

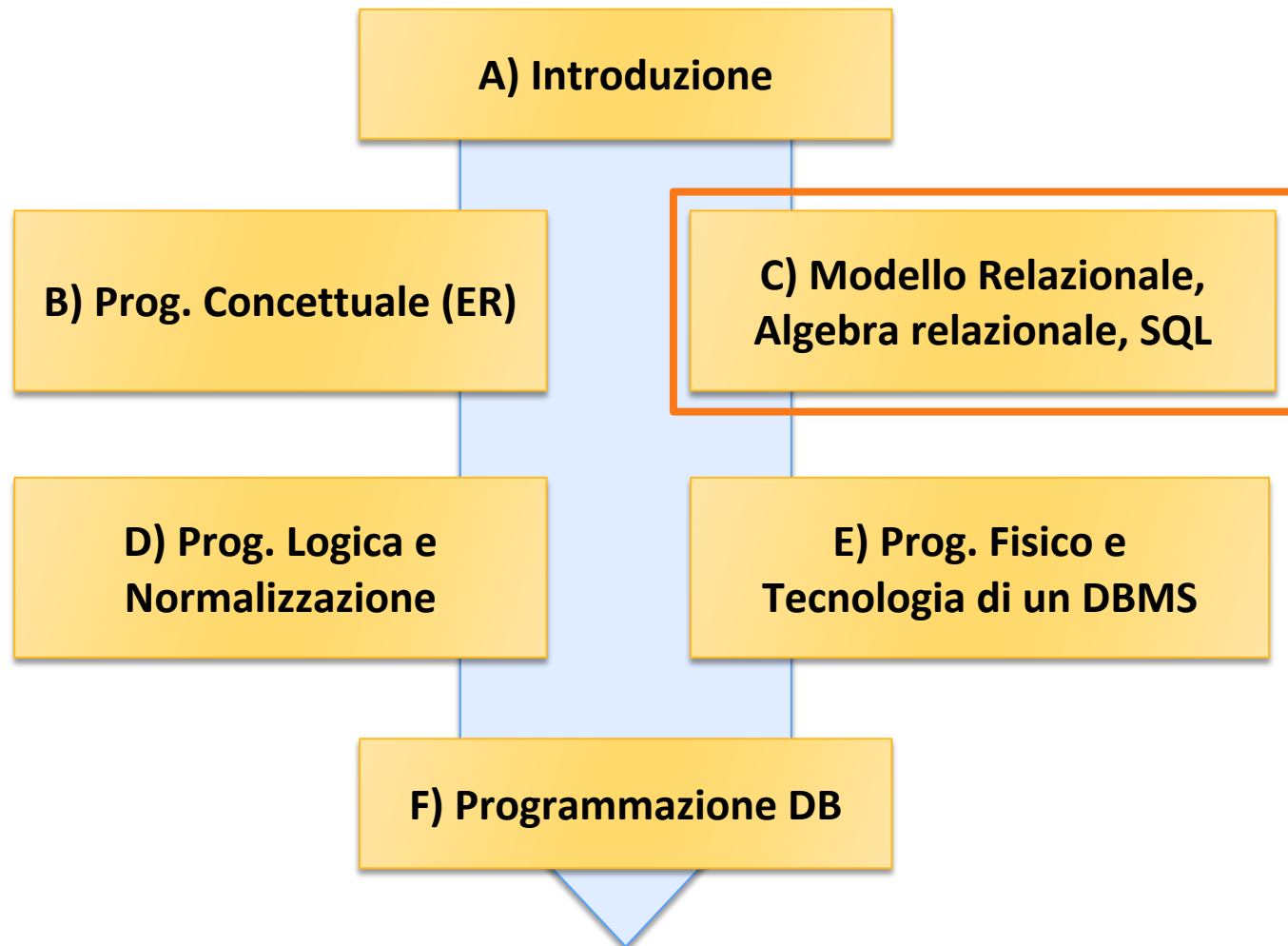


Basi di Dati



Il modello relazionale dei dati

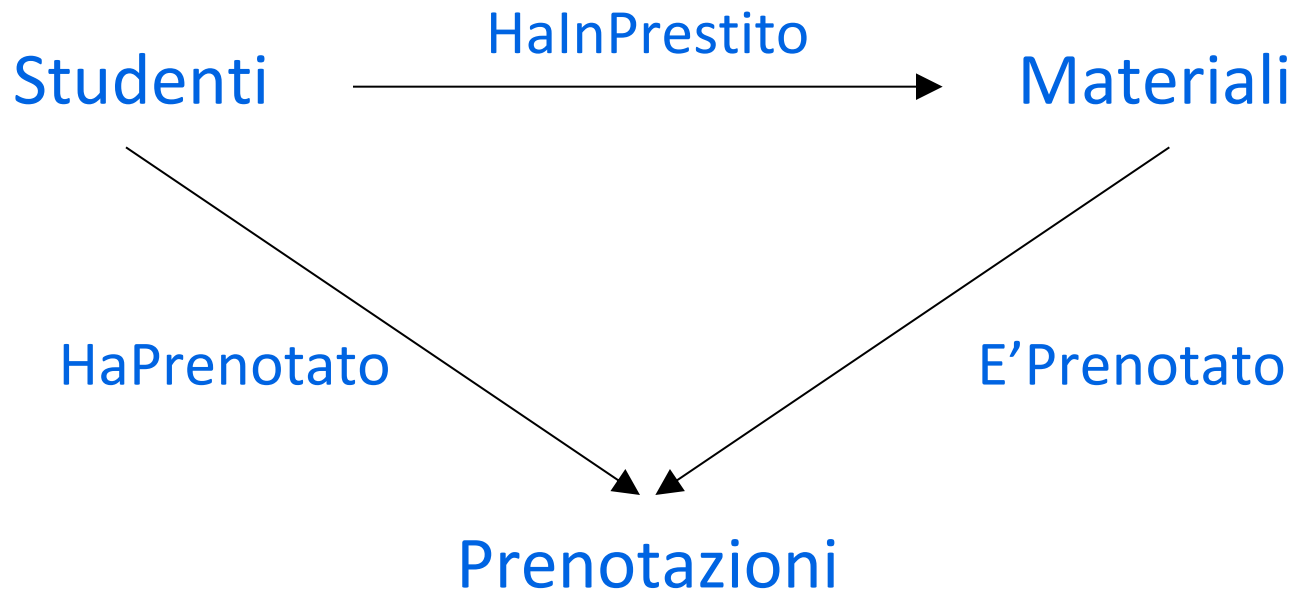
Basi di Dati – Dove ci troviamo?



