

MODELLO RELAZIONALE

AA 2006/2007

MODELLO DEI DATI

- insieme di costrutti utilizzati per organizzare i dati di interesse e descriverne la dinamica
- componente fondamentale: **meccanismi di strutturazione** (o **costruttori di tipo**)
- come nei linguaggi di programmazione esistono meccanismi che permettono di definire nuovi tipi, così ogni modello dei dati prevede alcuni costruttori

AA 2006/07

Modello relazionale

2

IL MODELLO RELAZIONALE

Permette di definire tipi per mezzo del costruttore relazione che permette di organizzare i dati in insiemi di record a struttura fissa.

Una relazione è spesso rappresentata da una tabella ove le righe rappresentano specifici record e le colonne corrispondono ai campi dei record, l'ordine di righe e colonne è sostanzialmente irrilevante.

In una base di dati relazionale ci sono in generale più relazioni.

I modelli logici dei dati

- Tre modelli logici tradizionali
 - gerarchico (organizzazione ad albero)
 - reticolare (organizzazione a grafo)
 - relazionale (organizzazione a tabella)
- Più recente (e poco diffuso)
 - a oggetti (organizzazione a oggetti)

Si chiamano modelli logici perchè pur basandosi su strutture astratte queste riflettono una particolare organizzazione

AA 2006/07

Modello relazionale

4

Modelli logici, caratteristiche

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

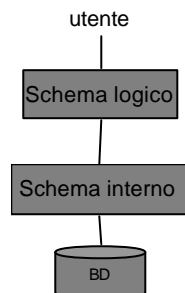
AA 2006/07

Modello relazionale

5

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)
- Basato:
 - sul concetto matematico di relazione (con una variante) –concetto formale–
 - tabelle –concetto intuitivo–
- Definisce come sono organizzati i dati e non come sono poi memorizzati e gestiti dal sistema informatico



AA 2006/07

Modello relazionale

6

Relazione: tre accezioni

- relazione matematica: come nella teoria degli insiemi
- relazione (dall'inglese relationship) che rappresenta una classe di fatti, nel modello Entity-Relationship; tradotto anche con associazione o correlazione
- relazione secondo il modello relazionale dei dati

AA 2006/07

Modello relazionale

7

Relazione matematica

- 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 ordinate (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

Il numero delle componenti del prodotto (n) è detto grado della relazione; il numero di n -uple della relazione è la cardinalità della relazione.

AA 2006/07

Modello relazionale

8

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

AA 2006/07

Modello relazionale

9

Relazione matematica, esempio

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

Mario Rossi	2345
Luca Verdi	2367
Anna Bianchi	2378
	2356
D_1	D_2

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

AA 2006/07

Modello relazionale

10

Relazione matematica, esempio

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

AA 2006/07

Modello relazionale

11

Relazione matematica, esempio

- 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

AA 2006/07

Modello relazionale

12

Relazione matematica, proprietà

- 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

AA 2006/07

Modello relazionale

13

Relazione matematica, proprietà

- ciascuna n -upla è ordinata: l' i -esimo valore proviene dall' i -esimo dominio

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

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

AA 2006/07

14

Relazione matematica, esempio

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

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

AA 2006/07

Modello relazionale

15

Relazioni nel modello relazionale

- A ciascun dominio si associa un nome (attributo), che ne descrive il "ruolo"
- Gli attributi possono essere usati come intestazione
- Struttura non posizionale

Casa	Fuori	RetiCasa	RetiFuori
Juve	Lazio	3	1
Lazio	Milan	2	0
Juve	Roma	0	2
Roma	Milan	0	1

AA 2006/07

Modello relazionale

16

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

Insegnamento	Docente	Aula	Ora
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

AA 2006/07

Modello relazionale

17

Il modello è basato su valori

- Una base di dati è in genere costituita da più relazioni
- Si possono creare corrispondenze fra le tuple di relazioni distinte per mezzo di valori degli attributi che compaiono nelle enunuple

