

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

RELAZIONI matematica: come nella teoria degli insiemi

- **Relazione matematica**: come nella teoria degli insiemi.

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, 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

Una **relazione matematica** è un insieme di n -uple ordinate:

(d_1, \dots, d_n) tali che $d_1 \in D_1, \dots, d_n \in D_n$

Oss: una relazione è un insieme; quindi:

- 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

$Partite \subseteq string \times string \times int \times int$

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**

Relazione matematica: struttura non posizionale

$$Partite \subseteq string \times string \times int \times int$$

A ciascun dominio si associa un nome (**attributo**), che ne descrive il "ruolo"

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

Una **tupla** su un insieme di attributi X è una funzione t che associa a ciascun attributo A un valore del suo dominio.

Una **relazione su X** è un insieme di tuple su X .

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

Il modello è basato su valori

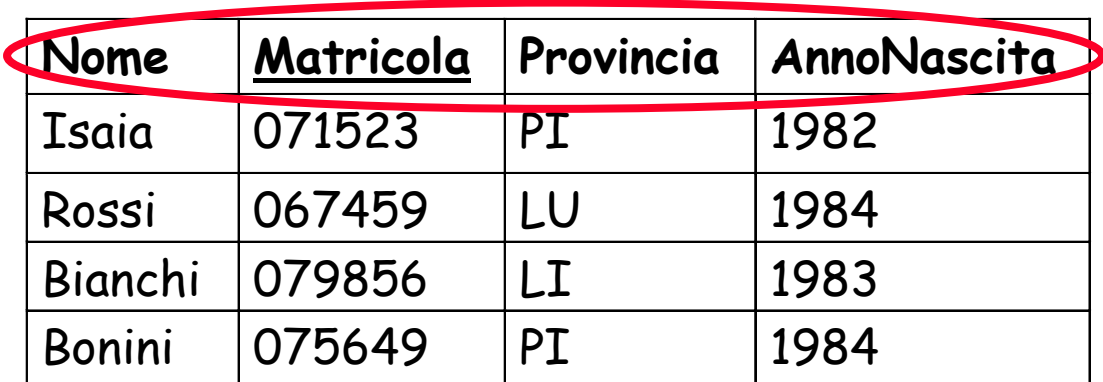
Il **modello relazionale** è basato su valori.

Ciò significa che i riferimenti fra dati in relazioni diverse sono rappresentati per mezzo di valori dei domini che compaiono nelle ennuple.

Schemi e istanze

- In ogni base di dati si distinguono:
 - lo **schema**, sostanzialmente invariante nel tempo, che ne descrive la struttura (aspetto intensionale)
 - le intestazioni delle tabelle
 - l'**istanza**, i valori attuali, che possono cambiare anche molto rapidamente (aspetto estensionale)
 - il "corpo" di ciascuna tabella

Studenti



Nome	<u>Matricola</u>	Provincia	AnnoNascita
Isaia	071523	PI	1982
Rossi	067459	LU	1984
Bianchi	079856	LI	1983
Bonini	075649	PI	1984

Schema di relazione

Istanza di
Relazione
o estensione
della relazione

IL MODELLO RELAZIONALE

- Definizione: I meccanismi per definire una base di dati con il modello relazionale sono l'**ennupla** e la **relazione**:
 - un **tipo ennupla** T è un insieme finito di coppie
(Attributo, Tipo elementare);
 - se T è un tipo ennupla, $R(T)$ è lo **schema della relazione R** ;
 - lo schema di una base di dati è un **insieme di schemi di relazione $R_i(T_i)$** ;
 - un'**istanza** di uno schema $R(T)$ è un insieme finito di ennuple di tipo T .
- Uguaglianza di due tipi ennupla, due ennuple, due tipi relazione

Vantaggi della struttura basata su valori

- **indipendenza dalle strutture fisiche** (si potrebbe avere anche con puntatori di alto livello) che possono cambiare dinamicamente. La rappresentazione logica dei dati (costituita dai soli valori) non fa riferimento a quella fisica
- si rappresenta solo ciò che è **rilevante** dal punto di vista dell'applicazione
- i dati sono **portabili** più facilmente da un sistema ad un altro
- i **puntatori sono direzionali**

Schema (riepilogo)

- Schema di relazione:

un nome R con un insieme T di attributi A_1, \dots, A_n :

$$R(T) = R(A_1, \dots, A_n)$$

- Schema di base di dati:

insieme di schemi di relazione:

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

Dove X_1, \dots, X_k sono insiemi di attributi

Tupla

- Una **tupla** su un insieme di attributi T , denotata con t , è una funzione che associa a ciascun attributo A in T un valore del dominio di A
- $t[A]$ denota il valore della tupla t sull'attributo $A \in T$
- $t[X]$ denota i valori della tupla t sugli attributi $X \in T$

$X = \{\text{Esame}, \text{Idoneità}\}$

Chi è $t[X]$?