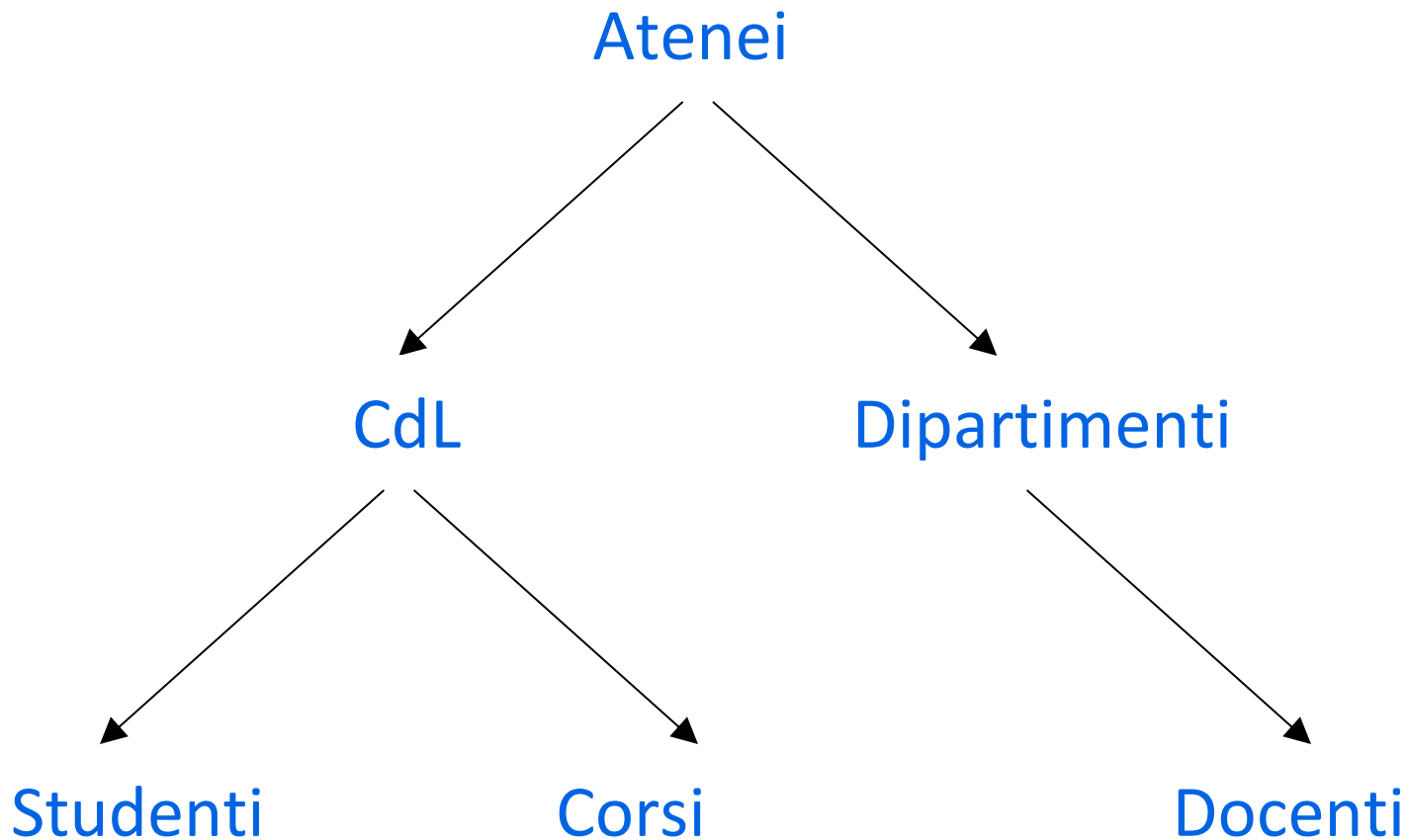
Cronologia dei modelli per la rappresentazione dei dati

- Modello gerarchico (anni 60)
- Modello reticolare (anni 70)
- Modello relazionale (anni 80)
- Modello a oggetti (anni 90)

Esempio di modello reticolare



Esempio di modello gerarchico



Cronologia del modello relazionale

- Inventato da Codd nel 1970
 - (IBM Research di Santa Teresa, Cal)
- Primi progetti:
 - SYSTEM R (IBM), Ingres (Berkeley Un.)
- Prima versione del linguaggio SQL (allora SEQUEL): 1974
- Primi sistemi commerciali: inizio anni '80 (Oracle, IBM-SQL DS e DB2, Ingres, Informix, Sybase)
- Successo commerciale: dal 1985.

Definizione informale

tabella



studente

colonna



schema



MAT	NOME	CITTA'	C-DIP
R	Carlo	Bologna	Inf
123	Giovanni	Milano	Log
307	Paola	Torino	Inf
415	Antonio	Roma	Log



istanza



riga



702

Definizione formale

- ✓ **Dominio D :**
 - ✓ un qualunque insieme di valori
- ✓ **Prodotto cartesiano** su n domini (non necessariamente distinti), $D1 \times D2 \times \dots \times Dn$:
 - ✓ insieme di tutte le n-ple (**tuple**) $\langle d1, d2, \dots, dn \rangle$, con $d_i \in Di, 1 \leq i \leq n$
- ✓ **Relazione R** su $D1, D2, \dots, Dn$: un qualunque sottoinsieme di $D1 \times D2 \times \dots \times Dn$
$$R \subseteq D1 \times D2 \times \dots \times Dn$$

