

# Progettazione di basi di dati



1

# Progettazione

---

Metodologie e modelli

2

Storia di un progetto



Ciò che ha chiesto il cliente

Basi di Dati corso A

3



Ciò che ha chiesto il cliente

Storia di un progetto



Ciò che ha capito il commerciale

Basi di Dati corso A

4

Ciò che ha chiesto il cliente

**Storia di un progetto**

Come ha risolto il problema la progettazione

Basi di Dati corso A

5

Ciò che ha chiesto il cliente

**Storia di un progetto**

Ciò che ha realizzato la fabbricazione

Basi di Dati corso A

6

Storia di  
un  
progetto

Ciò che ha chiesto il cliente

Basi di Dati corso A

7

Storia di  
un  
progetto

Ciò che ha chiesto il cliente

Ciò che realmente voleva il cliente

Basi di Dati corso A

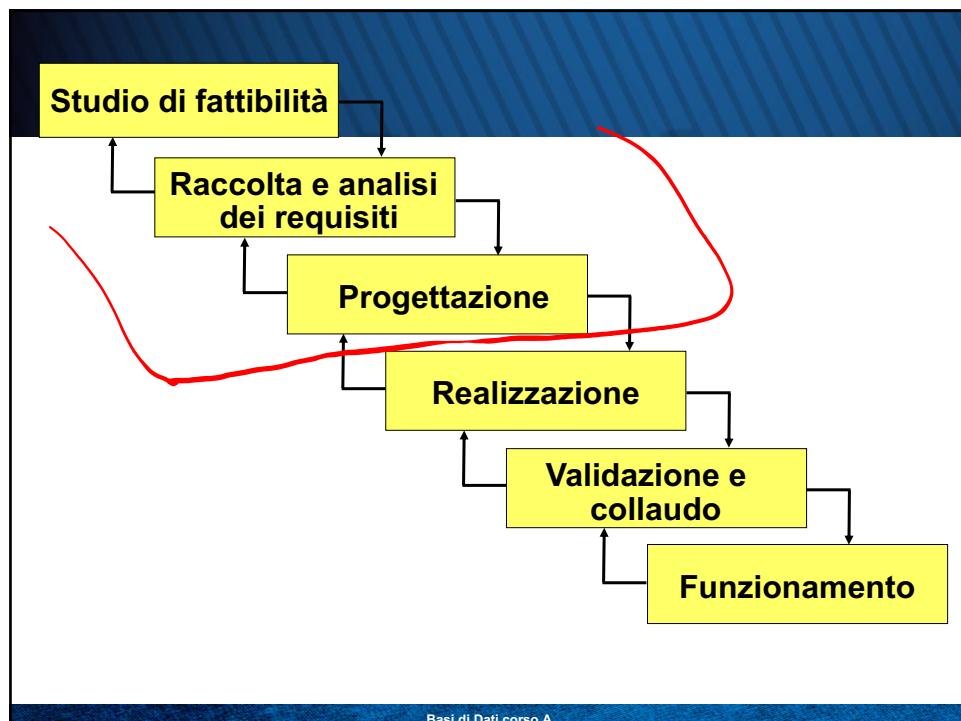
8



9

## Progettazione di basi di dati

- ▶ Una delle attività del processo di sviluppo dei sistemi informativi, va quindi inquadrata in un contesto più generale:
- ▶ **il ciclo di vita dei sistemi informativi:**
  - ▶ Insieme e sequenzializzazione delle attività svolte da analisti, progettisti, utenti, nello sviluppo e nell'uso dei sistemi informativi
  - ▶ **attività iterativa**



12

## Fasi del ciclo di vita

- **Studio di fattibilità:** definizione costi e priorità
- **Raccolta e analisi dei requisiti:** studio delle proprietà del sistema
- **Progettazione:** di dati e funzioni
- **Realizzazione**
- **Validazione e collaudo:** sperimentazione
- **Funzionamento:** il sistema diventa operativo

Basi di Dati corso A

13

## Esempio di cronoprogramma

ALLEGATO B: CRONOPROGRAMMA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Predicazione dei bandi per le forniture (Dipartimento di Chimica)																								
Licitazione delle gare e delle procedure autorizzative (Dipartimento di Chimica)																								
Installazione della strumentazione ed esecuzione delle opere edili (Dipartimento di Chimica)																								
Realizzazione del prototipo dei sistemi informatici (Dipartimento di Chimica e Dipartimento di Informatica)																								
Training dei personale (Dipartimento di Chimica)																								
Messa a punto delle macchine (Dipartimento di Chimica)																								
Studi applicativi sulle tipologie di materiali, sulle tecnologie e maggior impegno, sull'esperienza e conoscenze di campagne ad esempio per la qualità dell'aria (Dipartimento di Chimica e Dipartimento di Medicina)																								
Realizzazione del progetto dimostratore (Dipartimento di Chimica)																								

Basi di Dati corso A

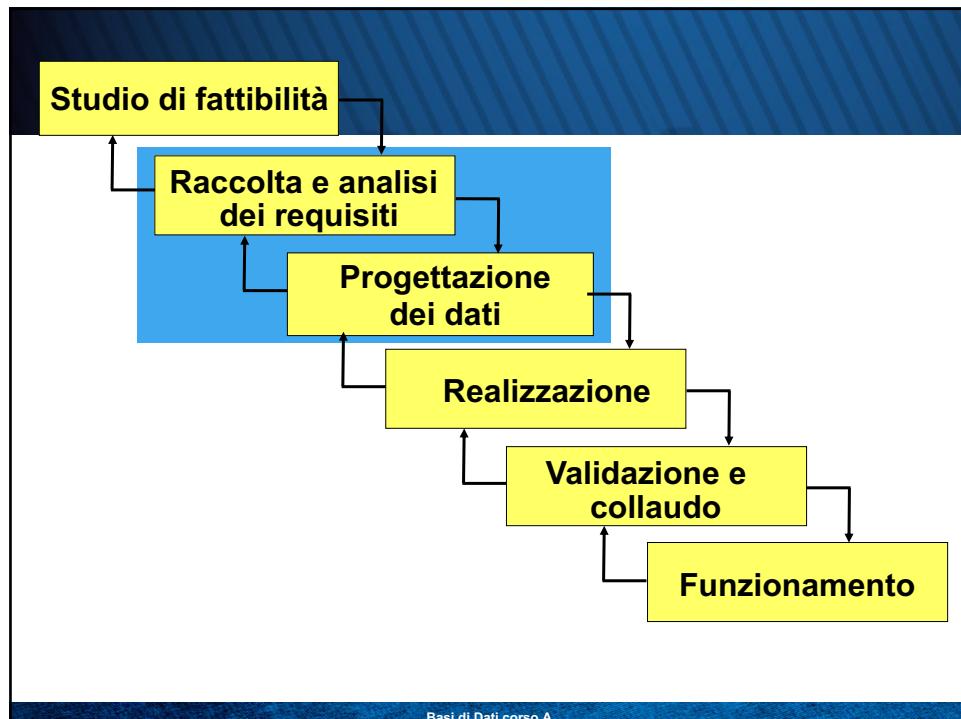
14

La progettazione di un sistema informativo riguarda due aspetti:

- ▶ progettazione dei dati
- ▶ progettazione delle applicazioni
- ▶ i dati hanno un ruolo centrale
- ▶ i dati sono più stabili

