l'insieme di tutte le n-uple (d₁, ..., dₙ) tali
    che d₁∈D₁, ..., dₙ ∈ Dₙ
- relazione matematica su D₁, ..., Dₙ:
  - un sottoinsieme di D₁×...×Dn.
- D<sub>1</sub>, ..., D<sub>n</sub> sono i domini della relazione



## Relazione matematica, proprietà



- una relazione matematica è un insieme di n-uple ordinate:
  - $(d_1, ..., d_n)$  tali che  $d_1 \in D_1, ..., d_n \in D_n$
- una relazione è un insieme:
  - 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 ⊆ string × string × int × int

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

- Ciascuno dei domini ha due 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



#### Tabelle e relazioni



- In una tabella che rappresenta una relazione
  - l'ordinamento tra le righe è irrilevante
  - l'ordinamento tra le colonne è irrilevante
- Una tabella rappresenta una relazione se
  - le righe sono diverse fra loro
  - le intestazioni delle colonne sono diverse tra loro
  - i valori di ogni colonna sono fra loro omogenei



#### Il modello è basato su valori



 I riferimenti fra dati in relazioni diverse sono rappresentati per mezzo di valori dei domini che compaiono nelle ennuple



studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978



esami

Studente	Voto	Corso
3456	30	04
3456	24	02
9283	28	01
6554	26	01

corsi

Codice	Titolo	Docente
01	Analisi	Mario
02	Chimica	Bruni
04	Chimica	Verdi



#### **Alternativa**



 Altri modelli (sia quelli "storici", reticolare e gerarchico, sia quello a oggetti) prevedono riferimenti espliciti, gestiti dal sistema



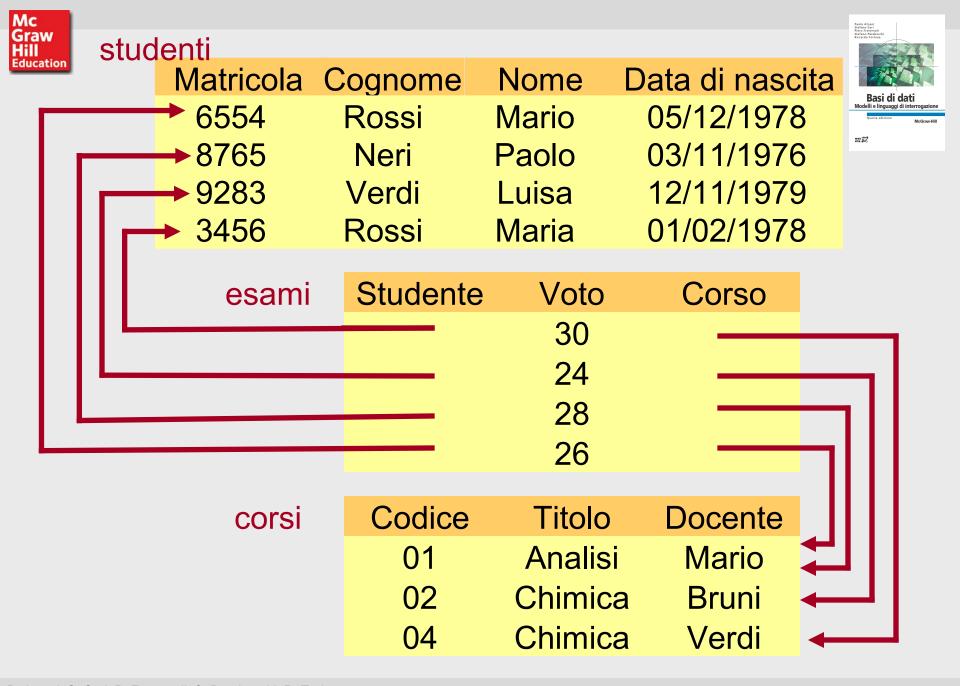
studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978



esami	Studente	Voto	Corso	
	3456	30	04	
	3456	24	02	
	9283	28	01	
	6554	26	01	

corsi	Codice	Titolo	Docente
	01	Analisi	Mario
	02	Chimica	Bruni
	04	Chimica	Verdi