AA 2006/07

Modello relazionale

18

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

AA 2006/07

Modello relazionale

19

Vantaggi della struttura basata su valori

- 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
- per accedere ai dati non serve sapere come sono memorizzati fisicamente

AA 2006/07

Modello relazionale

20

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
		30	
		24	
		28	
		26	

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

AA 2006/07

Modello relazionale

21

Definizioni

- Schema di relazione :
un nome R con un insieme di attributi $X=\{A_1, \dots, A_n\}$:
 $R(A_1, \dots, A_n)=R(X)$

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

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

AA 2006/07

Modello relazionale

22

Definizioni

- Schema di base di dati:
insieme di schemi di relazione con nomi diversi:

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

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

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

AA 2006/07

Modello relazionale

23

Definizioni, 2

- Una tupla su un insieme di attributi X è una funzione t che associa a ciascun attributo A in X un valore del dominio di A
- $t[A]$ (o $t.A$) denota il valore della tupla t sull'attributo A
- Per una relazione R con attributi A1, A2 e A3, e corrispondenti domini D1, D2 e D3 si può denotare anche come $R(A1 : D1, A2 : D2, A3 : D3)$

AA 2006/07

Modello relazionale

24

Definizioni, 3

- (Istanza di) relazione su uno schema $R(X)$: insieme r di tuple su X
- (Istanza di) base di dati su uno schema $R = \{R_1(X_1), \dots, R_n(X_n)\}$: insieme di relazioni $r = \{r_1, \dots, r_n\}$ (con r_i relazione su R_i)

Schema della relazione studenti

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

Istanza della relazione studenti

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
tupla $i \rightarrow$	3456	Rossi	Maria	01/02/1978

Da una tupla possiamo estrarre

il valore degli attributi t_i [matricola]=3456

Schema della relazione partite

PARTITE (Casa, Fuori, RetiCasa, RetiFuori)

Istanza della relazione partite

Partita	Casa	Fuori	RetiCasa	RetiFuori
t1	Juve	Lazio	3	1
t2	Lazio	Milan	2	0
t3	Juve	Roma	0	2
t4	Roma	Milan	0	1

t1[casa]= Juve

t2[Fuori]= Milan

t4[casa]= Roma

t4[Fuori]= Milan

t1[RetiCasa]= 3

t3[RetiFuori]= 2

Notazione

attributi: lettere iniziali dell'alfabeto, maiuscole:

A, B, C, A', A1, ...

insiemi di attributi: lettere finali dell'alfabeto, maiuscole:

X, Y, Z, X', X1, ...

giustapposizione dei nomi degli attributi: X=ABC (anziché X={A,B,C})

unioni di insiemi: XY anziché $X \cup Y$

nomi di relazione: R e lettere circostanti, maiuscole, anche con indici e pedici: R₁, S, S', ...

relazione: come il nome della relazione, ma in minuscolo

schema di base di dati: lettera maiuscola in grassetto **R, S**, ...

base di dati: stesso simbolo dello schema, ma in minuscolo

Riassunto delle definizioni

Termini informali		Termini formali
tabella		relazione
colonna		Attributo/dominio
riga		Tupla
Valori nella colonna		dominio
Definizione della tabella		Schema della relazione
Tabella popolata		Istanza della relazione

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

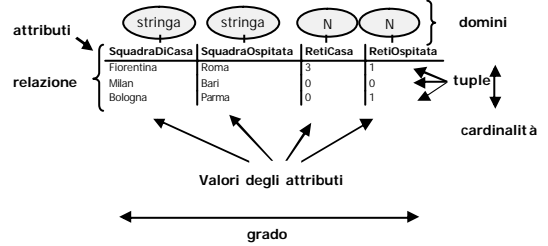
Possono esistere relazioni su solo attributo

Esempio di istanza della relazione Studenti

sid	nome	login	età	media
53666	Jones	jones@cs	18	3.4
53688	Smith	smith@eecs	18	3.2
53650	Smith	smith@math	19	3.8

- E' formalmente corretta ?
- E' stata progettata correttamente ?