Basi di Dati corso A

15



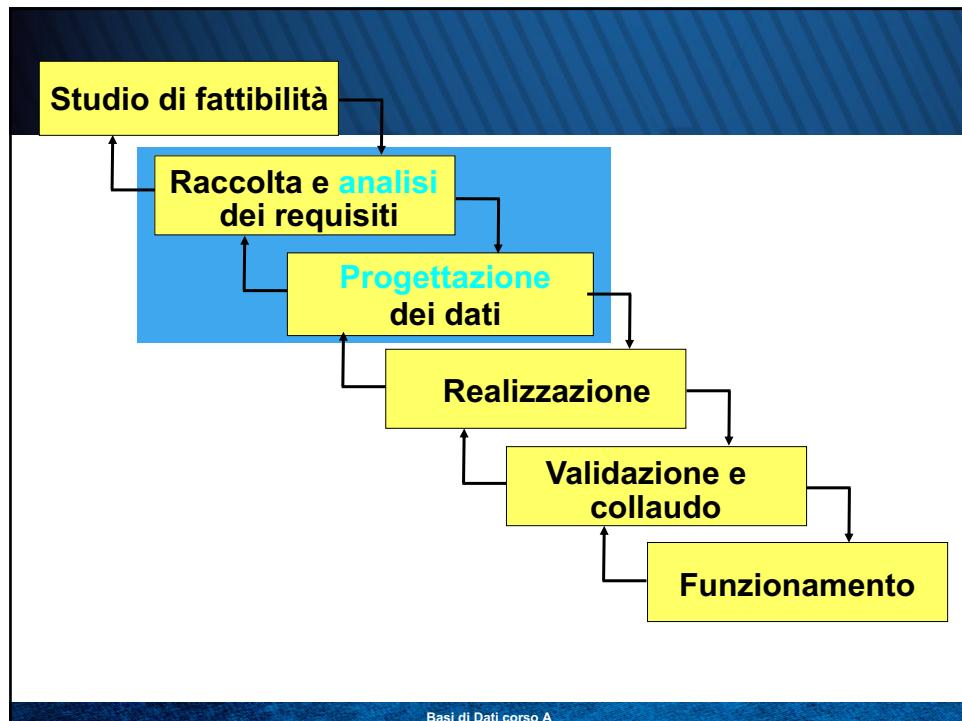
16

## Metodologia di progetto

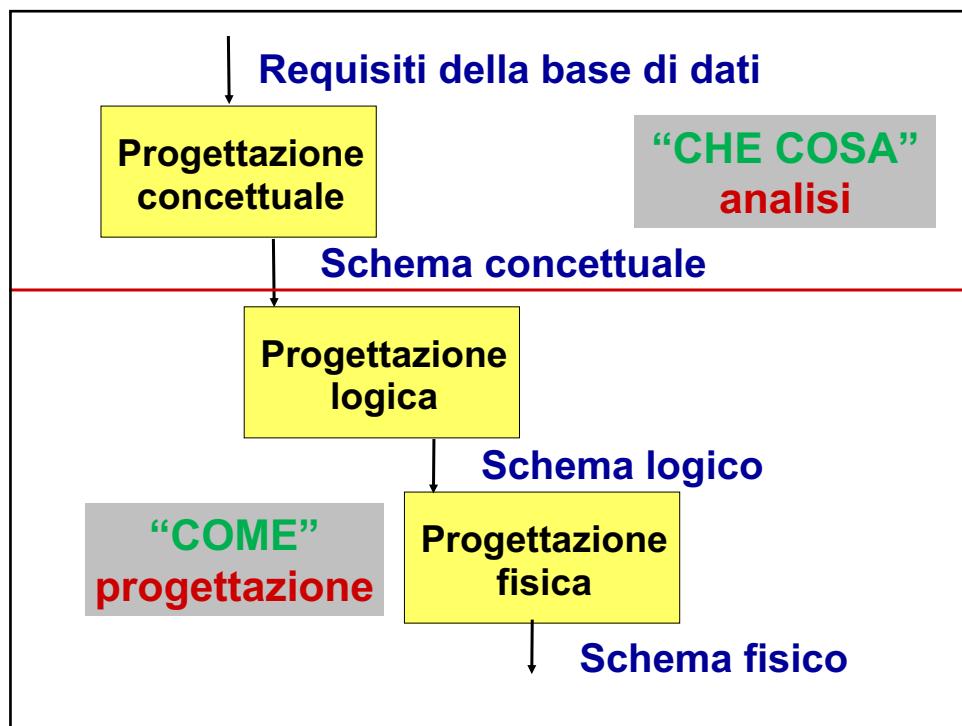
- ▶ Per garantire prodotti di buona qualità è opportuno seguire una
  - ▶ **metodologia di progetto**, con:
    - ▶ articolazione delle attività in fasi
    - ▶ criteri di scelta
    - ▶ modelli di rappresentazione
    - ▶ generalità e facilità d'uso

Basi di Dati corso A

17



18



19

I prodotti della varie fasi sono  
schemi di alcuni **modelli di dati**:

- ▶ Schema concettuale
- ▶ Schema logico
- ▶ Schema fisico

Basi di Dati corso A

20

## Progettare e implementare DB

- ▶ Phase 1: Requirements Collection and Analysis
- ▶ Phase 2: Conceptual Database Design
- ▶ Phase 3: Choice of DBMS
- ▶ Phase 4: Data Model Mapping (Logical Database Design)
- ▶ Phase 5: Physical Database Design
- ▶ Phase 6: Database System Implementation and Tuning

Basi di Dati corso A

21

## 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
- ▶ ad esempio, il **modello relazionale** prevede il costruttore **relazione**, che permette di definire insiemi di record omogenei

Basi di Dati corso A

22

## Schemi e istanze

- ▶ In ogni base di dati esistono:
  - ▶ lo **schema**, sostanzialmente invariante nel tempo, che ne descrive la struttura (aspetto intensionale)
    - ▶ nel modello relazionale, le intestazioni delle tabelle
  - ▶ l'**istanza**, i valori attuali, che possono cambiare anche molto rapidamente (aspetto estensionale)
    - ▶ nel modello relazionale, il “corpo” di ciascuna tabella

Basi di Dati corso A

23

## Due tipi (principali) di modelli

- ▶ **modelli logici:** utilizzati nei DBMS esistenti per l'organizzazione dei dati
  - ▶ utilizzati dai programmi
  - ▶ indipendenti dalle strutture fisiche
  - esempi: [relazionale](#), reticolare, gerarchico, a oggetti
- ▶ **modelli concettuali:** permettono di rappresentare i dati in modo indipendente da ogni sistema
  - ▶ cercano di descrivere i concetti del mondo reale
  - ▶ sono utilizzati nelle fasi preliminari di progettazione
  - il più noto è il modello [Entity-Relationship](#)

Basi di Dati corso A

24

## Modelli concettuali, perché?

- ▶ Proviamo a modellare una applicazione definendo direttamente lo schema logico della base di dati:
  - ▶ da dove cominciamo?
  - ▶ rischiamo di perderci subito nei dettagli
  - ▶ dobbiamo pensare subito a come correlare le varie tabelle (chiavi etc.)
  - ▶ i modelli logici sono rigidi

Basi di Dati corso A

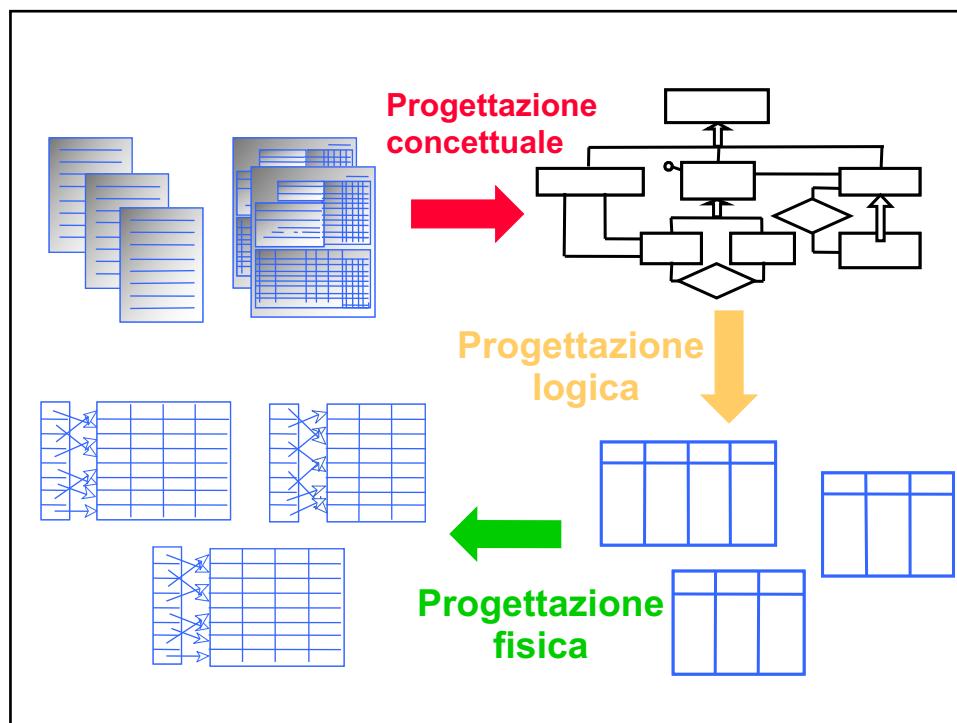
25

## Modelli concettuali, perché?

- ▶ servono per ragionare sulla realtà di interesse, indipendentemente dagli aspetti realizzativi
- ▶ permettono di rappresentare le classi di dati di interesse e le loro correlazioni
- ▶ prevedono efficaci rappresentazioni grafiche (utili anche per documentazione e comunicazione)

Basi di Dati corso A

26



27

## Modello Entity-Relationship (Entità-Relazione)

► Il più diffuso modello concettuale

- Ne esistono molte versioni
- (più o meno) diverse l'una dall'altra

Basi di Dati corso A

28

## I costrutti del modello E-R

- Entità
- Relationship
- Attributo
- Identificatore
- Generalizzazione
- ....

