

Il modello relazionale

Codd, 1970

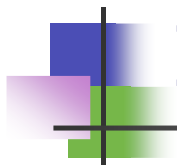
- Adottato dalla maggior parte dei DBMS in commercio
- Definisce come sono organizzati i dati e non come sono poi memorizzati e gestiti dal sistema informatico

Relazione - Tabella

Il concetto di **relazione** proviene dalla matematica mentre quello di **tabella** è intuitivo

Il modello relazionale permette di trattare i dati ad un **livello logico** senza preoccuparsi del **livello fisico** ovvero di come i dati sono effettivamente elaborati e memorizzati. Per accedere ai dati non è necessario conoscere le strutture fisiche con cui sono realizzati!

1



Implicazioni

- **Struttura**

I dati sono visti come tabelle

- **Integrità**

Le tabelle soddisfano dei vincoli di integrità

- **Manipolazione**

Gli operatori a disposizione per manipolare la tabelle (relazioni) sono operatori che derivano tabelle (relazioni) da tabelle (relazioni)

2



Prodotto cartesiano

- Consideriamo l'insieme dei Nomi e dei numeri di telefono dei dipendenti

Mario Rossi
Luca Verdi
Anna Bianchi

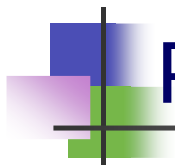
D_1

2345
2367
2378
2356

D_2

⌘ Il **prodotto cartesiano** $D_1 \times D_2$ è l'insieme di tutte le coppie ordinate (NOME, TELEFONO)

3



Prodotto cartesiano 2

Mario Rossi	2345
Mario Rossi	2367
Mario Rossi	2378
Mario Rossi	2356
Luca Verdi	2345
Luca Verdi	2367
Luca Verdi	2378
Luca Verdi	2356
Anna Bianchi	2345
Anna Bianchi	2367
Anna Bianchi	2378
Anna Bianchi	2356

- Il prodotto cartesiano contiene tutte le possibili associazioni fra gli elementi degli insiemi
- La rubrica dei numeri telefonici contiene solo alcune di tutte le possibili coppie

4



Relazioni

- Una **relazione matematica** sugli insiemi D_1 e D_2 (**domini**) è un sottoinsieme del prodotto cartesiano $D_1 \times D_2$
- Le relazioni si possono visualizzare efficacemente con una tabella in cui ogni colonna corrisponde ad un dominio e ogni riga a un elemento della relazione

La relazione RUBRICA

Mario Rossi	2345
Luca Verdi	2367
Luca Verdi	2378
Anna Bianchi	2356

La rubrica contiene solo le coppie (NOME,TELEFONO) che esistono

5



Definizione generale

- Si possono considerare n domini non necessariamente distinti $D_1, D_2, \dots, D_n$
- Il **prodotto cartesiano** $D_1 \times D_2 \times \dots \times D_n$ contiene tutte le possibile n -uple $(v_1, v_2, \dots, v_n)$ dove v_i è uno dei possibili valori presenti nel dominio D_i
- Una **relazione** R sui domini $D_1, D_2, \dots, D_n$ è un sottoinsieme del prodotto cartesiano $R \subseteq D_1 \times D_2 \times \dots \times D_n$
- Il numero delle componenti del prodotto (n) è detto **grado** della relazione; il numero di n -uple della relazione è la **cardinalità** della relazione.

Un dominio può avere dimensione infinita (es. $D_i = \mathbb{N}$). In tal caso il prodotto cartesiano è infinito, mentre possono essere definite relazioni finite.

6



Un esempio

D_1 = Squadre di Serie A

D_2 = Squadre di Serie A

D_3 = Numero

D_4 = Numero

Relazione

Risultati delle partite della II giornata del campionato

Lazio	Inter	1	1
Fiorentina	Roma	3	1
Milan	Bari	0	0
Bologna	Parma	0	1
Udinese	Napoli	2	1
Lecce	Perugia	1	1
Reggina	Juventus	3	0
Atalanta	Brescia	2	2
Verona	Vicenza	0	0