Esempio

- ❖ $D1 = (a,b)$
- ❖ $D2 = (1,2,3)$
- ❖ $D1 \times D2 = (\langle a,1 \rangle, \langle b,1 \rangle, \langle a,2 \rangle, \langle b,2 \rangle, \langle a,3 \rangle, \langle b,3 \rangle)$
- ❖ $R1 = (\langle a,1 \rangle, \langle b,3 \rangle)$
- ❖ $R2 = (\langle a,2 \rangle, \langle b,1 \rangle, \langle b,3 \rangle)$
- ❖ $R3 = ()$
- ❖ $R4 = (\langle a,1 \rangle, \langle b,1 \rangle, \langle a,2 \rangle, \langle b,2 \rangle, \langle a,3 \rangle, \langle b,3 \rangle)$

Proprietà

- **Grado** della relazione:
 - numero di domini (n)
- **Cardinalità** della relazione:
 - numero di tuple
- **Attributo**:
 - nome assegnato ad un dominio in una relazione

Proprietà

- Schema di una relazione R :

$$R(attr1, \dots, attrN)$$

[I nomi degli attributi in uno schema devono essere tutti **distinti** fra loro]

- (Istanza della) relazione R :

- un insieme r di tuple su $(attr1, \dots, attrN)$

- $t[attr]$ denota il valore della tupla t sull'attributo $attr$

A	B
a	1
b	3

C	D
c	1
b	3
a	2

Confronto della terminologia

DEFINIZIONE FORMALE	DEFINIZIONE INFORMALE
relazione attributo tupla, n-pla dominio cardinalita' grado	tabella colonna riga tipo di dato numero di righe numero di colonne

Una differenza
significativa:

DEFINIZIONE
FORMALE
assenza
di duplicati

DEFINIZIONE
INFORMALE
possibili duplicati

Base di dati

❑ Schema di base di dati:

- ❑ un insieme di schemi di relazione

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

[tutti i nomi di relazioni della base di dati devono essere **differenti**]

❑ Istanza della 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)

$R1(A,B)$

A	B
a	1
b	3

$R2(C,D)$

C	D
c	1
b	3
a	2

Esempio: gestione degli esami universitari

studente

MATR	NOME	CITTA'	C-DIP
123	Carlo	Bologna	Inf
415	Paola	Torino	Inf
702	Antonio	Roma	Log

Esempio: gestione degli esami universitari

corso

COD-CORSO	TITOLO	DOCENTE
1	matematica	Barozzi
2	informatica	Meo

Esempio: gestione degli esami universitari

esame

MATR	COD-CORSO	DATA	VOTO
123	1	7-9-14	30
123	2	8-1-15	28
702	2	7-9-14	20

Esempio: gestione degli esami universitari

studente

MATR	NOME	CITTA'	C-DIP
123	Carlo	Bologna	Inf
415	Paola	Torino	Inf
702	Antonio	Roma	Log

esame

MATR	COD-CORSO	DATA	VOTO
123	1	7-9-14	30
123	2	8-1-15	28
702	2	7-9-14	20

corso

COD-CORSO	TITOLO	DOCENTE
1	matematica	Barozzi
2	informatica	Meo

Interrogazioni

quali professori hanno esaminato Carlo?

studente

MATR	NOME	CITTA'	C-DIP
123	Carlo	Bologna	Inf
415	Paola	Torino	Inf
702	Antonio	Roma	Log

esame

MATR	COD-CORSO	DATA	VOTO
123	1	7-9-14	30
123	2	8-1-15	28
702	2	7-9-14	20

corso

COD-CORSO	TITOLO	DOCENTE
1	matematica	Barozzi
2	informatica	Meo

Interrogazioni

quali studenti hanno preso 30 in matematica?