Riassunto...



Strutture nidificate

Il modello relazionale non permette di usare domini arbitrari per la definizione delle relazioni; in particolare non è in generale possibile usare domini strutturati (array, set, liste, ...)

Concisaemente, una relazione in cui ogni dominio è "atomico" (non ulteriormente decomponibile) si dice che è in Prima Forma Normale, o 1NF (1st Normal Form)

In molti casi è pertanto richiesta un'attività di normalizzazione dei dati che dia luogo a relazioni in 1NF e che preservi l'informazione originale

Rappresentazione di Strutture nidificate per mezzo di relazioni

Da Filippo Via Roma 2, Roma			Da Filippo Via Roma 2, Roma		
Ricevuta Fiscale 1235 del 12/10/2000			Ricevuta Fiscale 1240 del 13/10/2000		
3	Coperti	3,00	2	Coperti	2,00
2	Antipasti	6,20	2	Antipasti	7,00
3	Primi	12,00	2	Primi	8,00
2	Bistecche	18,00	2	Orate	20,00
			2	Caffè	2,00
Totale		39,20	Totale		39,00

Relazioni che rappresentano strutture nidificate

Ricevute	Numero	Data	Totale
	1235	12/10/2000	39,20
	1240	13/10/2000	39,00

Dettaglio	Numero	Qtà	piatto	costo
	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!
 - l'ordine delle righe e' rilevante?
 - possono esistere linee ripetute in una ricevuta?
- Sono possibili rappresentazioni diverse

Rappresentazione alternativa per strutture nidificate

Ricevute	Numero	Data	Totale
	1235	12/10/2000	39,20
	1240	13/10/2000	39,00

Dettaglio	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

AA 2006/07

Modello relazionale

37

Informazione incompleta

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

AA 2006/07

Modello relazionale

38

Informazione incompleta: motivazioni

Nome	SecondoNome	Cognome
Franklin	Delano	Roosevelt
Winston	sconosciuto	Churchill
Charles	Inesistente	De Gaulle
Josip	senza informazione	Stalin

AA 2006/07

Modello relazionale

39

Tipi di valore nullo

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

AA 2006/07

Modello relazionale

40

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

AA 2006/07

Modello relazionale

41

Informazione incompleta nel modello relazionale

- Si adotta una tecnica rudimentale ma efficace:
 - 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 debbono) imporre restrizioni sulla presenza di valori nulli

AA 2006/07

Modello relazionale

42

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	Docente	
	01	Analisi	Mario	
	02	Chimica	NULL	
	NULL	Chimica	Verdi	

AA 2006/07

Modello relazionale

43

Vincoli di integrità

- E' evidente che solo alcune configurazioni di valori nulli sono ammissibili.
- Esistono istanze di basi di dati che, pur sintatticamente corrette, non rappresentano informazioni possibili per l'applicazione di interesse.

AA 2006/07

Modello relazionale

44

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	

AA 2006/07

Modello relazionale

45

Vincoli di integrità

Non tutte le tuple rappresentano informazione corretta per un'applicazione:

- valori nulli
- valori fuori del dominio di un attributo **Es: Voto = 32 in relazione Esami**
- tuple inconsistenti (valori di più attributi non simultaneamente assegnabili) **Es: Voto = 27-lode in relazione Esami**
- tuple con valori uguali per attributi identificanti **Es: tuple con stessa matricola in relazione Studenti**
- valori inesistenti in attributi usati per corrispondenze tra relazioni **Es: Studente in relazione Esami**

AA 2006/07

Modello relazionale

46

Vincoli di integrità

- 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
 - vero: istanza corretta (*ammissibile, lecita*)
 - falso: istanza inconsistente

AA 2006/07

Modello relazionale

47

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 di integrità esprimibili direttamente.

