

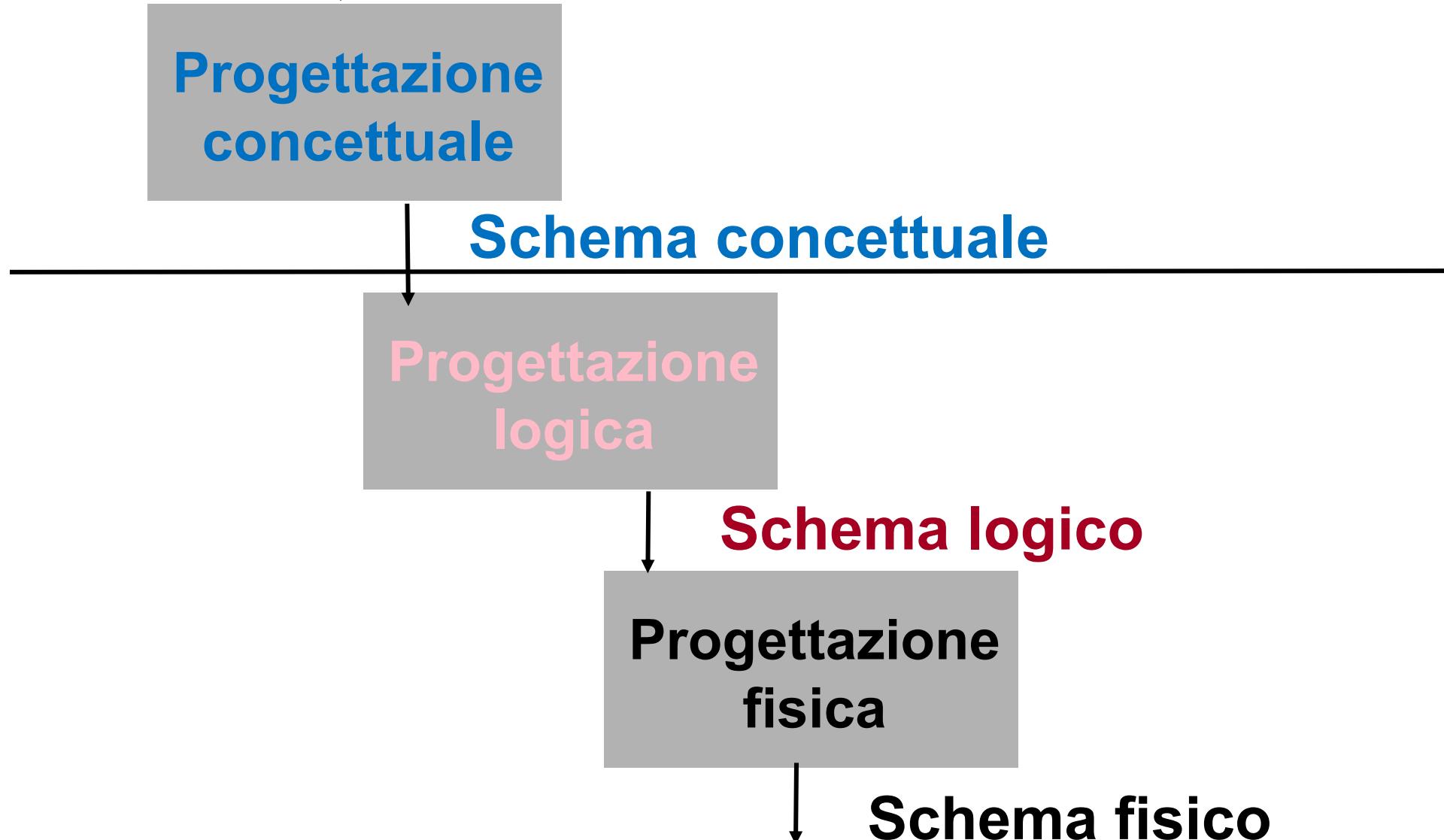


# Basi di Dati

## Modello relazionale (prima parte)

Marco Maratea  
Laurea in Informatica, DeMaCS, UNICAL

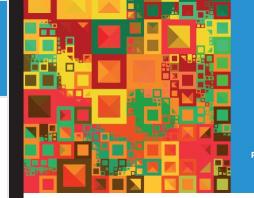
9 Ottobre 2025





# I modelli logici dei dati

- Tre modelli logici tradizionali
  - gerarchico
  - reticolare
  - relazionale
- Più recenti
  - a oggetti (poco diffuso)
  - basato su XML ("complementare" a quello relazionale)
  - NoSQL