studente

MATR	NOME	CITTA'	C-DIP
123	Carlo	Bologna	Inf
415	Paola	Torino	Inf
702	Antonio	Roma	Log

esame

MATR	COD-CORSO	DATA	VOTO
123	1	7-9-14	30
123	2	8-1-15	28
702	2	7-9-14	20

corso

COD-CORSO	TITOLO	DOCENTE
1	matematica	Barozzi
2	informatica	Meo

Esempio: gestione personale

impiegato

MATR	NOME	DATA-ASS	SALARIO	MATR-MGR
1	Piero	1-1-12	1500 €	2
2	Giorgio	1-1-14	2000 €	null
3	Giovanni	1-7-13	1000 €	2

assegnamento

MATR	NUM-PROG	PERC
1	3	50
1	4	50
2	3	100
3	4	100

progetto

NUM-PROG	TITOLO	TIPO
3	Idea	Esprit
4	Wide	Esprit

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



Interrogazioni

chi e' il manager di Piero?
impiegato

MATR	NOME	DATA-ASS	SALARIO	MATR-MGR
1	Piero	1-1-12	1500 €	2
2	Giorgio	1-1-14	2000 €	null
3	Giovanni	1-7-13	1000 €	2

assegnamento

MATR	NUM-PROG	PERC
1	3	50
1	4	50
2	3	100
3	4	100

progetto

NUM-PROG	TITOLO	TIPO
3	Idea	Esprit
4	Wide	Esprit

Interrogazioni

in quali tipi di progetti lavora Giovanni?
impiegato

MATR	NOME	DATA-ASS	SALARIO	MATR-MGR
1	Piero	1-1-12	1500 €	2
2	Giorgio	1-1-14	2000 €	null
3	Giovanni	1-7-13	1000 €	2