AA 2006/07

Modello relazionale

48

Tipi di vincoli

- vincoli intrarelazionali
 - vincoli su valori (o di dominio)
 - vincoli di tupla
 - vincoli di chiave (valuta le tuple nel complesso. es. non possono esistere due tuple con uno stesso valore per un particolare attributo A chiave)
- vincoli interrelazionali (coinvolge più relazioni)
 - vincoli di integrità referenziale

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 tupla

- Esprimono condizioni sui valori di ciascuna tupla, indipendentemente dalle altre tuple.
- Caso particolare:
 - Vincoli di dominio: coinvolgono un solo attributo

Sintassi ed esempi

- Possibile sintassi: espressione booleana che confronta valori di attributo o espressioni aritmetiche su di essi
- Vincolo di dominio. Es.:
(Voto \geq 18) AND (Voto \leq 30)
- Vincolo di tupla. Es.:
(Voto = 30) OR NOT (Lode = "e lode")

Vincoli di tupla, esempio

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 tuple

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

Chiave

- insieme di attributi che identificano le tuple di una relazione

Formalmente:

- un insieme K di attributi è superchiave per r se r non contiene due tuple 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
 - 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 tuple uguali su Cognome e Corso:
 - Cognome e Corso formano una chiave
- Ma è sempre vero?
- E' una chiave per caso! a noi però interessano le chiavi corrispondenti a vincoli di integrità soddisfatti da tutte le relazioni lecite dello schema.

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")

Vincoli sullo schema

Studenti

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

- chiavi:

Matricola
Cognome, Nome, Nascita

Vincoli su una istanza dello schema

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 tuple distinte ma uguali
- Il numero degli attributi è finito
segue che
- 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

Chiavi

Studenti	Matricola	CodiceFiscale	Cognome	Nome	DataNascita
	29323	BNCGRG78F21A	Bianchi	Giorgio	21/06/1978
	35467	RSSNNA78D13A	Rossi	Anna	13/04/1978
	39654	VRDMRC79K20A	Verdi	Marco	20/09/1979
	42132	VRDMRC79K20B	Verdi	Marco	20/09/1979

Il valore di Matricola identifica univocamente uno studente

Lo stesso vale per CodiceFiscale

...e per ogni insieme di attributi che include Matricola o CodiceFiscale

{Matricola,Cognome}, {CodiceFiscale,Nome}, ...

Viceversa, possono esistere due tuple uguali su {Cognome,Nome,DataNascita}

Chiavi

Studenti	Matricola	CodiceFiscale	Cognome	Nome	DataNascita
	29323	BNCGRG78F21A	Bianchi	Giorgio	21/06/1978
	35467	RSSNNA78D13A	Rossi	Anna	13/04/1978
	39654	VRDMRC79K20A	Verdi	Marco	20/09/1979
	42132	VRDMRC79K20B	Verdi	Marco	20/09/1979

Nella relazione Studenti:

{Matricola} e {CodiceFiscale} sono due chiavi

{Matricola,Cognome} e {CodiceFiscale,Nome} sono solo superchiavi

{Cognome,Nome,DataNascita} non è superchiave

Chiavi e valori nulli

- In presenza di valori nulli, i valori della chiave non permettono più
 - di identificare univocamente le tuple
 - di realizzare facilmente i riferimenti con 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

- questo schema di relazione ha come chiavi {Matricola} e {Cognome, Nome, DataNascita}
- La presenza di valori nulli nelle chiavi deve essere limitata

Chiave primaria

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

<u>Matricola</u>	<u>Cognome</u>	<u>Nome</u>	<u>Corso</u>	<u>Nascita</u>
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

Chiave primaria

- Usata per stabilire corrispondenze tra le tabelle
- se nessun attributo della relazione può giocare il ruolo di chiave primaria se ne definisce uno aggiuntivo allo scopo