7



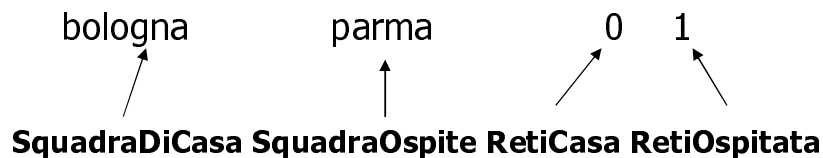
Proprietà delle relazioni

- una relazione matematica è un **insieme** di n -uple **ordinate** $(d_1, \dots, d_n)$ tali che $d_1 \in D_1, \dots, d_n \in D_n$
- 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
- ciascuno dei domini ha un **ruolo** diverso, dipendente dalla posizione (struttura **posizionale**)

8

Relazioni con attributi

- Ogni n -upla della relazione stabilisce un legame fra i dati (definisce l'esistenza di un fatto)
- Il significato dei valori dipende dall'ordine con cui sono elencati nell' n -upla



- Si può associare un nome alle componenti della n -upla
- I nomi associati ai domini si dicono **attributi** e descrivono il ruolo che il dominio rappresenta nella relazione

9

Record e campi

- Ciascuna n -upla della relazione può essere vista come un **record**, ovvero una relazione è sostanzialmente un insieme di record omogenei
- Gli attributi definiscono la struttura del record ovvero i suoi **campi**
- Si può accedere al valore di un attributo di un record usando il nome del campo

The diagram shows a table with four columns: **SquadraDiCasa**, **SquadraOspitata**, **RetiCasa**, and **RetiOspitata**. The rows are: Fiorentina, Milan, and Bologna. An arrow labeled 'tuple' points to the row containing 'Bologna'.

SquadraDiCasa	SquadraOspitata	RetiCasa	RetiOspitata
Fiorentina	Roma	3	1
Milan	Bari	0	0
Bologna	Parma	0	1

Si usa il nome **tupla** per indicare una riga di n valori corrispondenti ad attributi

10



Attributi e tuple

- Se **X** è un insieme di attributi e **D** un insieme di domini, si stabilisce una corrispondenza fra attributi e domini con una funzione

$$\text{DOM}: X \rightarrow D$$

che associa a ciascun attributo $A \in X$ un dominio $\text{DOM}(A) \in D$

- La funzione DOM corrisponde alla definizione del **tipo dati di ogni attributo**
- Una **tupla** su un insieme di attributi X è una funzione **t** che associa a ciascun attributo $A \in X$ un valore in $\text{DOM}(A)$

11



Domini e tipi di dato

- Un dominio è praticamente un tipo di dato
- La definizione di un dominio comprende
 - L'insieme di tutti i valori possibili che può assumere un elemento del dominio (e quindi i valori ammessi per l'attributo)
 - Gli operatori che possono essere applicati ai valori di quel dominio

Dominio: N (INTEGER)

Operatori: comparazione ($=, >, <$), aritmetici ($+, -, *, /$)

- Non è necessariamente legata alla rappresentazione fisica dei dati

12

Le tuple

Data una **tupla** posso estrarne il valore degli attributi

SquadraDiCasa	SquadraOspitata	RetiCasa	RetiOspitata
Fiorentina	Roma	3	1

$t[\text{SquadraOspitata}] = \text{Roma}$

$t[\text{SquadraDiCasa}] = \text{Fiorentina}$

Si possono estrarre anche insiemi di attributi ottenendo tuple

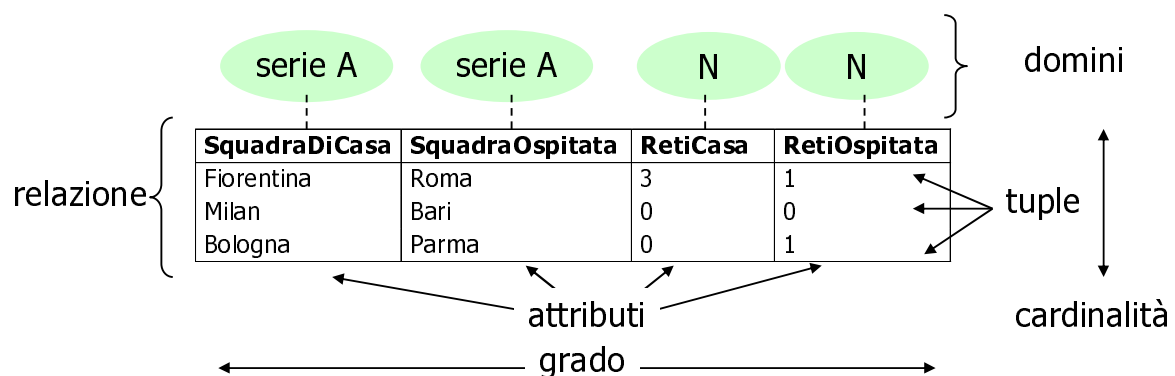
$t[\text{SquadraDiCasa}, \text{SquadraOspitata}] =$