(pag. 204 V ed. Atzeni l'elenco sintetico)

Basi di Dati corso A

29

## Entità

- ▶ Classe di oggetti (es. fatti, persone, cose) della applicazione di interesse con proprietà comuni e con esistenza “autonoma”
- ▶ Esempi:
  - ▶ impiegato, città, conto corrente, ordine, fattura

Basi di Dati corso A

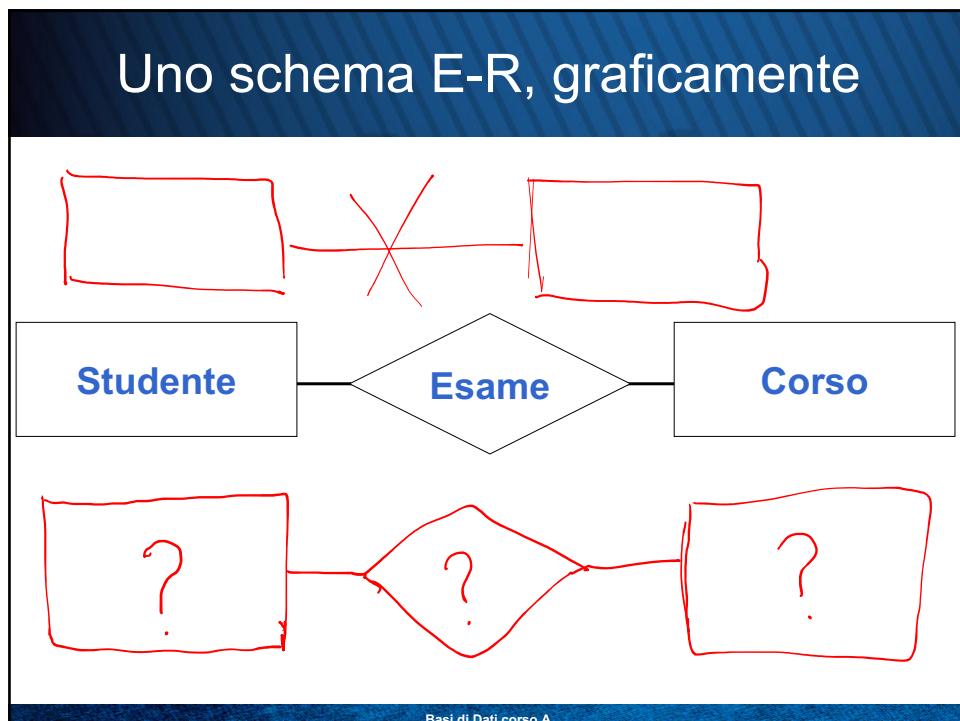
30

## Relationship

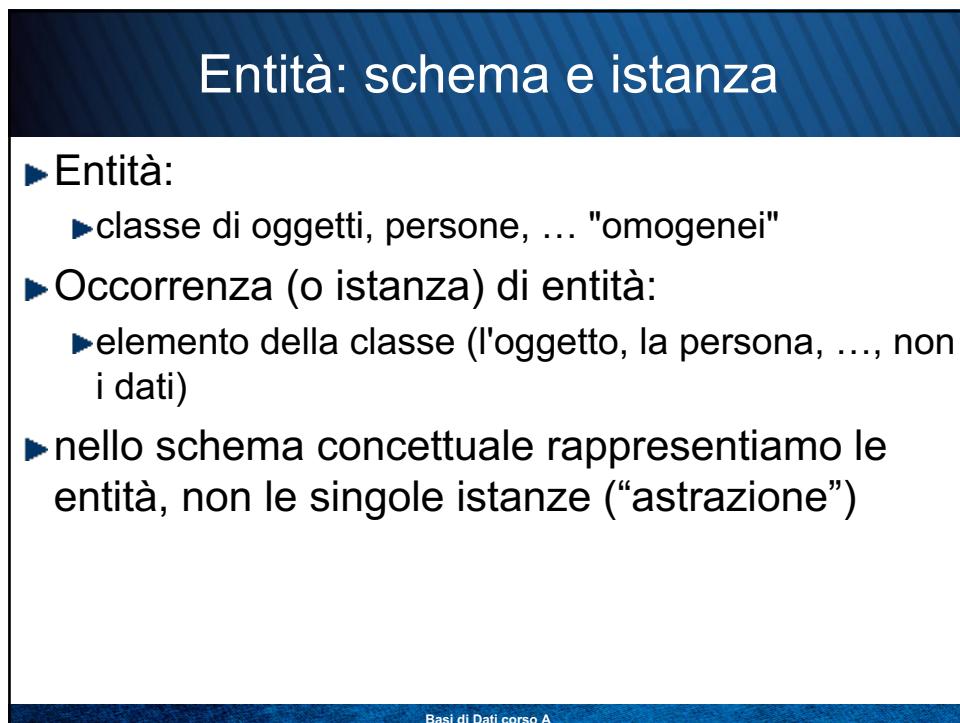
- ▶ Legame logico fra due o più entità, rilevante nell'applicazione di interesse
- ▶ Esempi:
  - ▶ Residenza (fra persona e città)
  - ▶ Esame (fra studente e corso)

Basi di Dati corso A

31



32



33

## Rappresentazione grafica di entità

Impiegato

Dipartimento

Città

Vendita

Basi di Dati corso A

34

## Entità, commenti

- Ogni entità ha un nome che la identifica univocamente nello schema:
  - nomi espressivi
  - opportune convenzioni
  - singolare

Basi di Dati corso A

35

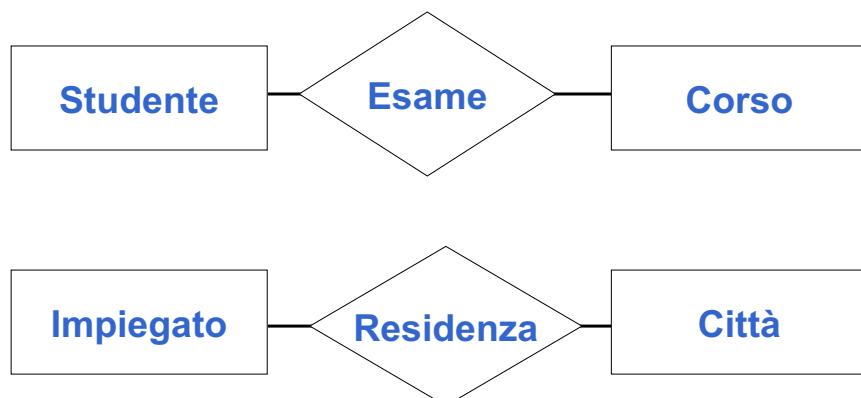
## Relationship

- ▶ Legame logico fra due o più entità, rilevante nell'applicazione di interesse
- ▶ Esempi:
  - ▶ Residenza (fra persona e città)
  - ▶ Esame (fra studente e corso)
- ▶ Chiamata anche:
  - ▶ relazione, correlazione, associazione

Basi di Dati corso A

36

## Rappresentazione grafica di relationship



Basi di Dati corso A

37

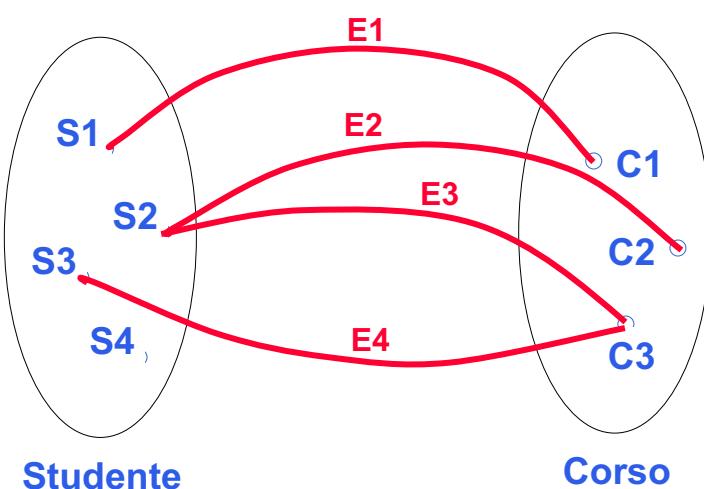
## Relationship, commenti

- ▶ Ogni relationship ha un nome che la identifica univocamente nello schema:
  - ▶ nomi espressivi
  - ▶ opportune convenzioni
    - ▶ singolare
    - ▶ sostantivi invece che verbi (se possibile)

Basi di Dati corso A

38

## Esempi di occorrenze



Basi di Dati corso A

39

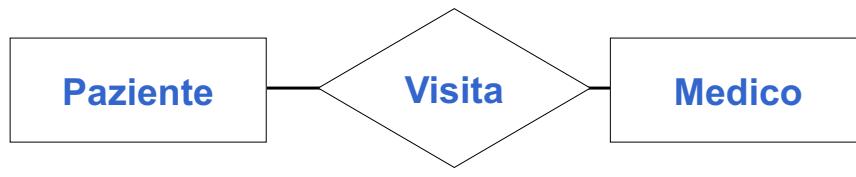
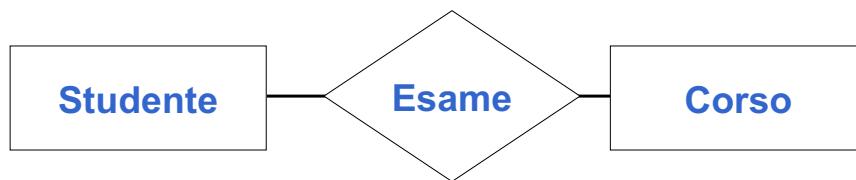
## Relationship, occorrenze

- ▶ Una occorrenza di una relationship binaria è coppia di occorrenze di entità, una per ciascuna entità coinvolta
- ▶ Una occorrenza di una relationship n-aria è una n-upla di occorrenze di entità, una per ciascuna entità coinvolta
- ▶ Nell'ambito di una relationship non ci possono essere occorrenze (coppie, ennuple) ripetute