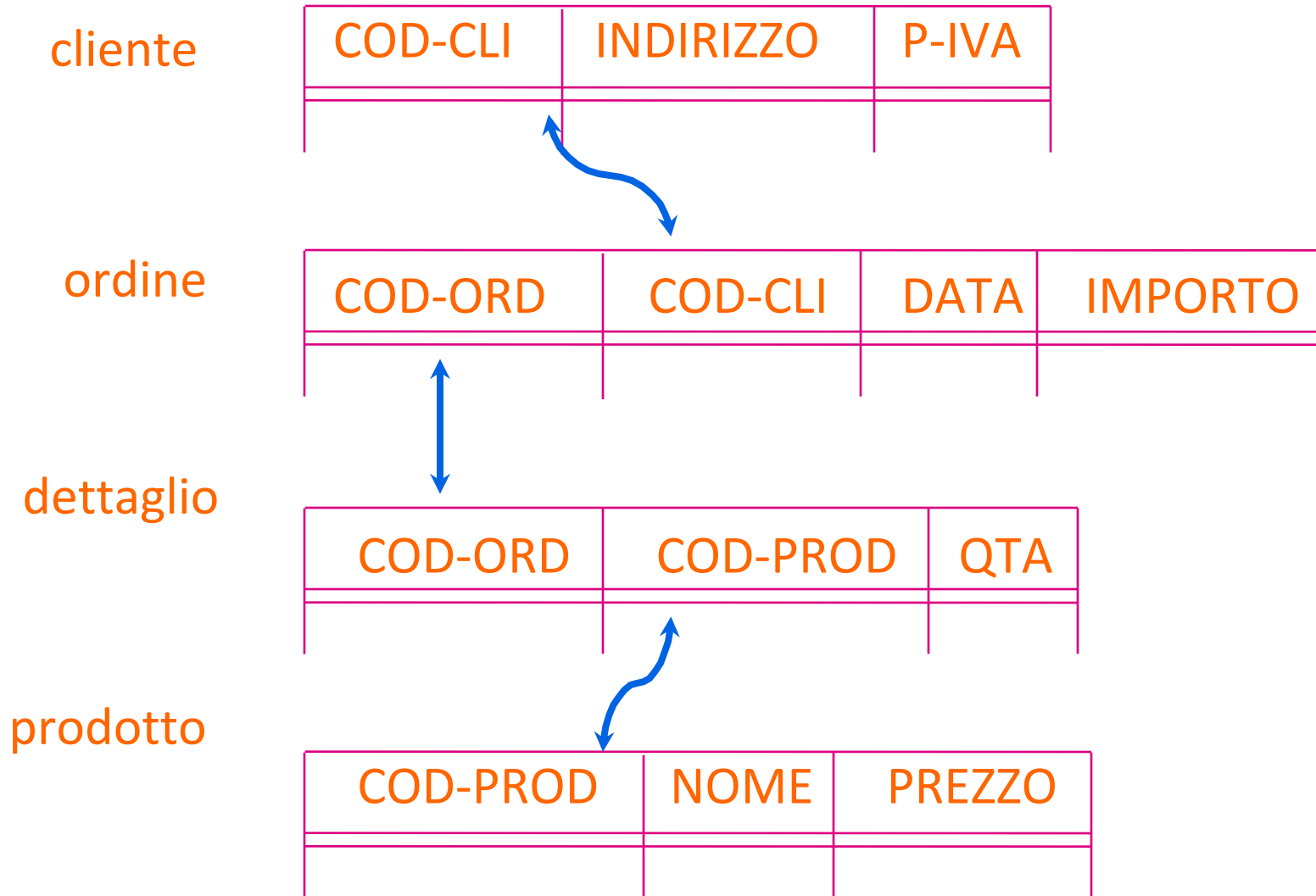
assegnamento

MATR	NUM-PROG	PERC
1	3	50
1	4	50
2	3	100
3	4	100

progetto

NUM-PROG	TITOLO	TIPO
3	Idea	Esprit
4	Wide	Esprit

Esempio: gestione ordini



Interrogazioni

- quali ordini ha emesso Paolo?
- quanti ordini ha emesso Paolo?
- quante candele sono state ordinate il 5/7/15?
- calcolare per ciascun cliente la somma degli importi di tutti gli ordini
- estrarre l'ordine di importo più alto

Riflessioni

- Differenza fra **schema** e **istanza**
- Due attività assai differenti:
 - progetto dello schema
 - gestione dell'istanza
- Passaggio dai dati all'informazione
 - **query language**

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



Come arricchire lo schema?

- VINCOLI DI INTEGRITA':
 - escludono alcune istanze in quanto, pur sintatticamente corrette, non rappresentano informazioni possibili per l'applicazione di interesse
- VINCOLI DI INTEGRITA', SCHEMI E ISTANZE
 - 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”)
- CHIAVI
- VINCOLI SUI VALORI NULLI (poi)
- INTEGRITA' REFERENZIALE (poi)
- VINCOLI GENERICI (poi)

Nozione di CHIAVE

- Sottoinsieme degli attributi dello schema che ha la proprietà di unicità e minimalità
- **unicità:**