Matricola	Esame	Voto	Idoneità
936463	Inglese	NULL	Si
936462	Basi di dati	30	No
...

Istanze

- Un'istanza di relazione o relazione su uno schema $R(X)$ è l'insieme r di tuple su X

- Un'istanza di base di dati su uno schema

$$R = \{R_1(X_1), \dots, R_n(X_n)\}$$

è l'insieme delle relazioni $r = \{r_1, \dots, r_n\}$ (con r_i relazione su R_i)

Relazione come una tabella bidimensionale

Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	9:45
Chimica	Nicola Mori	N1	9:45
Fisica I	Mario Bruni	N1	11:45
Fisica II	Mario Bruni	N3	9:45
Sistemi inform.	Piero Rossi	N3	8:00

Orario

Schema

Insegnamento	Docente	Aula	Ora
--------------	---------	------	-----

Istanza

Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	9:45
Chimica	Nicola Mori	N1	9:45
Fisica I	Mario Bruni	N1	11:45
Fisica II	Mario Bruni	N3	9:45
Sistemi inform.	Piero Rossi	N3	8:00

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

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

Valori nulli

Informazione incompleta nel modello relazionale

- **valore nullo**: denota l'assenza di un valore del dominio (e non è un valore del dominio)
- $t[A]$, per ogni attributo A , è un valore del dominio $\text{dom}(A)$ oppure il valore nullo **NULL**
- Si possono (e devono) imporre restrizioni sulla presenza di valori nulli

IdPersona	Stato civile	Coniuge
936463	celibe	NULL
936462	sposato	936465

Matricola	Esame	Voto	Idoneità
936463	Inglese	NULL	Si
936462	Basi di dati	30	NULL

Troppi valori nulli

studenti	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	Chimica	NULL
	NULL	Chimica	Verdi

Vincoli di integrità

Esistono istanze di basi di dati che, pur **sintatticamente corrette**, non rappresentano informazioni possibili per l'applicazione di interesse e che quindi generano **informazioni senza significato**.

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 integrità

- uno **schema relazionale** è costituito da un insieme di **schemi di relazione** e da un **insieme di vincoli** d'integrità sui possibili valori delle estensioni delle relazioni.
- Un **vincolo d'integrità** è una proprietà che deve essere soddisfatta dalle istanze che rappresentano informazioni corrette per l'applicazione.
- Un vincolo è espresso mediante 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
- Non tutte le proprietà di interesse sono rappresentabili per mezzo di vincoli formulabili in modo esplicito

Tipi di vincoli

- vincoli **intrarelazionali**:
 - sono i vincoli che devono essere rispettati dai valori contenuti nella relazione considerata
 - vincoli **su valori** (o **di dominio**)
 - vincoli **di ennupla**
- vincoli **interrelazionali**:
 - sono i vincoli che devono essere rispettati da valori contenuti in relazioni diverse

Vincoli di ennupla

- I **Vincoli di ennupla** esprimono condizioni sui valori di ciascuna ennupla, indipendentemente dalle altre ennuple
- Caso particolare:
 - **Vincoli di dominio**: coinvolgono un solo attributo

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

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

Stipendi	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

$\text{Lordo} = (\text{Ritenute} + \text{Netto})$

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

- Informalmente:
- Una **chiave** è un 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)

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*
- è *minimale* (in questo caso contiene un solo attributo)

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 possibili)
- 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

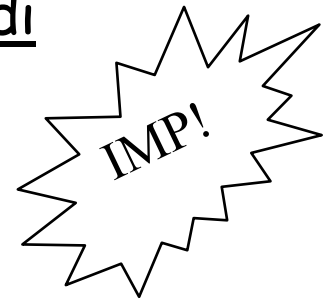
Cognome, Nome, Nascita

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 con valori uguali (una relazione è un sottoinsieme del prodotto cartesiano)
- Ogni relazione ha **sicuramente** come **superchiave** l'insieme di tutti gli attributi su cui è definita
- e quindi ogni relazione **ha (almeno) una chiave**
- 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

La presenza di valori nulli fra i valori di una chiave non permette

- di identificare le ennuple
- di realizzare facilmente i riferimenti da altre relazioni

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

Chiave primaria

- Una **chiave primaria** è una 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