# Struttura basata su valori: vantaggi



- indipendenza dalle strutture fisiche (si potrebbe avere anche con puntatori di alto livello) che possono cambiare dinamicamente
- si rappresenta solo ciò che è rilevante dal punto di vista dell'applicazione
- l'utente finale vede gli stessi dati dei programmatori
- i dati sono portabili più facilmente da un sistema ad un altro
- i puntatori sono direzionali



#### **Definizioni**



Schema di relazione:

un nome R con un insieme di attributi

$$A_1, ..., A_n$$

$$R(A_1,...,A_n)$$

 Schema di base di dati: insieme di schemi di relazione:

$$R = \{R_1(X_1), ..., R_k(X_k)\}$$



## Definizioni, 2



- Una ennupla su un insieme di attributi X è una funzione che associa a ciascun attributo A in X un valore del dominio di A
- t[A] denota il valore della ennupla t sull'attributo A



## Definizioni, 3



- (Istanza di) relazione su uno schema R(X): insieme r di ennuple su X
- (Istanza di) base di dati su uno schema R= {R<sub>1</sub>(X<sub>1</sub>), ..., R<sub>n</sub>(X<sub>n</sub>)}:
  insieme di relazioni r = {r<sub>1</sub>,..., r<sub>n</sub>} (con r<sub>i</sub> relazione su R<sub>i</sub>)



## Relazioni su singoli attributi



#### studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

studenti lavoratori

Matricola 6554 3456





	Da Filippo Via Roma 2, Roma					
	<i>Ricevuta Fiscale</i> 1235 <i>del</i> 12/10/2002					
3	Coperti	3,00				
2	Antipasti	6,20				
3	Primi	12,00				
2	Bistecche	18,00				
	<i>Totale</i> 39,20					

Da Filippo Via Roma 2, Roma							
	Ricevuta Fiscale 1240 del 13/10/2002						
2	Coperti	2,00					
2	Antipasti	7,00					
2	Primi	8,00					
2	Orate	20,00					
2	Caffè	2,00					
	<i>Totale</i> 39,00						





	Da Filippo Via Roma 2, Roma							
	Ricevuta Fiscale 1235 del 12/10/2002							
3	Coperti	3,00						
2	Antipasti	6,20						
3	Primi	12,00						
2	Bistecche	18,00						
	Totale 39,20							

	100 Dec					
	Da Filippo Via Roma 2, Roma					
	Ricevuta Fiscale 1240 del 13/10/2002					
2	Coperti	2,00				
2	Antipasti	7,00				
2	Primi	8,00				
2	Orate	20,00				
2	Caffè	2,00				
	Totale 39,00					





	Da Filippo Via Roma 2, Roma						
	Ricevuta Fiscale 1235 del 12/10/2002						
3	Coperti	3,00					
2	Antipasti	6,20					
3	Primi	12,00					
2	Bistecche	18,00					
	Totale 39,20						

Da Filippo Via Roma 2, Roma								
Ricevuta Fiscale 1240 del 13/10/2002								
2	Coperti	2,00						
2	Antipasti	7,00						
2	Primi	8,00						
2	Orate	20,00						
2	Caffè	2,00						
Totale 39,00								





#### Ricevute

Numero	Data	Qtà	Descrizione	Importo	Totale
1235	12/10/2002	3	Coperti	3,00	39,20
		2	Antipasti	6,20	
		3	Primi	12,00	
		2	Bistecche	18,00	
1240	13/10/2002	2	Coperti	2,00	39,00

• Ma i valori debbono essere semplici, non relazioni!