 - non esistono due tuple con chiave uguale
- **minimalità:**
 - sottraendo un qualunque attributo alla chiave si perde la proprietà di unicità

Definizione di SUPERCHIAVE e CHIAVE

- 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** per r se è una superchiave minimale per r (cioè non contiene un'altra superchiave)

Chiavi nell'esempio: gestione degli esami universitari

studente

<u>MATR</u>	NOME	CITTA'	C-DIP

esame

<u>MATR</u>	<u>COD-CORSO</u>	DATA	VOTO

corso

<u>COD-CORSO</u>	TITOLO	DOCENTE

Chiavi nell'esempio: gestione personale

impiegato

<u>MATR</u>	NOME	DATA-ASS	SALARIO	MATR-MIL

assegnamento

<u>MATR</u>	<u>NUM-PROG</u>	PERC

progetto

<u>NUM-PROG</u>	NOME	PREZZO

Chiavi nell'esempio: gestione ordini

cliente

<u>COD-CLI</u>	INDIRIZZO	P-IVA

ordine

<u>COD-ORD</u>	COD-CLI	DATA	IMPORTO

dettaglio

<u>COD-ORD</u>	<u>COD-PROD</u>	QTA

prodotto

<u>COD-PROD</u>	NOME	PREZZO

Con molteplici chiavi:

Una è definita **CHIAVE PRIMARIA** le rimanenti chiavi sono **SECONDARIE (ALTERNATIVE)**

CLIENTE (COD-CLIENTE,INDIRIZZO,P-IVA)

AK: P-IVA

Chiave primaria: COD-CLIENTE

(indicata sottolineando i relativi attributi)

Chiave secondaria: P-IVA

(indicata a parte dalla sigla AK, Alternative Key)