# Modelli logici, caratteristiche

- **Gerarchico e reticolare**
  - utilizzano riferimenti esplicativi (puntatori) fra record
- **Relazionale** "è basato su valori"
  - anche i riferimenti fra dati in strutture (relazioni) diverse sono rappresentati per mezzo dei valori stessi



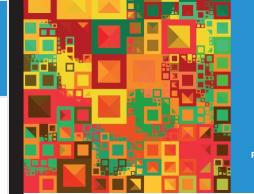
## Il modello relazionale

- 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 tavelle



# Relazione: tre accezioni

- **relazione matematica**: come nella teoria degli insiemi
- **relazione**: secondo il modello relazionale dei dati
- **relazione**: **relationship** nel modello Entità-Relazione;  
tradotto anche con **associazione** o **correlazione**



## 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



## 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



## 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:
  - 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



## Relazione matematica, esempio

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

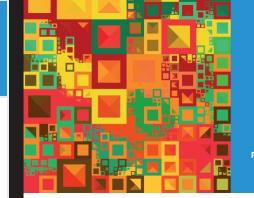
Juve	Lazio	3	1
Lazio	Milan	2	0
Juve	Roma	0	2
Roma	Milan	0	1

- Dal punto di vista della relazione, se noi invertissimo prima e seconda riga, o terza e quarta, la relazione sarebbe ancora corretta; ma cosa ne sarebbe del significato?



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

Casa	Fuori	Reti	Casa	Reti	Fuori
Juve	Lazio	3			1
Lazio	Milan	2			0
Juve	Roma	0			2
Roma	Milan	0			1



## Tabelle e relazioni

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



# 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	05/12/1978
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



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



# Alternativa

- Altri modelli (ad es. quelli "storici", reticolare e gerarchico) prevedono riferimenti esplicativi, gestiti dal sistema



studenti

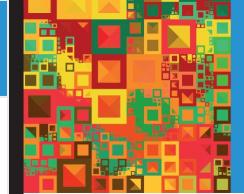
Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

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



**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



Paolo Atzeni  
Stefano Ceri  
Piero Fraternali  
Stefano Paraboschi  
Riccardo Torlone

Basi di dati

VI edizione

McGraw-Hill Connect

McGraw-Hill

# Struttura basata su valori: vantaggi

- indipendenza dalle strutture fisiche 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 programmati
- i dati sono portabili più facilmente da un sistema ad un altro



## Definizioni

- **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:  
$$R = \{R_1(X_1), \dots, R_k(X_k)\}$$
 con  $X_i$  insiemi di attributi



## Definizioni, 2

- 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**
- Es.: se  $X = (\text{Nome}, \text{Cognome}, \text{Età})$  e  $t = (\text{"Mario"}, \text{"Rossi"}, 34)$ , allora  $t(\text{Cognome}) = \text{"Rossi"}$



## Definizioni, 3

- 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 instanze di relazioni  $r = \{r_1, \dots, r_n\}$  (con  $r_i$  istanza di relazione di  $R_i$ , quindi insieme di ennuple su  $X_i$ )



# 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

studentiLav

Matricola
6554
3456



# Relazioni ed istanze

**Schema di basi di dati:** {studente(Matricola, Cognome, Nome, Data di Nascita), studenteLav(Matricola)}

**Instanza di basi di dati:**

```
{{("6554","Rossi","Mario",05/12/1978),  
 ("8765","Neri","Paolo",03/11/1976),  
 ("9284","Verdi","Luisa",12/11/1979),  
 ("3456","Rossi","Maria",01/02/1978)},{("6454"),("3456")}}
```



# Esempio

<b>DA FILIPPO</b> <b>VIA ROMA 2, ROMA</b>		
<b>RICEVUTA FISCALE</b>		
1235 <b>DEL</b> 12/10/2017		
3	Coperti	3,00
2	Antipasti	6,20
3	Primi	12,00
2	Bistecche	18,00
<b>TOTALE</b>		39,20

<b>DA FILIPPO</b> <b>VIA ROMA 2, ROMA</b>		
<b>RICEVUTA FISCALE</b>		
1240 <b>DEL</b> 13/10/2017		
2	Coperti	2,00
2	Antipasti	7,00
2	Primi	8,00
2	Orate	20,00
2	Caffè	2,00
<b>TOTALE</b>		39,00



Ricevute

## Tutto in una tabella

Numero	Data	Qtà	Descrizione	Importo	Totale
1235	12/10/2017	3	Coperti	3,00	39,20
1235	12/10/2017	2	Antipasti	6,20	39,20
...		3	Primi	12,00	
...		2	Bistecche	18,00	
1240	13/10/2017	2	Coperti	2,00	39,00
1240	13/10/2017	...	...	...	39,00

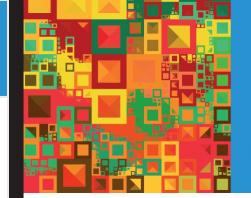
Numero, Data e Totale sono ripetute per ogni riga della ricevuta, avendo ridondanza. Organizziamo in due tabella con informazioni «fisse» e «variabili».