SquadraDiCasa	SquadraOspitata
Fiorentina	Roma

13

tuple e relazioni

- Una relazione su X è un insieme di tuple su X
- Si usa la definizione di tupla al posto di n-upla
 - in una n-upla gli elementi sono individuati per posizione
 - In una tupla gli elementi sono individuati per attributi



14

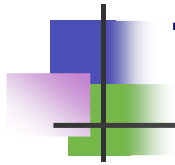


Tabelle e relazioni

- 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

15

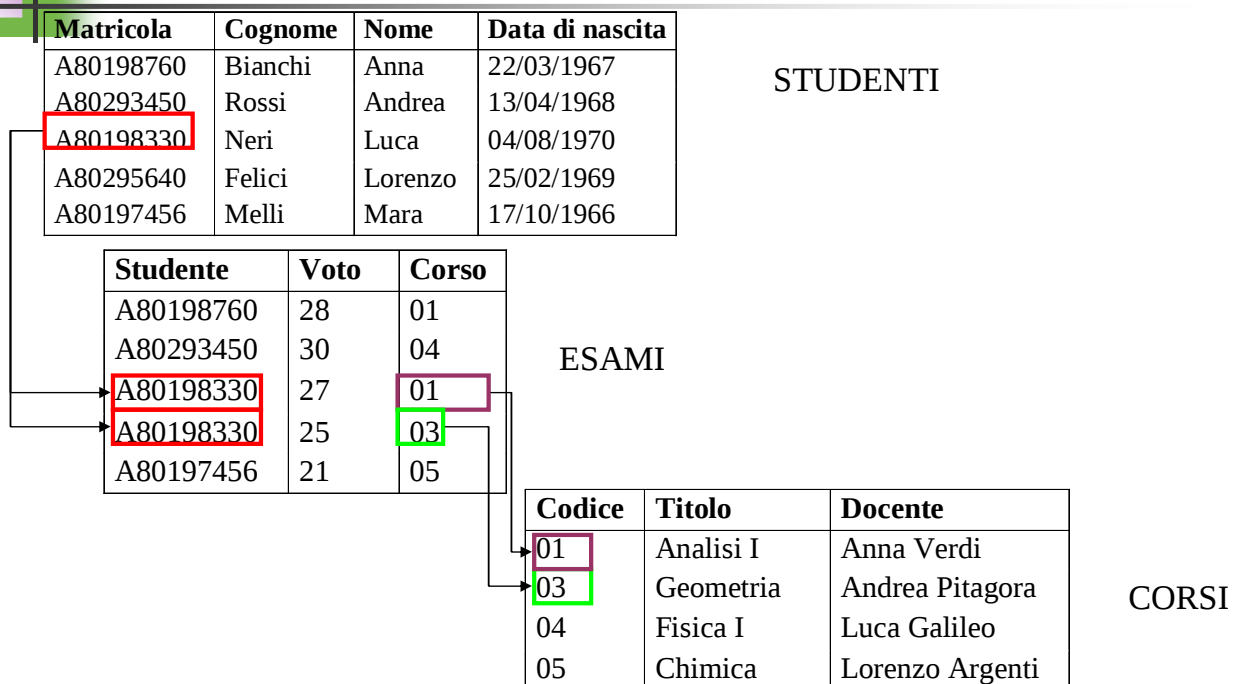


Relazioni e basi di dati

- Una base di dati è in genere costituita da più relazioni
- Si possono creare **corrispondenze** fra le tuple di relazioni distinte
- Si ottengono corrispondenze fra tuple di relazioni diverse aventi gli stessi valori su un insieme assegnato di attributi
- Il modello relazionale è **basato su valori**
 - indipendenza dalle strutture fisiche
 - si rappresenta solo ciò che è rilevante
 - 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

16

Corrispondenze fra relazioni



17

Schemi per basi di dati

- Uno **schema di relazione** è costituito dal nome della relazione e da un insieme di attributi $X = \{A_1, A_2, \dots, A_n\}$

STUDENTI(Matricola, Cognome, Nome, Data di Nascita)