Chiave: esercizio

Tab1	Matricola	Cognome	Nome	Nascita	Corso
	4328	Rossi	Luigi	29/04/59	Informatica
	6328	Rossi	Dario	29/04/59	Informatica
	4766	Rossi	Luca	01/05/61	Fisica
	4856	Neri	Luca	01/05/61	Economia
	5536	Neri	Luca	05/03/58	Economia

{Matricola} è chiave

{Cognome, Nome, Nascita} è chiave

{Matricola, Corso} è superchiave, ma non chiave

{Nome, Corso} non è superchiave

Chiave: esercizio

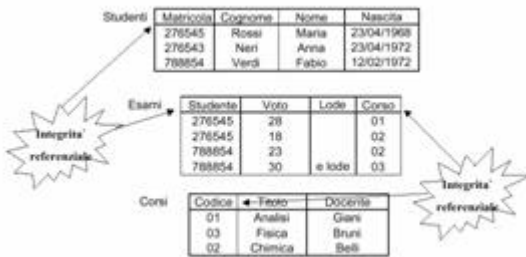
Tab2	Matricola	Cognome	Nome	Nascita	Corso
	6328	Rossi	Dario	29/04/59	Informatica
	4766	Rossi	Luca	01/05/61	Fisica
	4856	Neri	Luca	01/05/61	Economia
	5536	Neri	Luca	05/03/58	Informatica

{Nome, Corso} è "chiave per caso"

Integrità referenziale

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

Integrità referenziale



AA 2006/07

Modello relazionale

73

Infrazioni

Codice	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	Matricola	Cognome	Nome
	3987	Rossi	Luca
	3295	Neri	Piero
	9345	Neri	Mario
	7543	Mori	Gino

AA 2006/07

74

Infrazioni

Codice	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	Numero	Cognome	Nome
	MI	39548K	Rossi	Mario
	TO	E39548	Rossi	Mario
	PR	839548	Neri	Luca

AA 2006/07

Modello relazionale

75

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

AA 2006/07

Modello relazionale

76

Integrità referenziale: Graficamente

l'attributo Vigile della relazione
INFRAZIONI e la relazione VIGILI

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

R_1

Vigili (Matricola, cognome, nome)

R_2

AA 2006/07

Modello relazionale

77

Infrazioni

Codice	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	Matricola	Cognome	Nome
	3987	Rossi	Luca
	3295	Neri	Piero
	9345	Neri	Mario
	7543	Mori	Gino

AA 2006/07

78

vincoli di integrità referenziale fra: gli attributi Prov e Numero di INFRAZIONI e la relazione AUTO

Infrazioni

Codice	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	Numero	Cognome	Nome
MI	39548K	Rossi	Mario
TO	E39548	Rossi	Mario
PR	839548	Neri	Luca

AA 2006/07

Modello relazionale

79

Violazione di vincolo di integrità referenziale

Infrazioni

Codice	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	Numero	Cognome	Nome
MI	E39548	Rossi	Mario
TO	F34268	Rossi	Mario
PR	839548	Neri	Luca

AA 2006/07

Modello relazionale

80

Vincolo di integrità referenziale

Studenti	Matricola	Cognome	Nome	DataNascita
	29323	Bianchi	Giorgio	21/06/1978
	35467	Rossi	Anna	13/04/1978
	39654	Verdi	Mario	20/08/1978
	42132	Neri	Luca	15/02/1978

Corsi	CodCorso	Titolo	Docente	Anno
	483	Analisi	Biondi	1
	729	Analisi	Neri	1
	913	System Informatici	Castelli	2

Esami	Matricola	CodCorso	Voto	Lode
	29323	483	28	NO
	39654	729	30	SI
	29323	913	28	NO
	35467	913	30	NO

In Esami, {Matricola} è una foreign key, così come {CodCorso}

AA 2006/07

Modello relazionale

81

Vincolo di integrità referenziale

Corsi	Codice	Titolo	Docente	Anno
	483	Analisi	Biondi	1
	729	Analisi	Neri	1

Esami	NumMatricola	CodCorso	Voto	Lode
	29323	483	28	NO

In generale la foreign key e la chiave primaria possono includere attributi con nomi diversi