Basi di Dati corso A

40

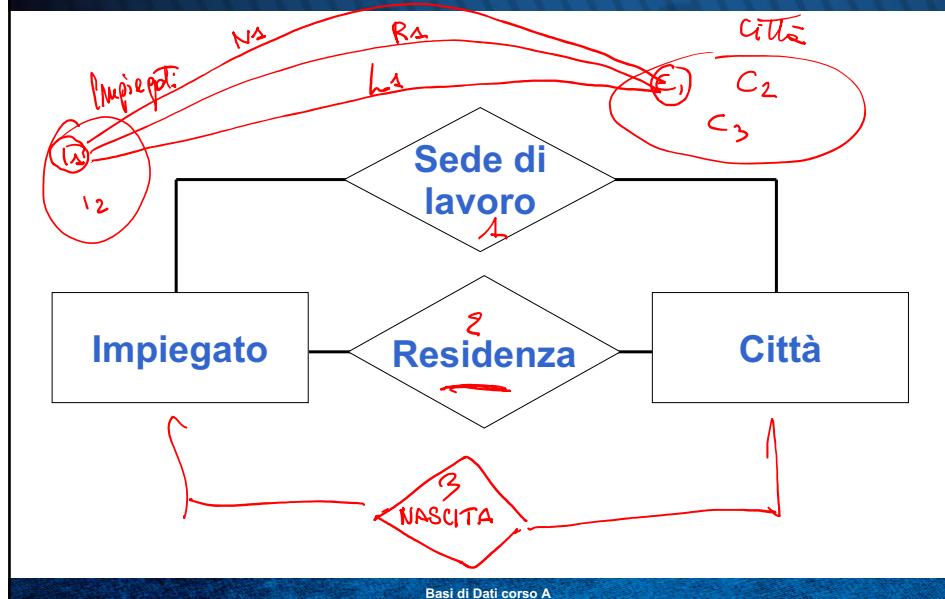
## Relationship corrette?



Basi di Dati corso A

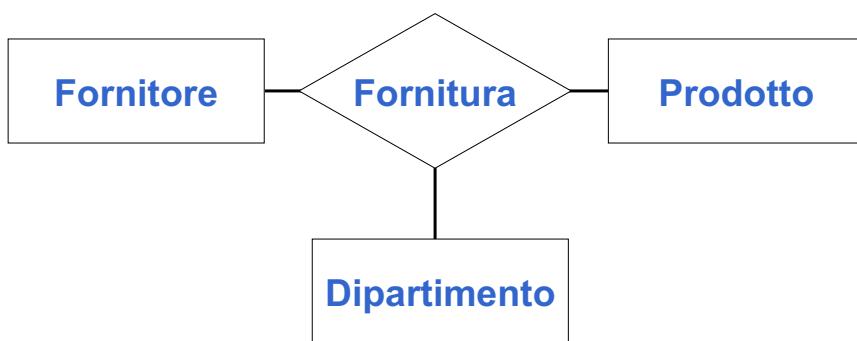
41

## Due relationship sulle stesse entità



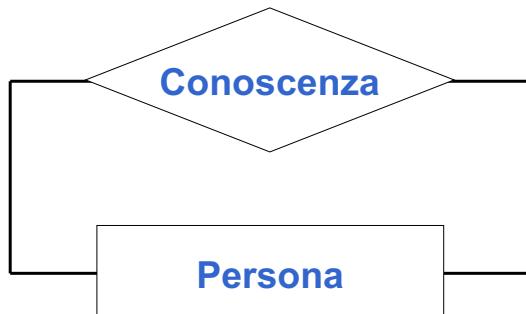
42

## Relationship n-aria



43

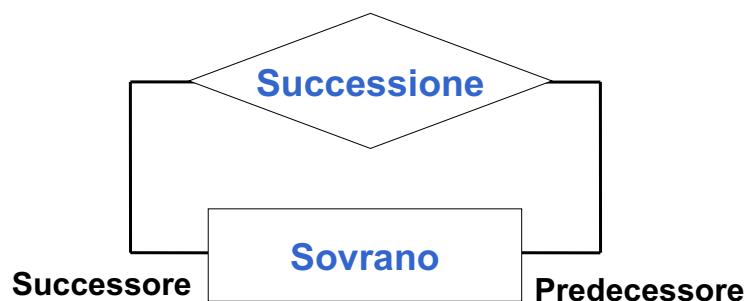
## Relationship ricorsiva: coinvolge “due volte” la stessa entità



Basi di Dati corso A

44

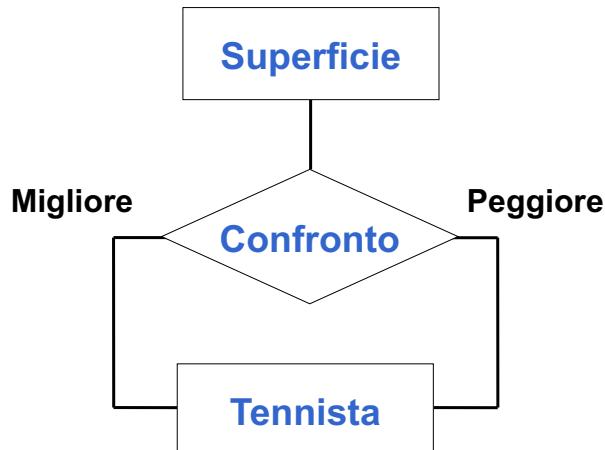
## Relationship ricorsiva con “ruoli”



Basi di Dati corso A

45

## Relationship ternaria ricorsiva



Basi di Dati corso A

46

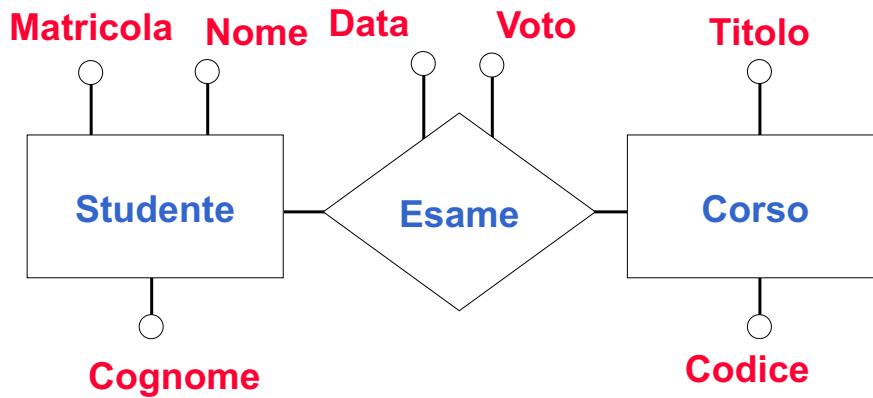
## Attributo

- ▶ Proprietà elementare di un'entità o di una relationship, di interesse ai fini dell'applicazione
- ▶ Associa ad ogni occorrenza di entità o relationship un valore appartenente a un insieme detto **dominio** dell'attributo