- Uno **schema di base di dati** è un insieme di schemi di relazioni con nomi diversi:

$$\mathbf{R} = \{R_1(X_1), R_2(X_2), \dots, R_m(X_m)\}$$

$\mathbf{R} = \{ \text{STUDENTI(Matricola, Cognome, Nome, Data di Nascita),}$
 $\text{ESAMI(Studente, Voto, Corso),}$
 $\text{CORSI(Codice, Titolo, Docente) } \}$

18

Istanze di basi di dati

- Una **istanza di relazione** su uno schema $R(X)$ è un insieme di r tuple su X

- Un'**istanza di base di dati** su uno schema

$$\mathbf{R} = \{R_1(X_1), R_2(X_2), \dots, R_m(X_m)\}$$

è un insieme di relazioni

$$\mathbf{r} = \{r_1, r_2, \dots, r_n\},$$

dove ogni r_i , per $1 \leq i \leq n$, è una relazione dello schema $R_i(X_i)$

19

Archivio forniture

Codice	Descrizione	Magazzino	Costo	PRODOTTI
A0001	Chiodi 35mm	234000	15	
A0004	Viti 12mm	102400	25	
P0010	Pinza	200	7000	
P0230	Cacciavite	600	3500	
Articolo	Quantita	Data	Cliente	FORNITURE
A0001	3000	25/06/2000	0034	
A0004	500	25/06/2000	0034	
P0010	3	12/07/2000	0001	
P0230	2	13/07/2000	0101	
A0004	100	15/07/2000	0034	
Codice	Nome	Indirizzo	PIVA	CLIENTI
0001	Carlo Berti	Via Roma 6	02332002	
0034	BCD Spa	Via Verdi 4	04554303	
0101	A&G Srl	Viale Morgagni 16	10920393	
0076	Luca Nelli	Piazza Bixio 5	08832822	

20

Relazioni con un solo attributo

- Sono ammesse relazioni con un solo attributo
- Può essere utilizzata per selezionare un sottoinsieme delle tuple di un'altra relazione, se contiene valori che appaiono come valori di un attributo dell'altra relazione
 - Se gli elementi del sottoinsieme sono pochi rispetto al totale può essere più efficiente questa soluzione piuttosto che utilizzare un attributo booleano nella relazione



21

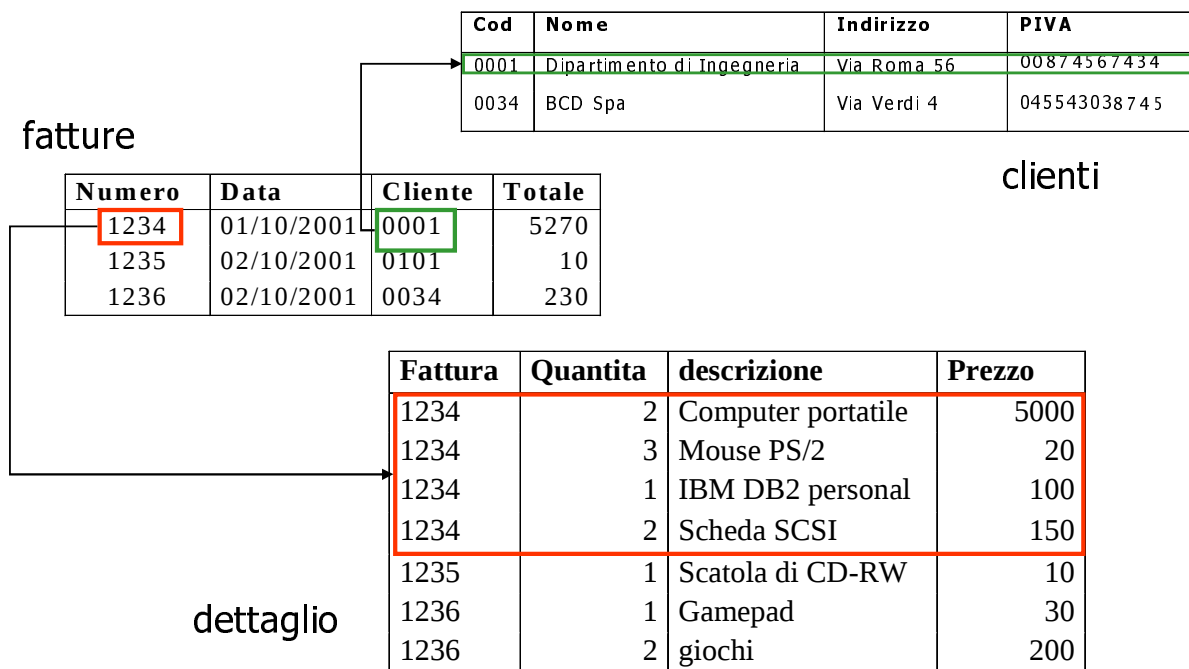
Strutture nidificate

- Rappresentare col modello relazionale un archivio di fatture emesse in seguito ad una fornitura di materiale