AA 2006/07

Modello relazionale

82

Vincolo di integrità referenziale

Personale	Codice	Nome	...	CodResponsabile
	123	Mario Rossi	...	325
	134	Gino Verdi	...	325
	325	Anna Neri

Foreign key e chiave primaria possono far parte della stessa relazione, ovviamente su (insiemi di) attributi diversi

AA 2006/07

Modello relazionale

83

Vincolo di integrità referenziale

Personale	Codice	Nome	...	CodResponsabile
	123	Mario Rossi	...	325
	134	Gino Verdi	...	325
	325	Anna Neri

In presenza di valori nulli, i vincoli di integrità referenziale si possono parzialmente rilassare

Studenti	Matricola	CodiceFiscale	Cognome	Nome	DataNascita
	29323	BNCGRG78F21A	Bianchi	Giorgio	21/06/1978
	35467	RSSNNA78D13A	Rossi	Anna	13/04/1978

Redditi	CF	Imponibile
	BNCGRG78F21A	10000

Nei DBMS un vincolo di integrità referenziale può anche esprimersi con riferimento ad una generica chiave (quindi anche non primaria)

AA 2006/07

Modello relazionale

84

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 ➡

AA 2006/07

Modello relazionale

85

Integrità referenziale e valori nulli

Impiegati	Matricola	Cognome	Progetto
	34321	Rossi	IDEA
	53524	Neri	XYZ
	64521	Verdi	NULL
	73032	Bianchi	IDEA

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

AA 2006/07

Modello relazionale

86

Azioni compensative

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

AA 2006/07

Modello relazionale

87

Eliminazione in cascata

Impiegati	Matricola	Cognome	Progetto
	34321	Rossi	IDEA
	53524	Neri	XYZ
	64521	Verdi	NULL
	73032	Bianchi	IDEA

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

AA 2006/07

Modello relazionale

88

Introduzione di valori nulli

Impiegati	Matricola	Cognome	Progetto
	34321	Rossi	IDEA
	53524	Neri	NULL
	64521	Verdi	NULL
	73032	Bianchi	IDEA

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

AA 2006/07

Modello relazionale

89

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

Auto	Prov	Numero	Cognome	Nome
	MI	39548K	Rossi	Mario
	TO	E39548	Rossi	Mario
	PR	839548	Neri	Luca

AA 2006/07

Modello relazionale

90

- 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

Ciascuno degli attributo in X deve corrispondere ad un preciso attributo della chiave primaria K di R_2 .

E' necessario introdurre un ordinamento.

Indicando con $X=A_1A_2...A_p$ e $K=B_1B_2...B_p$

Il vincolo di integrità referenziale è soddisfatto se, per ognitupla t_1 in R_1 senza nulli su X, esiste una tupla t_2 in R_2 tale che $t_1[A_i]=t_2[B_i]$ per ogni i compreso tra 1 e p.

Integrità referenziale: Graficamente

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

Auto (Prov, Numero, Cognome, Nome)

Riassunto modello relazionale

Il modello relazionale è basato sul concetto di relazione, che estende quello di relazione matematica tra n domini associando a ciascuna occorrenza di dominio un nome, detto attributo

Lo schema di una relazione consiste di un nome e di un insieme di attributi; l'istanza di una relazione è un insieme di tuple, ovvero funzioni che associano a ogni attributo dello schema un valore del corrispondente dominio

In assenza di informazioni si fa uso di un particolare valore, detto valore nullo (NULL), che non appartiene a nessun dominio

Per garantire l'integrità dei dati si possono specificare diversi tipi di vincoli, che definiscono quali sono le istanze legali (ammissibili)

I vincoli intra-relazionali includono quelli sui domini, sulle tuple e i vincoli di chiave; i vincoli inter-relazionali quelli di integrità referenziale. Questi ultimi permettono di stabilire le principali correlazioni tra i dati di diverse relazioni