

Basi di dati - Prof. G. Polese Capitolo 2

Modelli logici dei dati

- Tre modelli logici tradizionali
 - gerarchico
 - reticolare
 - relazionale
- Più recente (e poco diffuso)
 - a oggetti

Università degli Studi di Salerno 2

Basi di dati - Prof. G. Polese Capitolo 2

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 di valori

Università degli Studi di Salerno 3

Basi di dati - Prof. G. Polese Capitolo 2

Il modello relazionale

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

Università degli Studi di Salerno 4

Basi di dati - Prof. G. Polese Capitolo 2

Relazione: 2 accezioni

- **relazione matematica**: sottoinsieme del prodotto cartesiano di insiemi
- **relazione** secondo il modello relazionale dei dati

Università degli Studi di Salerno 5

Basi di dati - Prof. G. Polese Capitolo 2

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 (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

Università degli Studi di Salerno 6

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 7

Basi di dati - Prof. G. Polese Capitolo 2

Relazione matematica, proprietà

- Essendo la relazione matematica un insieme, esso gode delle seguenti proprietà:
 - non c'è ordinamento fra le n-uple (d_1, \dots, d_n)
 - le n-uple sono distinte
 - ciascuna n-upla è ordinata: l' i-esimo valore proviene dall' i-esimo dominio

Università degli Studi di Salerno 8

Basi di dati - Prof. G. Polese Capitolo 2

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 (string ed int) ha due ruoli diversi, distinguibili attraverso la posizione:
 - La struttura è **posizionale**

Università degli Studi di Salerno 9

Basi di dati - Prof. G. Polese Capitolo 2

Struttura non posizionale

- A ciascun dominio si associa un nome (**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

Università degli Studi di Salerno 10

Basi di dati - Prof. G. Polese Capitolo 2

Tabelle e Relazioni

- Una tabella rappresenta una relazione se
 - i valori di ogni colonna sono fra loro omogenei (dello stesso tipo)
 - le righe sono diverse fra loro (altrimenti non è un insieme)
 - 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

Università degli Studi di Salerno 11

Basi di dati - Prof. G. Polese Capitolo 2

Il modello relazionale è basato su valori

- i riferimenti fra dati in relazioni diverse (es. i dati di uno studente e quelli dei suoi esami) sono rappresentati per mezzo di valori dei domini che compaiono nelle ennuple

Università degli Studi di Salerno 12

Basi di dati - Prof. G. Polese Capitolo 2

Riferimenti tra ennuple, esempio

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

Università degli Studi di Salerno 13

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 14

Basi di dati - Prof. G. Polese Capitolo 2

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, i valori no

Università degli Studi di Salerno 15

Basi di dati - Prof. G. Polese Capitolo 2

Schema di Relazione e di BD

- Nel modello relazionale i concetti di Schema (intensione) ed Istanza (estensione) di Base di Dati si traducono in
 - Schema di relazione: un nome R con un insieme di attributi A_1, \dots, A_n : $R(A_1, \dots, A_n)$
 - 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

Università degli Studi di Salerno 16

Basi di dati - Prof. G. Polese Capitolo 2

Ennuple (Tuple) di Relazioni

- 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
- Le **ennuple** sono anche dette **tuple**

Università degli Studi di Salerno 17

Basi di dati - Prof. G. Polese Capitolo 2

Istanze di Relazione e di BD

- (Istanza di) **relazione** su uno schema $R(X)$: insieme r di ennuple 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 $R_i(X_i)$)

Università degli Studi di Salerno 18

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 19

Basi di dati - Prof. G. Polese Capitolo 2

Informazione incompleta

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

Università degli Studi di Salerno 20

Basi di dati - Prof. G. Polese Capitolo 2

Informazione incompleta: esempio

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

Università degli Studi di Salerno

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 22

Basi di dati - Prof. G. Polese Capitolo 2

Informazione incompleta nel modello relazionale

- 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 , diventa un valore del dominio $\text{dom}(A)$ oppure il valore nullo NULL
- Si possono (e si debbono) imporre restrizioni sulla presenza di valori nulli

Università degli Studi di Salerno 23

Basi di dati - Prof. G. Polese Capitolo 2

Eccesso di 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	NULL	NULL
04	Chimica	Verdi

Università degli Studi di Salerno

Basi di dati - Prof. G. Polese Capitolo 2

Tipi di valore nullo

- Esistono varie interpretazioni di un valore nullo, ad esempio
 - valore **sconosciuto**
 - valore **inesistente**
- I DBMS non distinguono i diversi tipi di valore nullo

Università degli Studi di Salerno 25

Basi di dati - Prof. G. Polese Capitolo 2

Vincoli di integrità

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

Università degli Studi di Salerno 26

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno

Basi di dati - Prof. G. Polese Capitolo 2

Vincolo 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**

Università degli Studi di Salerno 28

Basi di dati - Prof. G. Polese Capitolo 2

Vincoli di integrità, perché?