- Nel modello relazionale le informazioni in relazioni diverse sono correlate attraverso valori comuni
- in particolare, vengono spesso presi in considerazione i valori delle chiavi (primarie).
- le correlazioni debbono essere "coerenti"

Il modello relazionale

Studenti

Nome	<u>Matricola</u>	Provincia	AnnoNascita
Isaia	071523	PI	1982
Rossi	067459	LU	1984
Bianchi	079856	LI	1983
Bonini	075649	PI	1984

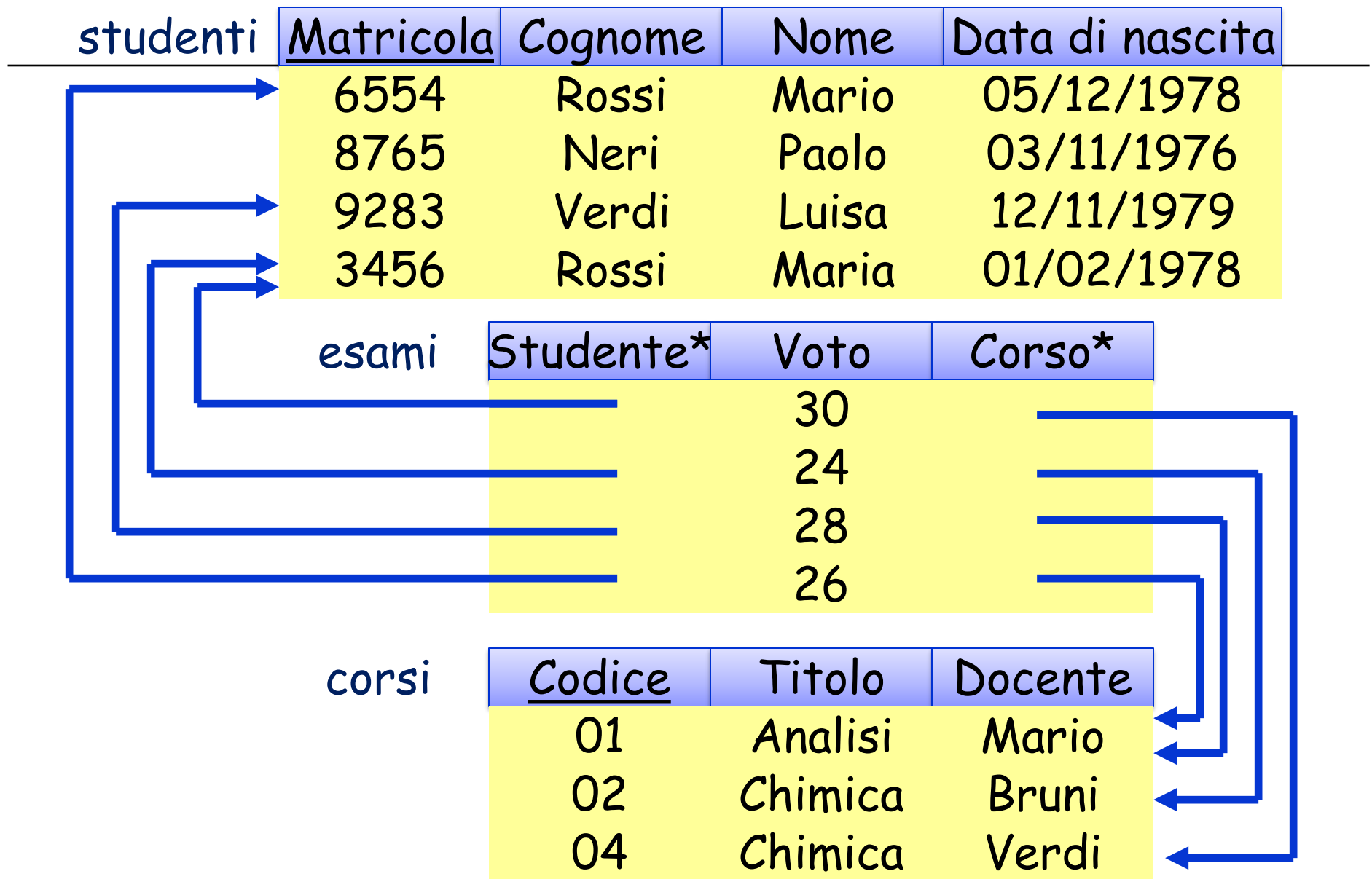
Chiave esterna

Vincolo di integrità referenziale



Esami

<u>Materia</u>	<u>Candidato*</u>	Data	Voto
BD	071523	12/01/06	28
BD	067459	15/09/06	30
FP	079856	25/10/06	30
BD	075649	27/06/06	25
LMM	071523	10/10/06	18