Basi di Dati corso A

47

## Attributi, rappresentazione grafica



Basi di Dati corso A

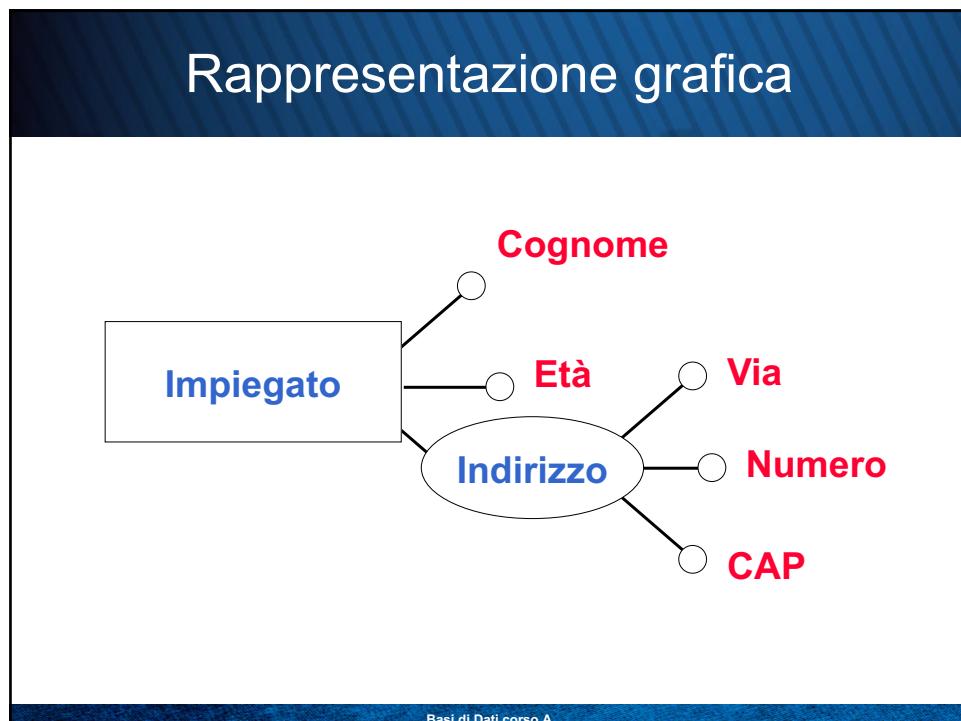
48

## Attributi composti

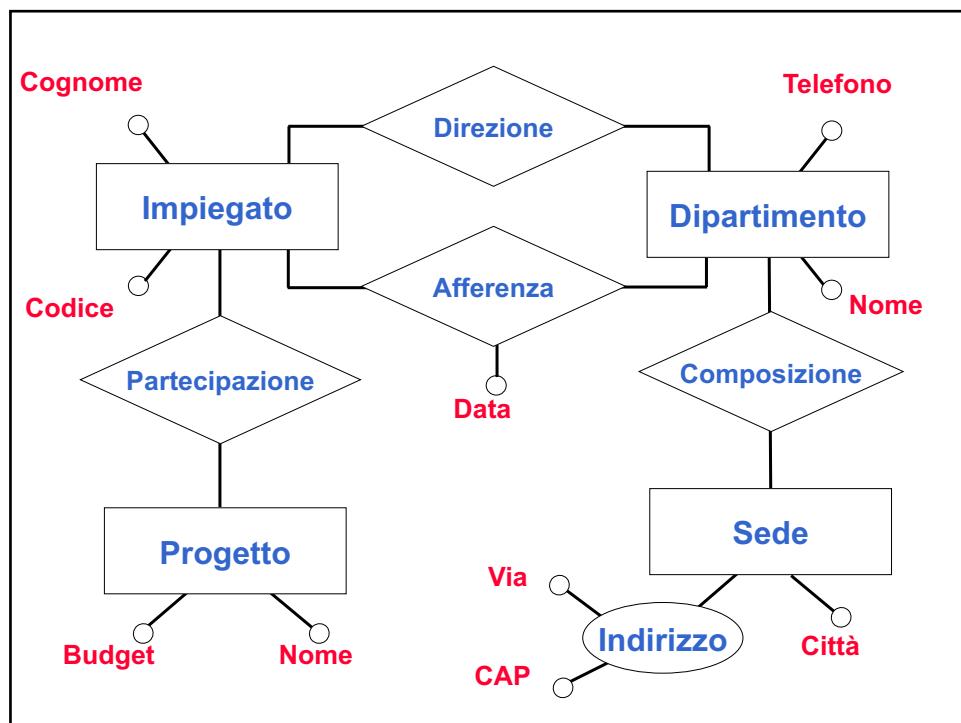
- ▶ Raggruppano attributi di una medesima entità o relationship che presentano affinità nel loro significato o uso
- ▶ Esempio:
  - ▶ Via, Numero civico e CAP formano un Indirizzo

Basi di Dati corso A

49



50



51

## Altri costrutti del modello E-R

- ▶ **Cardinalità**
  - ▶ di relationship
  - ▶ di attributo
- ▶ **Identificatore**
  - ▶ interno
  - ▶ esterno
- ▶ **Generalizzazione**

Basi di Dati corso A

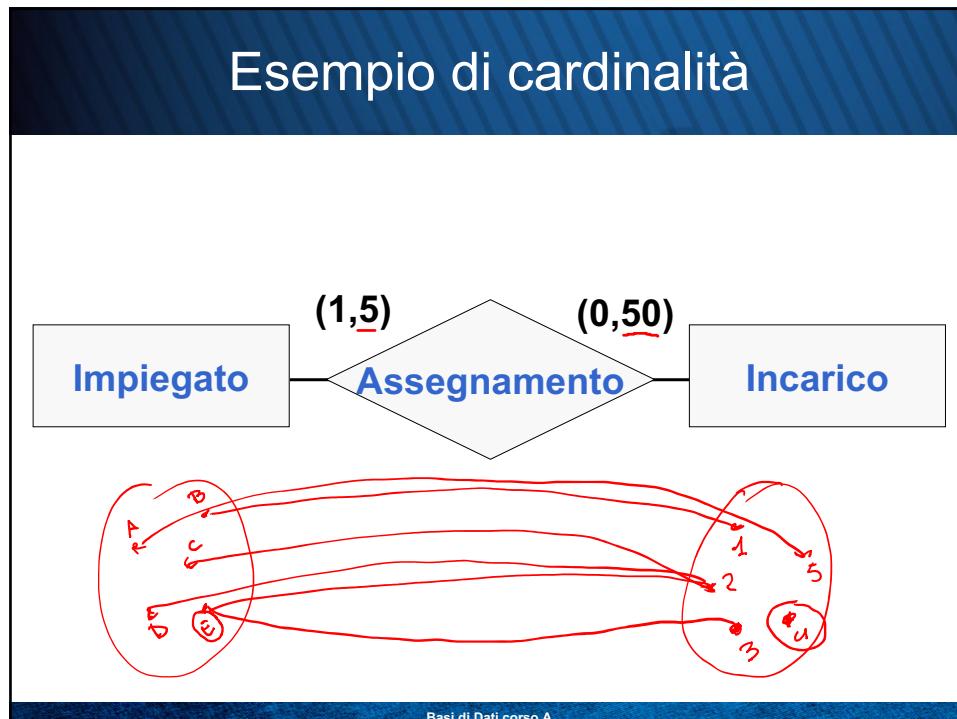
52

## Cardinalità di relationship

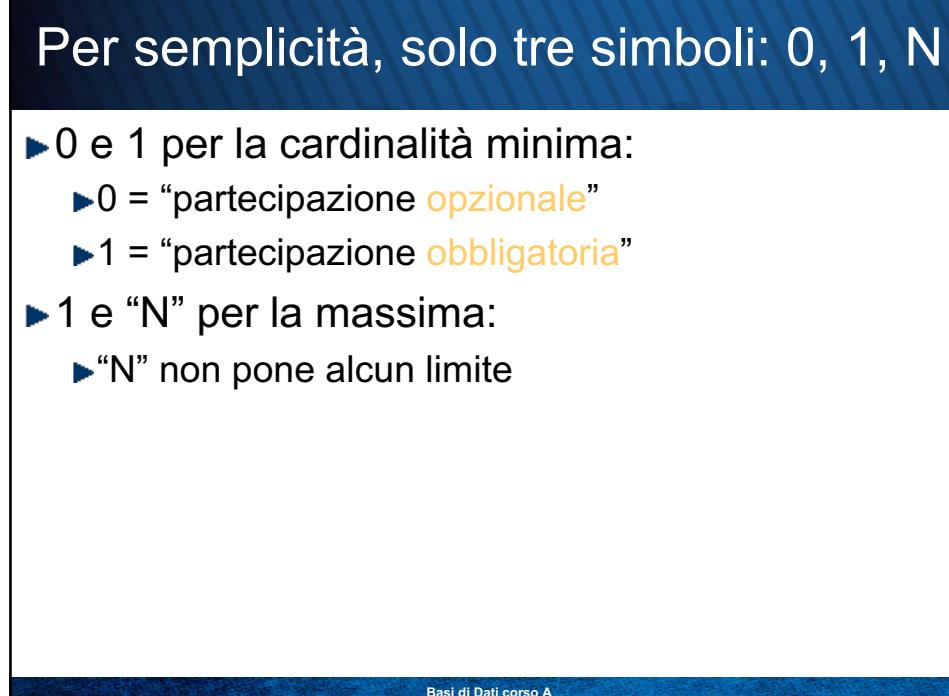
- ▶ Coppia di valori associati a ogni entità che partecipa a una relationship
- ▶ specificano il numero minimo e massimo di occorrenze delle relationship cui ciascuna occorrenza di una entità può partecipare

