

Basi di dati

2

IL MODELLO RELAZIONALE

Organizzazione dei dati

- Come sono organizzati i dati in una base di dati?

Modello dei dati

- Insieme di costrutti secondo cui i dati di interesse sono organizzati e utilizzati
- Esempio:
 - il **modello relazionale** prevede il costrutto **relazione** (o **tabella**), che permette di definire insiemi di record (righe) omogenei

I modelli logici dei dati

- “Tradizionali”
 - ... relazionale
- Altri più recenti:
 - a oggetti, XML, "NoSQL"

studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1998
8765	Neri	Paolo	03/11/1996
9283	Verdi	Luisa	12/11/1999
3456	Rossi	Maria	01/02/1998

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

Modello "relazionale"

- relazione matematica ... (alla lavagna)

DOMINI

$$D_1 = \{a, b\}$$

$$D_2 = \{x, y, z\}$$

PROD. CARTESIANO
 $D_1 \times D_2$

$$\left\{ \begin{array}{l} (a, x), \\ (a, y), \\ (a, z), \\ (b, x), \\ (b, y), \\ (b, z) \end{array} \right\}$$

01/10/2021

RELAZIONE BINARIA
 $R \subseteq D_1 \times D_2$

$$\left\{ \begin{array}{l} (a, x), \\ (b, x), \\ (b, z) \end{array} \right\}$$

a	x
b	x
b	z

$D_1 \quad D_2 \quad D_3$

$D_1 \times D_2 \times D_3$ tutte le tuple
(d_1, d_2, d_3)

tali che

$d_1 \in D_1, d_2 \in D_2, d_3 \in D_3$

$R \subseteq D_1 \times D_2 \times D_3$

Due osservazioni importanti

- Struttura non posizionale
- Modello basato su valori

Struttura "posizionale" o "non posizionale"

Partite* \subseteq *string* \times *string* \times *int* \times *int

Manchester	Arsenal	3	1
Arsenal	Liverpool	2	0
Manchester	Chelsea	0	2
Chelsea	Liverpool	0	1

Struttura non posizionale

- A ciascun dominio si associa un nome:
attributo (o **colonna**)

Casa	Fuori	RetiCasa	RetiFuori
Manchester	Arsenal	3	1
Arsenal	Liverpool	2	0
Manchester	Chelsea	0	2
Chelsea	Liverpool	0	1

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	05/12/1998
8765	Neri	Paolo	03/11/1996
9283	Verdi	Luisa	12/11/1999
3456	Rossi	Maria	01/02/1998

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

studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1998
8765	Neri	Paolo	03/11/1996
9283	Verdi	Luisa	12/11/1999
3456	Rossi	Maria	01/02/1998

esami

Studente	Voto	Corso
	30	
	24	
	28	
	26	

corsi

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

Basi di dati: schema e istanza

Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	10:00
Chimica	Nicola Mori	N1	10:00
Fisica I	Mario Bruni	N1	12:00
Fisica II	Mario Bruni	N3	10:00
Sistemi inform.	Piero Rossi	N3	8:00

Basi di dati: schema e istanza

Lo **schema** della relazione

Orario

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

Basi di dati: schema e istanza

Lo **schema** della relazione

Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	10:00
Chimica	Nicola Mori	N1	10:00
Fisica I	Mario Bruni	N1	12:00
Fisica II	Mario Bruni	N3	10:00
Sistemi inform.	Piero Rossi	N3	8:00

L'**istanza** della relazione

Definizioni e notazioni (alla lavagna)

- Schema di relazione
- Ennupla su uno schema di relazione
 - Valore di una ennupla su un attributo
- Relazione (“istanza di relazione“)

- Schema di base di dati
- (Istanza di) base di dati

Scheme di relazione

R nome della relazione

A_1 A_2 ... A_n attributi

D_1 D_2 D_2 domini

$R(A_1, A_2, \dots, A_n)$

STUDENTI (MATR, COGN, NOME, ...)

Schemi di base di dati

$$\mathbb{R} = \{ R_1 (A_{11}, A_{12}, \dots), \\ R_2 (A_{21}, A_{22}, \dots), \\ \dots \}$$

STUDENTI (MATR, ...)

ESAMI (STUD, CORSO, VOTO)

CORSI (CODICE, ...)

ISTANZA

(istanza di) relazione su $R(A B C)$

insieme di esemplari riga
word

con un valore per ciascuno tupla
degli attributi $A B C$

(istanza di) base di dati una relazione

01/10/2021 per ciascuna sequenza di relazioni

ennupla (o tupla) t

	A	B	C
$t \rightarrow$	a_1	b_2	c_1

$$t[A] = a_1$$

(oppure
 $t.A = a_1$)

VALORE DELLA
ENNUPLA t
SULL'ATTRIBUTO
A

- Usiamo $t[A]$ per indicare il valore della ennupla t sull'attributo A

Altra osservazione

- La struttura è rigida e non sempre corrisponde ai dati:
 - "informazione incompleta"

Informazione incompleta

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

Informazione incompleta

Nome	Ufficio	Età
Mario	Roma	34
Luca	Milano	
Piero	Firenze	42

Informazione incompleta: soluzioni?

- Potremmo pensare:
 - stringa nulla
 - "0"
 - "99"
- Si può fare?
 - Meglio di no!!

Informazione incompleta nel modello relazionale

- Tecnica semplice ma efficace:
 - **valore nullo** (un valore **diverso** dai valori del dominio, indicato qui con *NULL*)
 - $t[A]$, per ogni attributo A , è
 - un valore del dominio $\text{dom}(A)$
 - oppure il valore nullo *NULL*
- Commenti
 - tecnica semplice
 - attenzione ai troppi valori nulli

Troppi valori nulli

studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1998
9283	Verdi	Luisa	12/11/1999
<i>NULL</i>	Rossi	Maria	01/02/1998

esami

Studente	Voto	Corso
<i>NULL</i>	30	<i>NULL</i>
<i>NULL</i>	24	02
9283	28	01

corsi

Codice	Titolo	Docente
01	Analisi	Mario
02	<i>NULL</i>	<i>NULL</i>
04	Chimica	Verdi

Un'anticipazione parziale

- Riassumendo, quanto visto finora:
 - La definizione dello schema di una relazione in SQL (rivedremo più avanti)

```
postgres on paolo@PostgreSQL 11
```

```
create table esami(  
  studente integer not null,  
  corso integer not null,  
  voto integer not null  
)
```