Infrazioni

<u>Codice</u>	Data	Vigile*	Stato	Numero
34321	1/2/17	3987	Fra	CB 123 AA
53524	4/3/18	3295	Ita	AB 222 CF
64521	5/4/17	3295	Ita	AX 424 DA
73321	5/2/19	9345	Fra	AX 424 DA

Vigili

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

Infrazioni

<u>Codice</u>	Data	Vigile	Stato*	Numero*
34321	1/2/17	3987	Fra	AX 424 DA
53524	4/3/18	3295	Ita	AB 222 CF
64521	5/4/17	3295	Ita	AX 424 DA
73321	5/2/19	9345	Fra	AX 424 DA

Auto

<u>Stato</u>	<u>Numero</u>	Cognome	Nome
Ita	AB 222 CF	Rossi	Mario
Ita	AX 424 DA	Rossi	Mario
Fra	AX 424 DA	Jacques	Dupont

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 chiave primaria di R_2

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

Violazione del vincolo di integrità referenziale

Infrazioni

<u>Codice</u>	Data	Vigile	Stato*	Numero*
34321	1/2/17	3987	Fra	CB 123 AA
53524	4/3/18	3295	Fra	AB 222 CF
64521	5/4/17	3295	Ita	AX 424 DA
73321	5/2/19	9345	Fra	AX 424 DA

Auto

<u>Stato</u>	<u>Numero</u>	Cognome	Nome
Ita	AB 222 CF	Rossi	Mario
Ita	AX 424 DA	Rossi	Mario
Fra	AX 424 DA	Jacques	Dupont

Azioni compensative

- Esempio:
 - Viene eliminata una ennupla dalla tabella riferita causando così una violazione
- Azioni
 - Rifiuto dell'operazione
 - Eliminazione in cascata
 - Introduzione di valori nulli

Eliminazione in cascata

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

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

Vincoli multipli su più attributi

Incidenti	<u>Codice</u>	Data	StatoA*	NumeroA*	StatoB*	NumeroB*
	34321	1/2/95	Ita	AB 222 CF	Ita	AX 424 DA
	64521	5/4/96	Ita	AX 424 DA	Fra	CB 123 AA

Auto

<u>Stato</u>	<u>Numero</u>	Cognome	Nome
Ita	AX 424 DA	Bianchi	Giovanni
Ita	AB 222 CF	Rossi	Mario
Fra	CB 123 AA	Jacques	Dupont

- vincoli di integrità referenziale fra:
 - La coppia di attributi StatoA e NumeroA di INCIDENTI e la relazione AUTO
 - La coppia di attributi StatoB e NumeroB di INCIDENTI e la relazione AUTO

Riepilogo: CHIAVI ED ASSOCIAZIONI

- Superchiave
- Chiave
- Chiave primaria

Esempio: (Matricola) e (Nome,Indirizzo) sono chiavi in:

- Studenti(Matricola: Int, Nome: String, Indirizzo: String)
- Chiave esterna
- Associazioni

Esercizio

- Rappresentare per mezzo di una o più relazioni le informazioni per la gestione delle Prenotazioni Mediche dei Pazienti di uno Studio Medico Associato

Esercizio

- Definire uno schema di basi di dati per organizzare le informazioni di una azienda che ha impiegati (ognuno con un codice fiscale, cognome, nome e data di nascita), filiali (con codice, sede e direttore (che è impiegato)). Ogni impiegato 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

Esercizio

- Si considerino le informazioni per la gestione dei prestiti di una biblioteca personale.
- Il proprietario presta i libri ai propri amici, che indica con i loro nomi o soprannomi, e i cui numeri di telefono sono contenuti in una rubrica.
- I libri sono individuati attraverso i titoli.
- Quando si presta un libro si prende nota della data presunta di restituzione.
- Definire uno schema per rappresentare queste informazioni indicandone i vincoli.

Esercizio

- Rappresentare con una o più relazioni le informazioni contenute nell'orario dei treni in partenza di una stazione ferroviaria: numero, orario, destinazione finale, categoria, fermate intermedie, di tutti i treni in partenza

Esercizio

- Definire uno schema di base di dati che organizzi i dati necessari a generare la pagina dei programmi radiofonici di un quotidiano, con stazioni, ora e titoli dei programmi.
- Per ogni stazione sono memorizzati, oltre al nome anche la frequenza di trasmissione e la sede.