Basi di Dati corso A

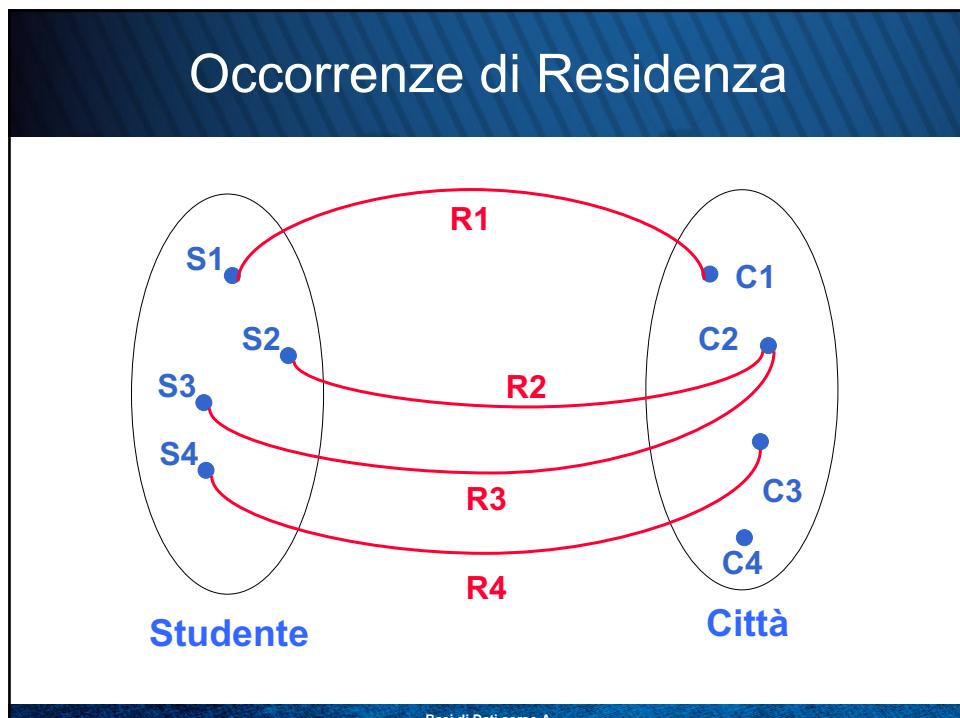
53



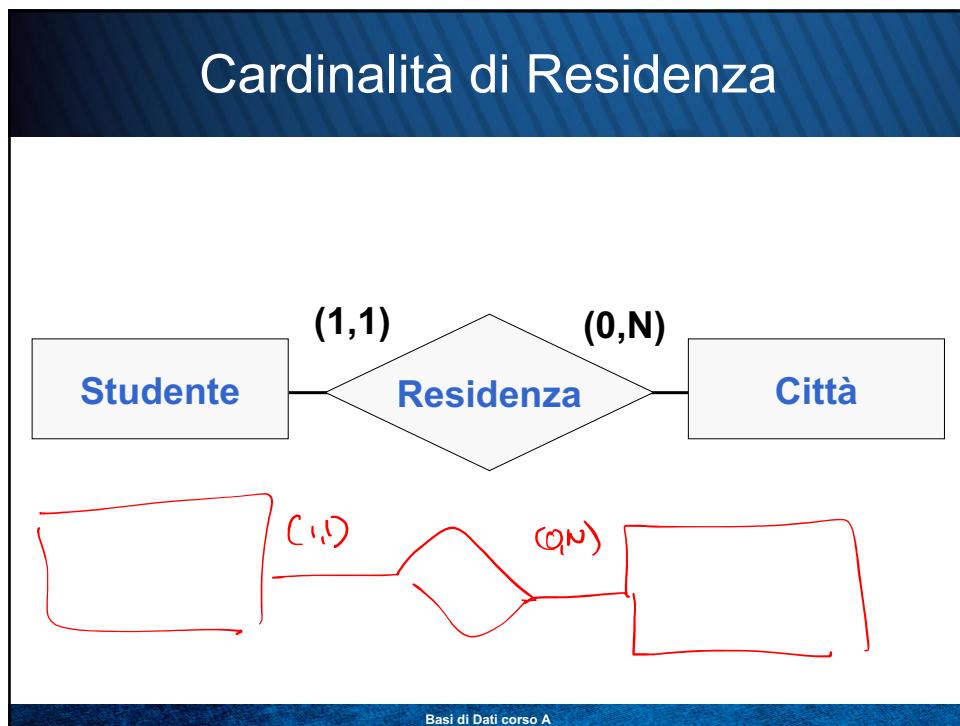
54



55



56



57

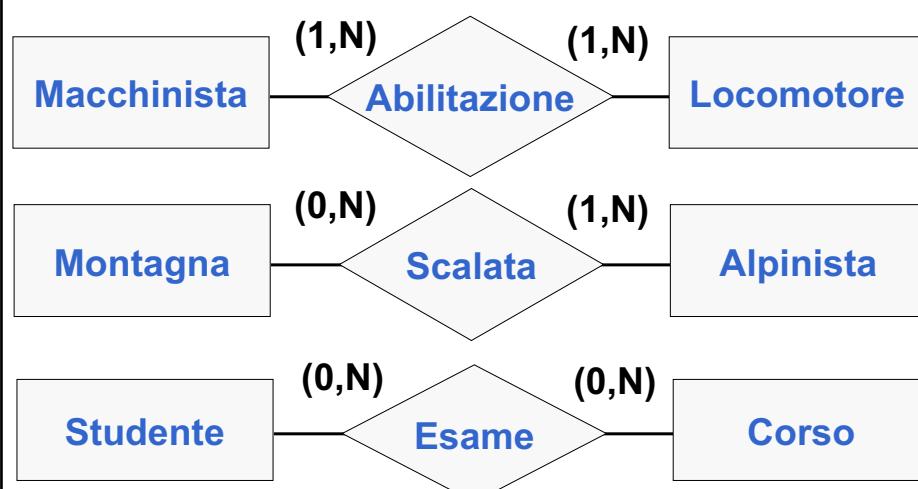
## Tipi di relationship

- ▶ Con riferimento alle cardinalità **massime**, abbiamo relationship:
  - ▶ uno a uno
  - ▶ uno a molti
  - ▶ molti a molti

Basi di Dati corso A

58

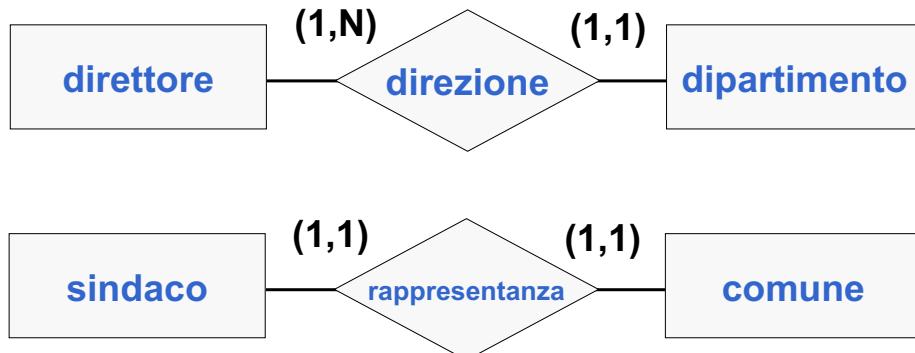
## Relationship “molti a molti”