# Relazioni che rappresentano strutture nidificate



Ricevute	Num
	123

NumeroDataTotale123512/10/200239,20124013/10/200239,00

#### Dettaglio

Numero	Qtà	Descrizione	Importo
1235	3	Coperti	3,00
1235	2	Antipasti	6,20
1235	3	Primi	12,00
1235	2	Bistecche	18,00
1240	2	Coperti	2,00



## Strutture nidificate, riflessione



- Abbiamo rappresentato veramente tutti gli aspetti delle ricevute?
- Dipende da che cosa ci interessa!
  - possono esistere linee ripetute in una ricevuta?
    - Al bar, servizio al tavolo, ad un gruppo:
      - Cliente 1: "Una birra"
      - Cameriere: "Se volete altre birre, ditelo, altrimenti non posso aggiungerle"
  - l'ordine delle righe è rilevante?
- Sono possibili rappresentazioni diverse



# Rappresentazione alternativa per strutture nidificate



Ricevute	Numero	Data	Totale
	1235	12/10/2002	39,20
	1240	13/10/2002	39,00

LIOTTO	
<b>Dettag</b>	
Dottag	•••

Numero	Riga	Qtà	Descrizione	Importo
1235	1	3	Coperti	3,00
1235	2	2	Antipasti	6,20
1235	3	3	Primi	12,00
1235	4	2	Bistecche	18,00
1240	1	2	Coperti	2,00



## Informazione incompleta



- Il modello relazionale impone ai dati una struttura rigida:
  - le informazioni sono rappresentate per mezzo di ennuple
  - solo alcuni formati di ennuple sono ammessi: quelli che corrispondono agli schemi di relazione
- I dati disponibili possono non corrispondere al formato previsto







Nome	SecondoNome	Cognome
Franklin	Delano	Roosevelt
Winston		Churchill
Charles		De Gaulle
Josip		Stalin



#### Informazione incompleta: soluzioni?



- Non conviene (anche se spesso si fa) usare valori del dominio (0, stringa nulla, "99", ...):
  - potrebbero non esistere valori "non utilizzati"
  - valori "non utilizzati" potrebbero diventare significativi
  - in fase di utilizzo (nei programmi) sarebbe necessario ogni volta tener conto del "significato" di questi valori



# Informazione incompleta nel modello relazionale



- Tecnica rudimentale ma efficace:
  - valore nullo: denota l'assenza di un valore del dominio (ma non è un valore del dominio)
- t[A], per ogni attributo A, è un valore del dominio dom(A) oppure il valore nullo (che indichiamo qui con NULL)
- Si possono (e debbono) imporre restrizioni sulla presenza di valori nulli



## Tipi di valore nullo



- (Almeno) tre casi differenti
  - valore sconosciuto
  - valore inesistente
  - valore senza informazione
- I DBMS non distinguono i tipi di valore nullo



## Troppi valori nulli



	1		ı	4.5
C1	ш			nti
3	LU	IU		ııı

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
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	Docente
	01	Analisi	Mario
	02	NULL	NULL
	04	Chimica	Verdi



# Vincoli di integrità



 Esistono istanze di basi di dati che, pur sintatticamente corrette, non rappresentano informazioni possibili per l'applicazione di interesse



#### Una base di dati "scorretta"



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







- Proprietà che deve essere soddisfatta dalle istanze che rappresentano informazioni corrette per l'applicazione
- Un vincolo è una funzione booleana (un predicato): associa ad ogni istanza il valore vero o falso





# Vincoli di integrità, perché?

- descrizione più accurata della realtà
- contributo alla "qualità dei dati"
- utili nella progettazione (vedremo)
- usati dai DBMS nella esecuzione delle interrogazioni



# Vincoli di integrità, nota



- alcuni tipi di vincoli (ma non tutti) sono "supportati" dai DBMS:
  - possiamo quindi specificare vincoli di tali tipi nella nostra base di dati e il DBMS ne impedisce la violazione
- per i vincoli "non supportati", la responsabilità della verifica è dell'utente o del programmatore



# Tipi di vincoli



- vincoli intrarelazionali
  - vincoli su valori (o di dominio)
  - vincoli di ennupla
- vincoli interrelazionali