Fattura n. 1234 del 1/10/2001		
A: Dipartimento di Ingegneria		
PI: 00874567434		
2	Computer portatile	5000
3	Mouse PS/2	20
1	IBM DB2 personal	100
2	schede SCSI	150
Totale		5270

22

Relazioni per strutture nidificate



23

Strutture nidificate: note

- La base di dati precedente rappresenta correttamente le fatture purché:
 - Non ci siano righe ripetute in una stessa fattura (le tuple devono essere uniche)
 - Non conti l'ordine con cui devono essere riportate le righe nella fattura
- Si può arricchire la rappresentazione in modo che sia rappresentato l'ordine del dettaglio
 - Si aggiunge l'attributo **Riga** alla relazione **Dettaglio**
- Il campo **Totale** della relazione **Fatture** è un attributo ridondante (potrebbe essere calcolato?)

24

Informazione incompleta

- Può accadere che il valore di un attributo non sia disponibile per tutte le tuple
- Non è corretto usare un valore del dominio non utilizzato (es. 0, "99", stringa vuota), per rappresentare l'assenza dell'informazione
 - Non è detto che esista un valore del dominio che non sia mai usato come valore valido (es. se l'attributo è una data)
 - L'uso dei valori del dominio può generare confusione: quando si accede ai dati si deve aver chiaro quali sono i valori veri e quelli fittizi
- Si deve estendere il concetto di relazione....

25

Il valore nullo

- Si introduce un **valore nullo** (NULL) che denota l'assenza di informazione sul valore di un attributo per una data tupla
- Il valore di un attributo A può appartenere a DOM(A) o essere NULL
- Il valore di un attributo può mancare perché
 - è **sconosciuto**
 - è realmente **inesistente** (l'attributo non è applicabile ad una data tupla)
 - è **senza informazione** (non si sa se esiste o meno)

Codice	Nome	Indirizzo	PIVA
0001	Carlo Berti	Via Roma 6	NULL
0034	BCD Spa	Via Verdi 4	04554303
0101	A&G Srl	Viale Morgagni 16	NULL

inesistente

sconosciuto

26

Tuple non corrette

- Occorre che le tuple rappresentino informazioni corrette per l'applicazione

Ex: Valori nulli

Non ammesso!

Matricola	Cognome	Nome	Data di nascita
NULL	Bianchi	Anna	22/03/1967
A80293450	NULL	NULL	NULL
A80197456	Melli	Mara	NULL

STUDENTI

????

Può essere accettabile
se non è un attributo
essenziale

Alcuni campi non possono assumere valori nulli!

27

Vincoli di integrità

- Un **vincolo di integrità** è una proprietà che deve essere soddisfatta dalle istanze della base di dati che rappresentano informazioni corrette per l'applicazione
- Un vincolo può essere visto come un **predicato** che associa ad un'**istanza di base di dati** il valore **vero** o **falso**
- Se il predicato assume il valore vero, l'istanza **soddisfa** il vincolo
- Ad uno **schema di base di dati** si può associare un insieme di vincoli di integrità. Un'istanza è **ammissibile** se soddisfa tutti i vincoli definiti.

28

Tipi di vincolo

Vincolo intrarelazionale

E' definito rispetto ad una singola relazione. Può riguardare:

- le tuple nel complesso (es. non possono esistere due tuple con uno stesso valore per un dato attributo A - **vincoli di chiave**)
- le tuple indipendentemente le une dalle altre (**vincolo di tupla**)
- i valori ammessi per un attributo (**vincolo di dominio**)

Vincolo interrelazionale

Coinvolge più relazioni

- "riferimenti appesi" (**vincoli di integrità referenziale**)

29

Vincoli intrarelazionali

- Sono vincoli che coinvolgono il valore degli attributi all'interno di una stessa relazione

ESAMI

Studente	Voto	Lode	Corso
A80198760	37		01
A80293450	30	L	04
A80198330	22		01
A80198330	25	L	03

Valore non ammesso

Combinazione non ammessa

Non ci possono essere due studenti con matricola uguale!