Basi di Dati corso A

59

## Relationship “uno a molti e uno a uno”



Basi di Dati corso A

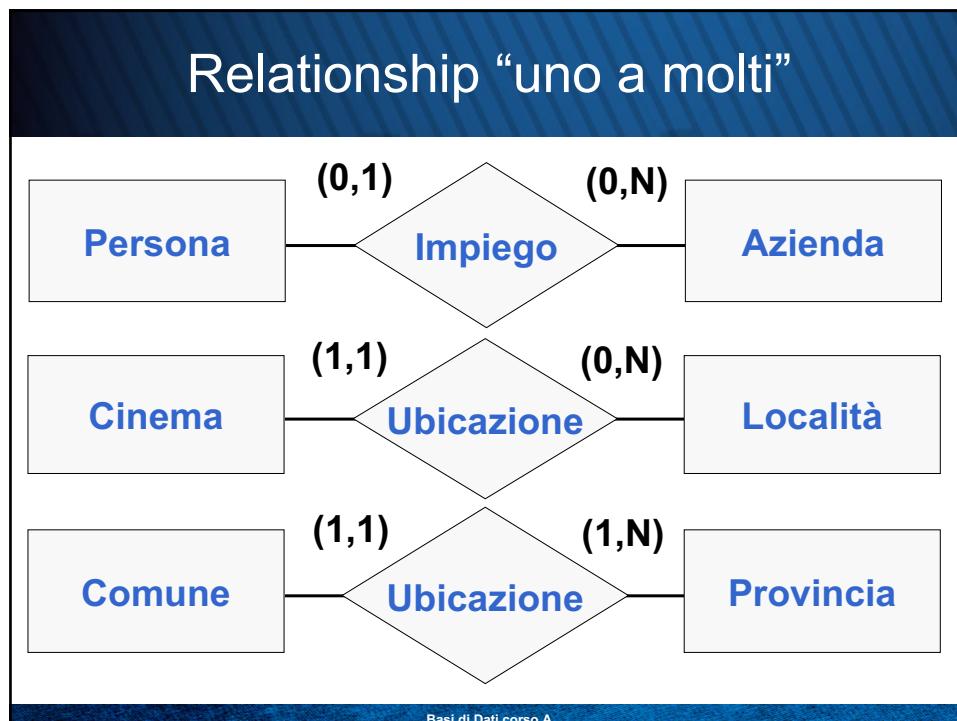
60

## Due avvertenze

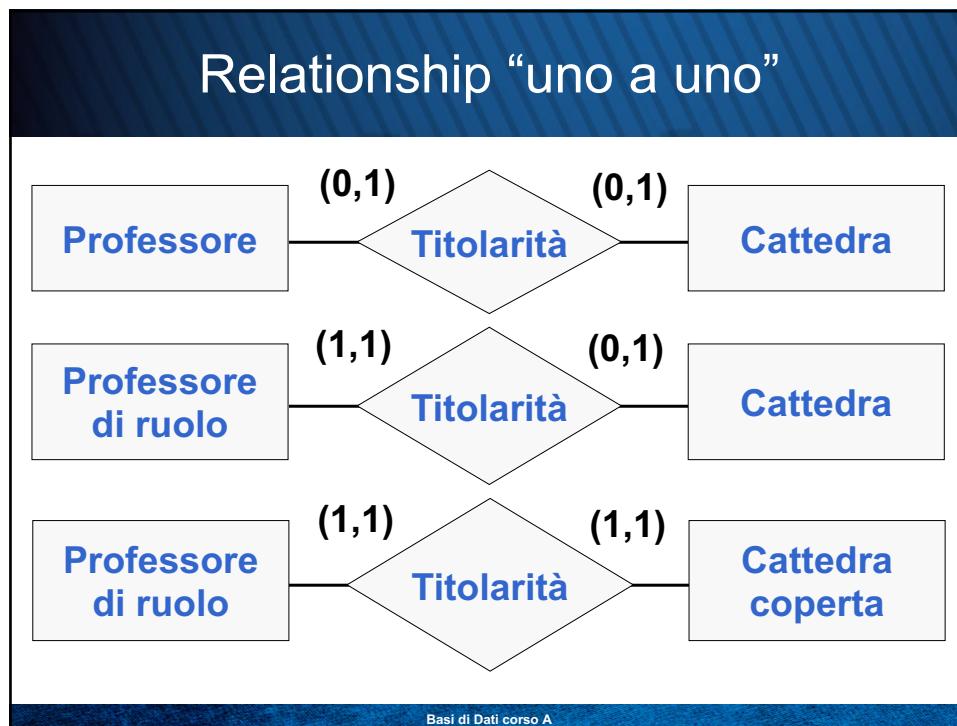
- ▶ Attenzione al "verso" nelle relationship uno a molti
- ▶ le relationship obbligatorie-obbligatorie sono molto rare

Basi di Dati corso A

61



62



63

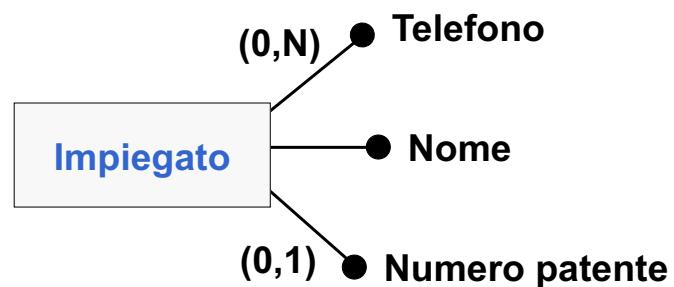
## Cardinalità di attributi

- ▶ È possibile associare delle cardinalità anche agli attributi, con due scopi:
  - ▶ indicare opzionalità ("informazione incompleta")
  - ▶ indicare attributi multivalore

Basi di Dati corso A

64

## Rappresentazione grafica



Basi di Dati corso A

65

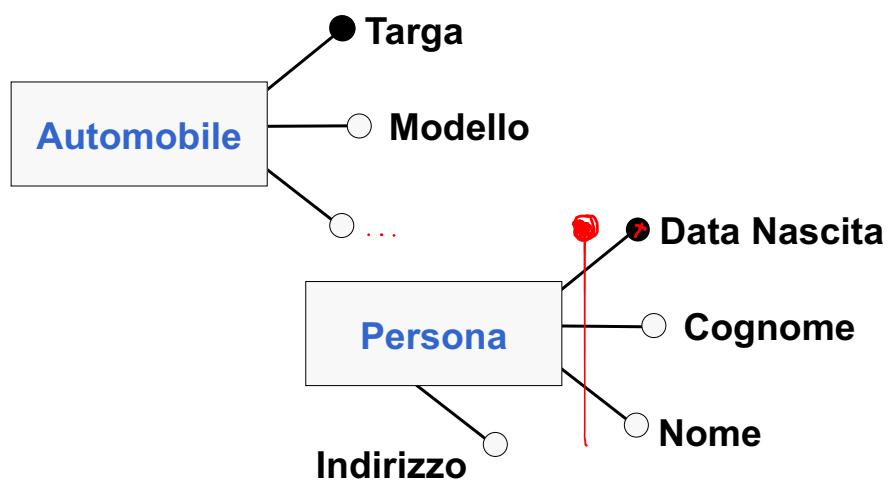
## Identificatore di una entità

- ▶ “strumento” per l’identificazione univoca delle occorrenze di un’entità
- ▶ costituito da:
  - ▶ attributi dell’entità
    - ▶ identificatore interno
  - ▶ (attributi +) entità esterne attraverso relationship
    - ▶ identificatore esterno