## Strutture nidificate

<b>DA FILIPPO</b> <b>VIA ROMA 2, ROMA</b>		
<b>RICEVUTA FISCALE</b>		
1235 <b>DEL</b> 12/10/2017		
3	Coperti	3,00
2	Antipasti	6,20
3	Primi	12,00
2	Bistecche	18,00
<b>TOTALE</b>		39,20

<b>DA FILIPPO</b> <b>VIA ROMA 2, ROMA</b>		
<b>RICEVUTA FISCALE</b>		
1240 <b>DEL</b> 13/10/2017		
2	Coperti	2,00
2	Antipasti	7,00
2	Primi	8,00
2	Orate	20,00
2	Caffè	2,00
<b>TOTALE</b>		39,00



## Strutture nidificate

<b>DA FILIPPO</b> <b>VIA ROMA 2, ROMA</b>		
<b>RICEVUTA FISCALE</b>		
1235 <b>DEL</b> 12/10/2017		
3	Coperti	3,00
2	Antipasti	6,20
3	Primi	12,00
2	Bistecche	18,00
<b>TOTALE</b>		39,20

<b>DA FILIPPO</b> <b>VIA ROMA 2, ROMA</b>		
<b>RICEVUTA FISCALE</b>		
1240 <b>DEL</b> 13/10/2017		
2	Coperti	2,00
2	Antipasti	7,00
2	Primi	8,00
2	Orate	20,00
2	Caffè	2,00
<b>TOTALE</b>		39,00



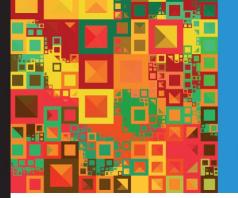
# Relazioni che rappresentano strutture nidificate

## Ricevute

	Numero	Data	Totale
	1235	12/10/2017	39,20
	1240	13/10/2017	39,00

## Dettaglio

	Numero	Qtà	Descrizione	Importo
	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?  
La rappresentazione è corretta?
- Dipende da che cosa ci interessa rappresentare...
  - possono esistere linee ripetute in una ricevuta?
  - possiamo ricostruire esattamente le ricevute partendo dalla base di dati?



# Rappresentazione alternativa per strutture nidificate

Ricevute

	Numero	Data	Totale
	1235	12/10/2017	39,20
	1240	13/10/2017	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
	...	...	...	...	...



## 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
- I dati disponibili possono non corrispondere al formato previsto



# Informazione incompleta: motivazioni

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



## Informazione incompleta: soluzioni?

- Non conviene (anche se ancora lo si fa) usare valori del dominio (0, stringa nulla, “99”, ...):
  - potrebbero non esistere valori “non utilizzati” (es. date)
  - valori “non utilizzati” potrebbero diventare significativi (es. matricola “999999”)
  - in fase di utilizzo (nei programmi) sarebbe necessario ogni volta tener conto del “significato” di questi valori



## Informazione incompleta nel modello relazionale

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



# Tipi di valore nullo

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



## Troppi valori nulli

studenti

	Matricola	Cognome	Nome	Data di nascita
	6554	Rossi	Mario	<i>NULL</i>
	9283	Verdi	Luisa	12/11/1993
	<i>NULL</i>	Rossi	Maria	01/02/1994

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



# Vincoli di integrità

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



## 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



## 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**



# 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



## Vincoli di integrità, nota

- alcuni tipi di vincoli (ma non tutti) sono "supportati" dai DBMS:
  - possiamo quindi specificare vincoli di tali tipi nella nostra base di dati e il DBMS ne impedisce la violazione
  - per i vincoli "non supportati", la responsabilità della verifica è dell'utente o del programmatore



## Tipi di vincoli

- vincoli **intrarelazionali**, con casi particolari:
  - vincoli su valori (o di **dominio**)
  - vincoli di **ennupla**
- vincoli **interrelazionali**



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 ennupla

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

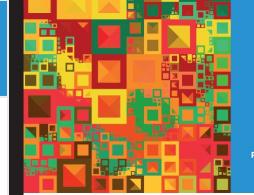


## Sintassi ed esempi

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

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

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



## Vincoli di ennupla, altro 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)

Esempio



## Vincoli di ennupla, violazione

Stipendi	Impiegato	Lordo	Ritenute	Netto
	Rossi	55.000	12.500	42.500
	Neri	45.000	10.000	35.000
	Bruni	50.000	11.000	36.000

Lordo = (Ritenute + Netto)