Matricola	Cognome	Nome	Data di nascita
A80198760	Bianchi	Anna	22/03/1967
A80293450	Rossi	Andrea	13/04/1968
A80198760	Felici	Lorenzo	25/02/1969
A80197456	Melli	Mara	17/10/1966

STUDENTI

30



Vincoli interrelazionali

- Sono vincoli che coinvolgono più relazioni
- Garantiscono l'integrità dei riferimenti fra tabelle

Studente	Voto	Corso
A80198760	28	01
A80293450	30	04
A80198330	27	01
A80198330	25	03
A80197456	21	05

ESAMI

Codice	Titolo	Docente
01	Analisi I	Anna Verdi
03	Geometria	Andrea Pitagora
04	Fisica I	Luca Galileo

CORSI

?

31



Vincoli di tupla

- Esprimono condizioni sui valori degli attributi ciascuna tupla indipendentemente dalle altre
- Sono espressi con predicati che devono essere veri per tutte le tuple

$(Voto \geq 18) \text{ AND } (Voto \leq 30)$
 $\text{NOT}((Lode = 'L') \text{ AND } (Voto \neq 30))$

- Possono anche essere definite da espressioni aritmetiche che legano fra loro i valori degli attributi

$\text{PAGAMENTI}(\text{Data}, \text{Importo}, \text{Ritenute}, \text{Netto})$
 $\text{Netto} = \text{Importo} - \text{Ritenute}$

32

Chiavi

- I vincoli di chiave sono fondamentali
- Una **chiave** è un insieme di attributi utilizzato per identificare **univocamente** le tuple di una relazione

Matricola	Cognome	Nome	Data di nascita	Corso
A80198760	Rossi	Andrea	22/03/1967	Ingegneria
A80293450	Rossi	Luca	22/03/1967	Ingegneria
A80293856	Rossi	Filippo	25/02/1969	Fisica
A80198777	Bianchi	Filippo	25/02/1969	Lettere
A80197456	Bianchi	Filippo	22/10/1966	Lettere

Matricola

Nome, Cognome, Data di Nascita

SI

NO

Nome, Cognome, Corso
Cognome, Data di Nascita

33

Superchiavi e chiavi

- Un insieme **K** di attributi è **superchiave** per una relazione r se r non contiene due tuple distinte t_1 e t_2 con $t_1[K]=t_2[K]$
- Un insieme di attributi **K** è **chiave** per una relazione r se è una **superchiave minimale** (cioè non esiste un'altra superchiave K' che è contenuta propriamente in K)

CHIAVI (Superchiavi Minimali)

{Matricola}

{Nome, Cognome, Data di Nascita}

SUPERCHIAVI

{Matricola, Nome}

{Nome, Cognome, Data di Nascita, Corso}

{Matricola, Corso}

.....

34

Chiavi... per caso!

- Alcune combinazioni di attributi possono essere chiavi "per caso" per una data istanza di relazione

Matricola	Cognome	Nome	Data di nascita	Corso
A 80198760	Rossi	Andrea	22/03/1967	Ingegneria
A 80293450	Rossi	Luca	22/03/1967	Ingegneria
A 80293856	Rossi	Filippo	25/02/1969	Fisica
A 80198777	Bianchi	Filippo	25/02/1969	Lettere
A 80197456	Bianchi	Filippo	22/10/1966	Lettere

- Niente vieta che prima o poi possano iscriversi allo stesso corso due persone nate lo stesso giorno e con lo stesso nome

35

Scelta delle chiavi

- Ogni relazione $r(X)$ ha sempre una chiave
 - L'insieme di tutti gli attributi X è una superchiave dato che non possono esistere 2 tuple uguali
- La scelta della chiave deve tenere conto delle proprietà del mondo reale da cui provengono i dati
 - I vincoli sono definiti a livello di schema e devono essere validi per ogni istanza corretta nel mondo reale
- Può essere aggiunto un **campo ad hoc** (ad esempio la matricola) da usare come chiave

36

Importanza delle chiavi

- l'esistenza delle chiavi garantisce l'accessibilità a ciascun dato della base di dati
- le chiavi permettono di correlare le tuple in relazioni diverse (il modello relazionale è basato su valori e non su puntatori)
- ammettere valori nulli per gli attributi di una chiave può causare ambiguità

Matricola	Cognome	Nome	Data di nascita	Corso
A80198777	Bianchi	Filippo	25/02/1969	Lettere
NULL	Bianchi	Filippo	NULL	Lettere

??