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

Università degli Studi di Salerno 29

Basi di dati - Prof. G. Polese Capitolo 2

Vincoli di integrità, nota

- non tutte le proprietà di interesse sono rappresentabili per mezzo di vincoli formulabili in modo esplicito

Università degli Studi di Salerno 30

Basi di dati - Prof. G. Polese Capitolo 2

Tipi di vincoli

- vincoli **intrarelazionali**
 - vincoli su valori (o di **dominio**)
 - vincoli di **enunzia**
- vincoli **interrelazionali**

Università degli Studi di Salerno 31

Bas di dati - Prof. G. Polese Capitolo 2

Violazione di vincoli, esempi

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

Università degli Studi di Salerno

Bas di dati - Prof. G. Polese Capitolo 2

Vincoli di ennupla

- Esprimono condizioni sui valori di ciascuna ennupla, in modo indipendente dalle altre ennuple
- Caso particolare:
 - Vincoli di dominio che coinvolgono un solo attributo

Università degli Studi di Salerno 33

Bas di dati - Prof. G. Polese Capitolo 2

Sintassi ed esempi

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

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

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

Università degli Studi di Salerno 34

Bas di dati - Prof. G. Polese Capitolo 2

Vincoli di ennupla, 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

Lordo = (Ritenute + Netto)

Università degli Studi di Salerno

Bas di dati - Prof. G. Polese Capitolo 2

Identificazione di 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

Università degli Studi di Salerno 36

Bas di dati - Prof. G. Polese Capitolo 2

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_1 e t_2 con $t_1[K] = t_2[K]$
- K è **chiave candidata** per r se è una superchiave minimale per r (cioè non contiene un'altra superchiave)

Università degli Studi di Salerno 37

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 38

Basi di dati - Prof. G. Polese Capitolo 2

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
 - è minimale

Università degli Studi di Salerno 39

Basi di dati - Prof. G. Polese Capitolo 2

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?

Università degli Studi di Salerno 40

Basi di dati - Prof. G. Polese Capitolo 2

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) solo le istanze che soddisfano tutti i vincoli

Università degli Studi di Salerno 41

Basi di dati - Prof. G. Polese Capitolo 2

Studenti

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

- chiavi candidate:
 - Matricola
 - Cognome, Nome, Nascita

Università degli Studi di Salerno 42

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 44

Basi di dati - Prof. G. Polese Capitolo 2

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 (modello basato su valori)

Università degli Studi di Salerno 45

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 46

Basi di dati - Prof. G. Polese Capitolo 2

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

- La presenza di valori nulli nelle chiavi non deve essere permessa

Università degli Studi di Salerno 47

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 48

Basi di dati - Prof. G. Polese Capitolo 2

Integrità referenziale

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

Università degli Studi di Salerno 49

Basi di dati - Prof. G. Polese Capitolo 2

Infrazioni

<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

Università degli Studi di Salerno 50

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 51

Basi di dati - Prof. G. Polese Capitolo 2

Vincolo di integrità referenziale

- Un vincolo di **integrità referenziale** ("Chiave esterna") 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

Università degli Studi di Salerno 52

Basi di dati - Prof. G. Polese Capitolo 2

Chiave Esterna (foreign key)

- In altre parole una chiave esterna è un gruppo di attributi di una relazione che costituisce anche la chiave primaria di un'altra relazione dello schema di database.

Università degli Studi di Salerno 53

Basi di dati - Prof. G. Polese Capitolo 2

Vincolo di integrità referenziale, esempio

- 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

Università degli Studi di Salerno 54

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 55

Basi di dati - Prof. G. Polese Capitolo 2

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 ➡

Università degli Studi di Salerno 56

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno 57

Basi di dati - Prof. G. Polese Capitolo 2

Azioni compensative del DBMS

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

Università degli Studi di Salerno 58

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno

Basi di dati - Prof. G. Polese Capitolo 2

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

Università degli Studi di Salerno

Basi di dati - Prof. G. Polese Capitolo 2

Ordine Attributi in Vincoli Multipli

- 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

Università degli Studi di Salerno 62