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, indipendentemente dalle altre ennuple
- Caso particolare:
  - Vincoli di dominio: coinvolgono un solo attributo



# Sintassi ed esempi



- Una possibile sintassi:
  - espressione booleana di atomi che confrontano valori di attributo o espressioni aritmetiche su di essi

(Voto 
$$\geq$$
 18) AND (Voto  $\leq$  30)







- "Se piove prendo l'ombrello"
  - Se non piove posso prenderlo o non prenderlo!
  - NOT piove OR prendo l'ombrello



# Vincoli di ennupla, altro esempio



Impiegato	Lordo	Ritenute	Netto
Rossi	55.000	12.500	42.500
Neri	45.000	10.000	35.000
Bruni	47.000	11.000	36.000

Lordo = (Ritenute + Netto)

Esempio





## Vincoli di ennupla, violazione

Stipendi	Impiegato	Lordo	Ritenute	Netto
	Rossi	55.000	12.500	42.500
	Neri	45.000	10.000	35.000

50.000 11.000

Lordo = (Ritenute + Netto)

Bruni

36.000



### Identificazione delle ennuple



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

- 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<sub>1</sub>
  e t<sub>2</sub> con t<sub>1</sub>[K] = t<sub>2</sub>[K]
- K è chiave per r se è una superchiave minimale per r (cioè non contiene un'altra superchiave)



#### **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?



#### Vincoli, schemi e istanze



- i vincoli corrispondono a proprietà del mondo reale modellato dalla base di dati
- interessano a livello di schema (con riferimento cioè a tutte le istanze)
- ad uno schema associamo un insieme di vincoli e consideriamo corrette (valide, ammissibili) le istanze che soddisfano tutti i vincoli
- un'istanza può soddisfare altri vincoli ("per caso")





#### 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



- Una relazione non può contenere ennuple 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

Esempio



#### Chiavi e valori nulli



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





 La presenza di valori nulli nelle chiavi deve essere limitata

Mario

**NULL** 

**NULL** 

Neri

5/12/78



# **Chiave primaria**



- Chiave su cui non sono ammessi 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

Esempio



# Integrità referenziale



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



## Infrazioni



Quarta edizione McGraw-H

elte (A)
----------

<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



wab 🎉

<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	Prov	<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<sub>1</sub> e un'altra relazione R<sub>2</sub> impone ai valori su X in R<sub>1</sub> di comparire come valori della chiave primaria di R<sub>2</sub>





- 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



# Violazione di vincolo di integrità referenziale



#### Infrazioni

Co	<u>dice</u>	Data	Vigile	Prov	Numero
34	321	1/2/95	3987	MI	39548K
53	524	4/3/95	3295	TO	E39548
64	521	5/4/96	3295	PR	839548
73	321	5/2/98	9345	PR	839548

Λ		1	_
Δ		H	$\bigcap$
$oldsymbol{\cap}$	u	LV	U

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



# Vincoli di integrità referenziale: commenti



- Giocano un ruolo fondamentale nel concetto di "modello basato su valori"
- In presenza di valori nulli i vincoli possono essere resi meno restrittivi



- Sono possibili meccanismi per il supporto alla loro gestione ("azioni" compensative a seguito di violazioni)
- Attenzione ai vincoli su più attributi





#### Integrità referenziale e valori nulli



#### **Impiegati**

<b>Matricola</b>	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





# **Azioni compensative**



- Esempio:
  - Viene eliminata una ennupla causando una violazione
- Comportamento "standard":
  - Rifiuto dell'operazione
- Azioni compensative:
  - Eliminazione in cascata
  - Introduzione di valori nulli



#### Eliminazione in cascata



**Impiegati** 

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

Progetti

<u>Codice</u>	Inizio	Durata	Costo
IDEA	01/2000	36	200
ВОН	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
ВОН	09/2001	24	150





# Vincoli multipli su più attributi



#### Incidenti

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

Λ		1	
$\Delta$		ы	$\cap$
$\Box$	u	ш	U

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



## Vincoli multipli su più attributi, 2



- 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