37

Chiavi primarie

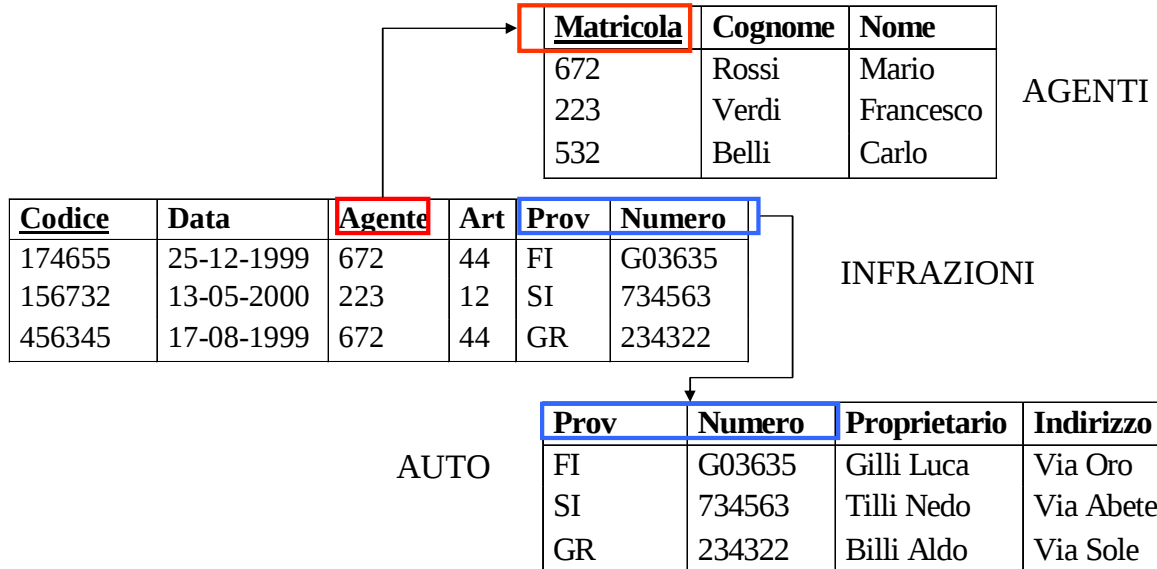
- E' una chiave su cui sono vietati i valori nulli
- Si può sempre definire un attributo che ha la proprietà di essere una chiave primaria
 - non è detto che tale attributo abbia un significato per l'applicazione
 - può essere generato in modo automatico all'atto dell'inserimento (es. codice progressivo)

<u>Matricola</u>	Cognome	Nome	Data di nascita	Corso
A80198777	Bianchi	Filippo	25/02/1969	Lettere
A80710111	Bianchi	Filippo	22/09/1971	Lettere

38

Integrità referenziale

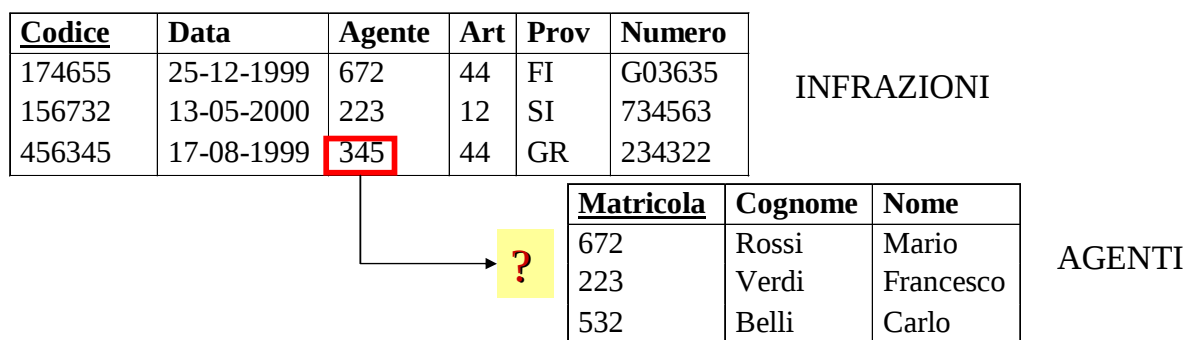
- Le corrispondenze fra i dati in relazioni diverse si stabiliscono per mezzo dei valori di chiavi delle tuple



39

Vincoli di integrità referenziale

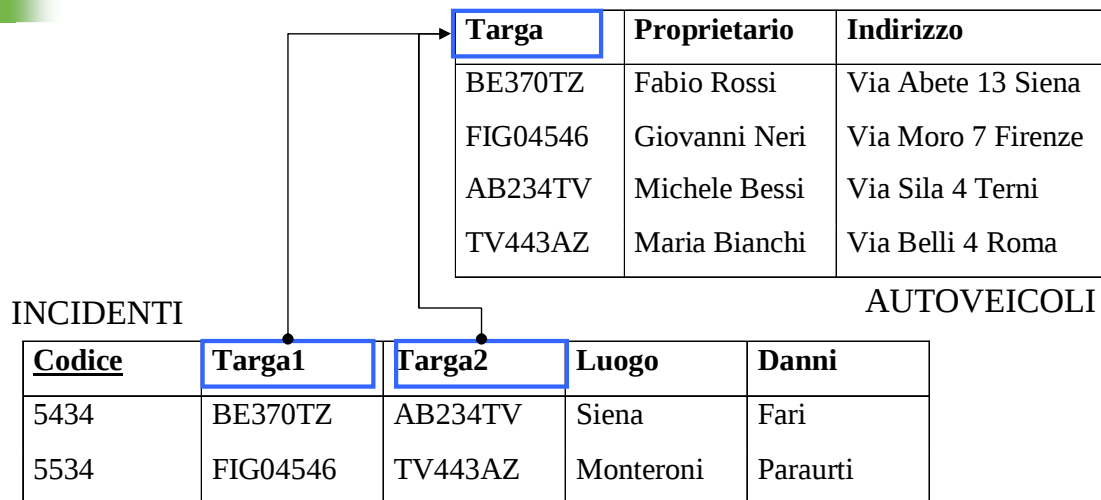
- Un **vincolo di integrità referenziale** (**foreign key**) fra un insieme di attributi X di una relazione R_1 e un'altra relazione R_2 è soddisfatto se i valori su X di ciascuna tupla in R_1 compaiono come valori della chiave (primaria) dell'istanza di R_2



Occorre definire un vincolo di integrità referenziale fra l'attributo Agente della relazione INFRAZIONI e la relazione AGENTI (chiave primaria Matricola)

40

Un esempio

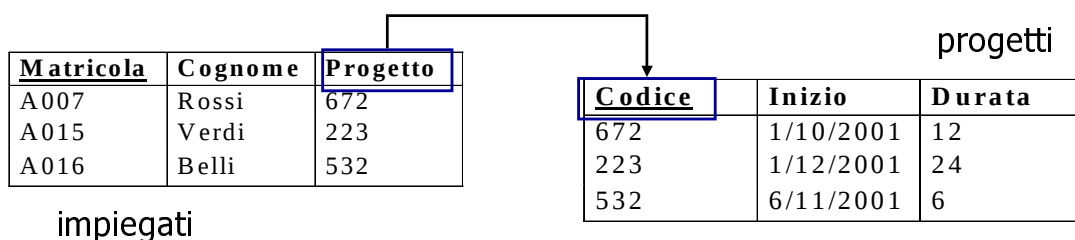


- 1) vincolo di integrità referenziale fra l'attributo Targa1 della relazione INCIDENTI e la relazione AUTOVEICOLI
- 2) vincolo di integrità referenziale fra l'attributo Targa2 della relazione INCIDENTI e la relazione AUTOVEICOLI

41

Azioni compensative

- un'operazione causa una violazione di un vincolo di integrità referenziale... cosa accade?
- Es. cancellazione di una tupla referenziata
 - Rifiuto dell'operazione
 - Eliminazione in cascata
 - Introduzione di valori nulli



42

Azioni compensative [esempio]

impiegati

<u>Matricola</u>	Cognome	Progetto
A007	Rossi	672
A015	Verdi	NULL
A016	Belli	532

progetti

<u>Codice</u>	Inizio	Durata
672	1/10/2001	12
223	1/12/2001	24
532	6/11/2001	6

sostituzione con valori nulli

impiegati

<u>Matricola</u>	Cognome	Progetto
A007	Rossi	672
A015	Verdi	223
A016	Belli	532

progetti

<u>Codice</u>	Inizio	Durata
672	1/10/2001	12
223	1/12/2001	24
532	6/11/2001	6

cancellazione in cascata