Basi di Dati corso A

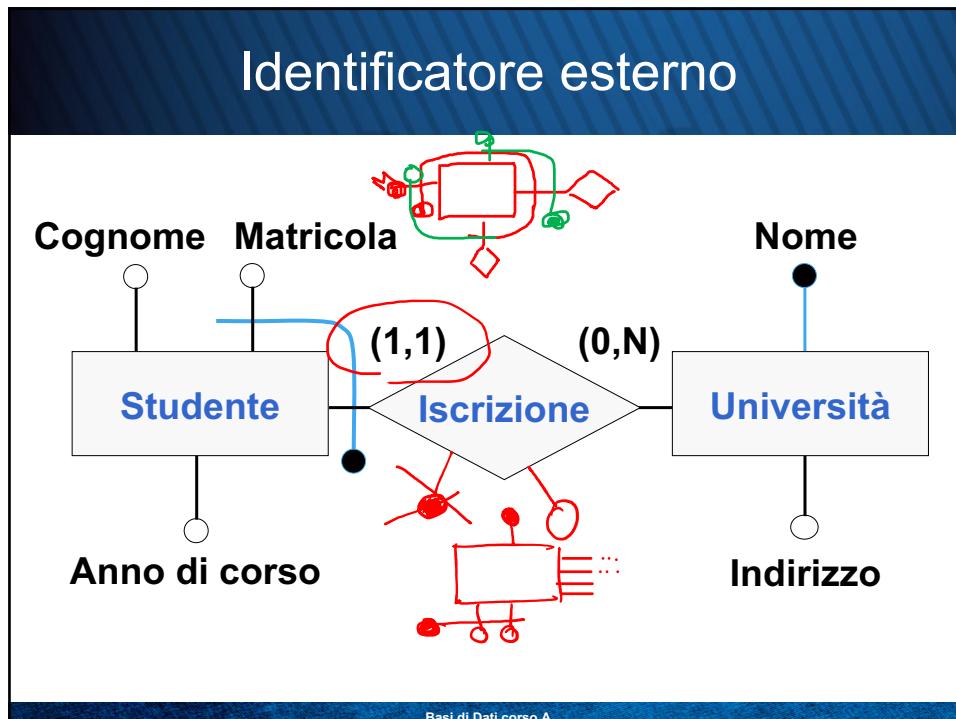
66

## Identifieri interni

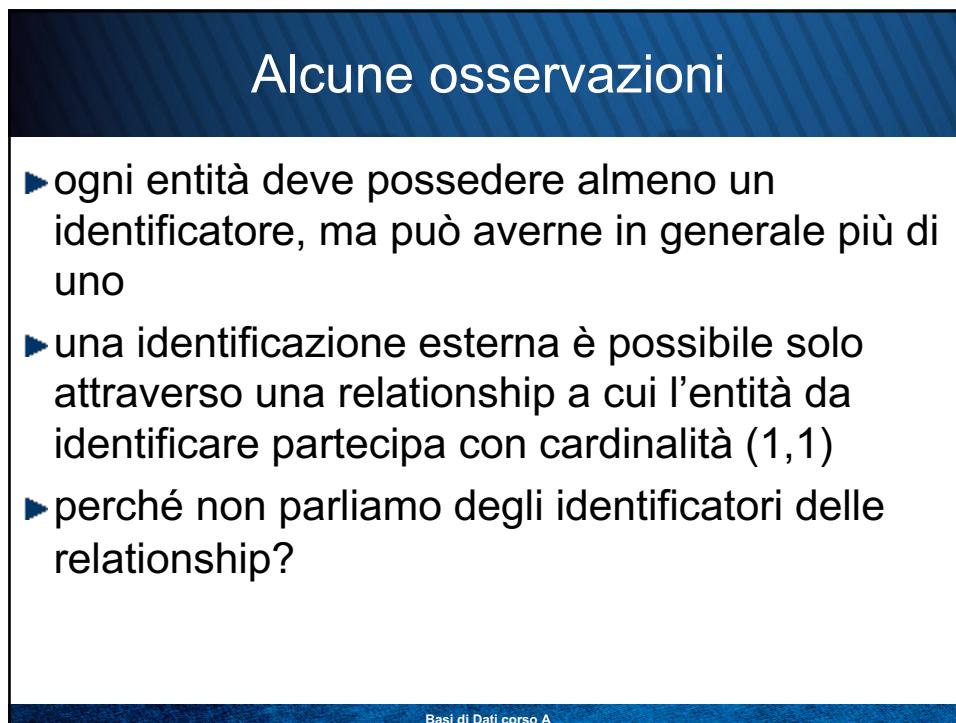


Basi di Dati corso A

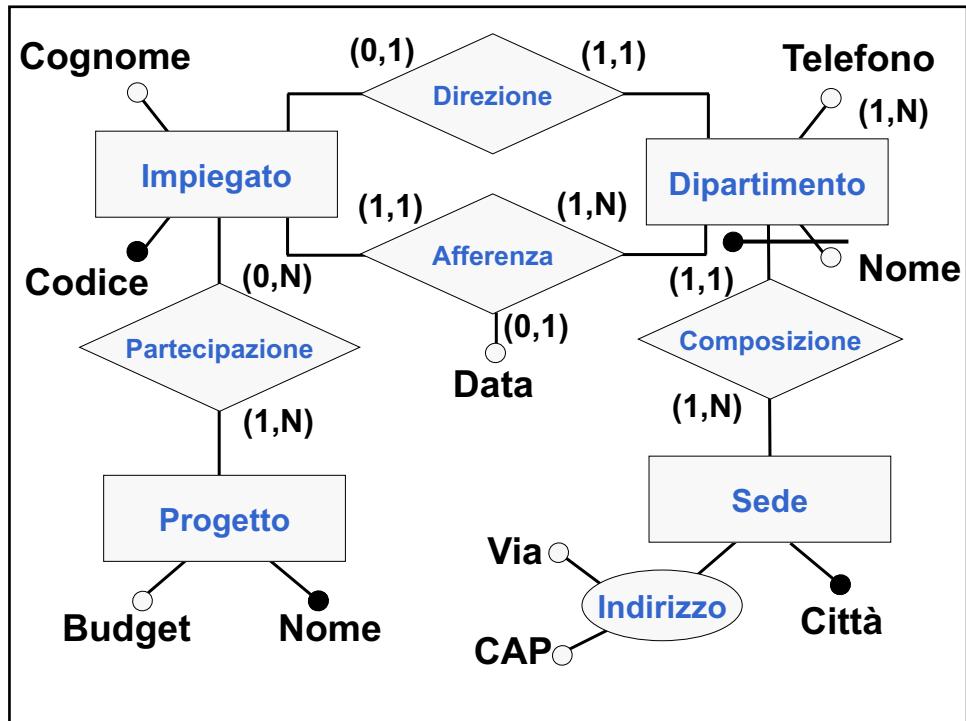
67



68



69



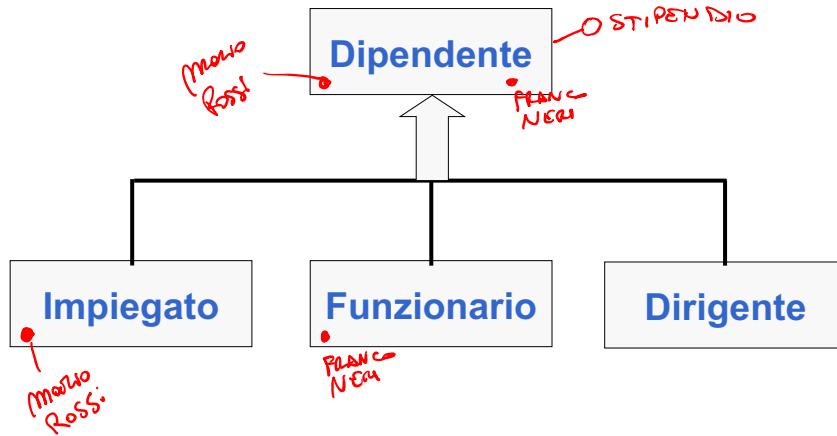
70

## Generalizzazione

- ▶ mette in relazione una o più entità E1, E2, ..., En con una entità E, che le comprende come casi particolari
  - ▶ E è **generalizzazione** di E1, E2, ..., En
  - ▶ E1, E2, ..., En sono **specializzazioni** (o sottotipi) di E

71

## Rappresentazione grafica



Basi di Dati corso A

72

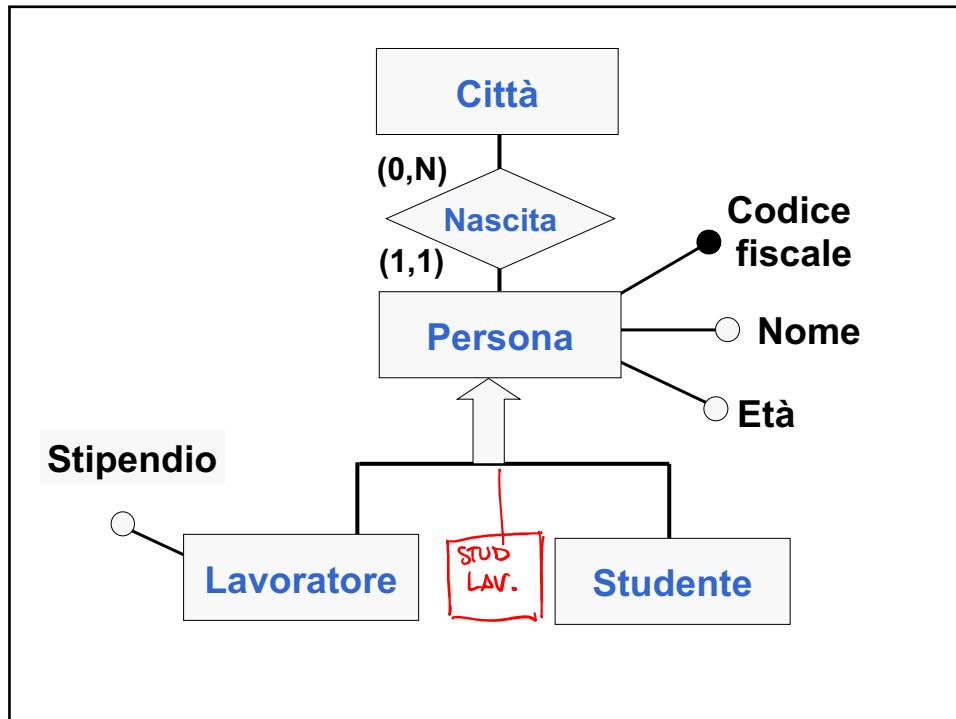
## Proprietà delle generalizzazioni

Se E (genitore) è generalizzazione di E<sub>1</sub>, E<sub>2</sub>, ..., E<sub>n</sub> (figlie):

- ▶ ogni proprietà di E è significativa per E<sub>1</sub>, E<sub>2</sub>, ..., E<sub>n</sub>
- ▶ ogni occorrenza di E<sub>1</sub>, E<sub>2</sub>, ..., E<sub>n</sub> è occasione anche di E

Basi di Dati corso A

73



74

## Ereditarietà

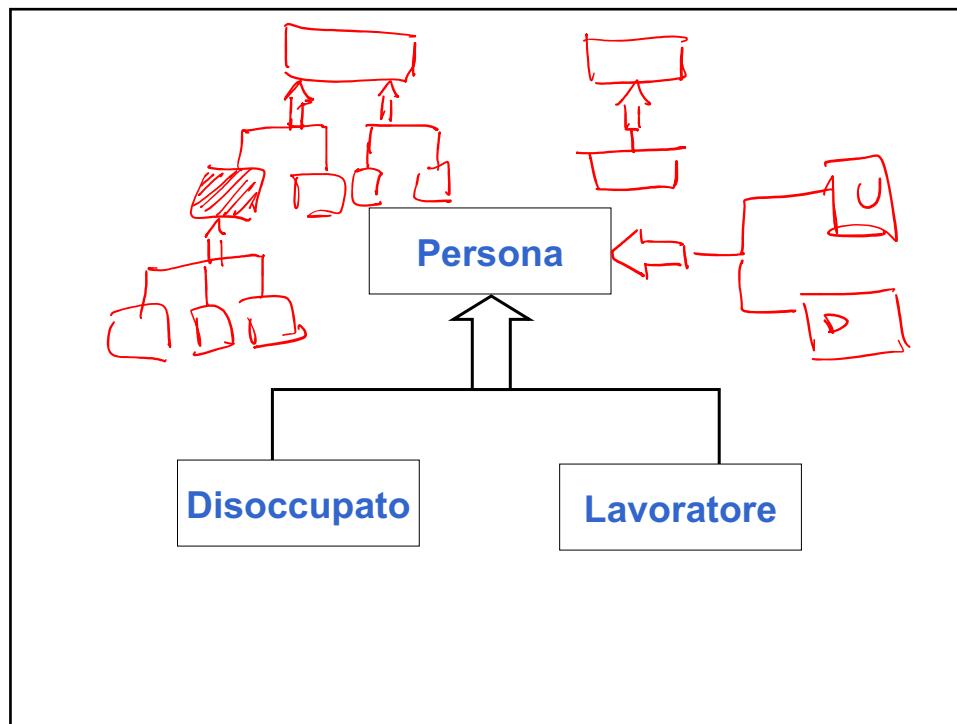
- ▶ tutte le proprietà (attributi, relationship, altre generalizzazioni) dell'entità genitore vengono **ereditate** dalle entità figlie e non rappresentate esplicitamente

## Tipi di generalizzazioni

- ▶ **totale** se ogni occorrenza dell'entità genitore è  
occorrenza di almeno una delle entità figlie,  
altrimenti è **parziale**
- ▶ **esclusiva** se ogni occorrenza dell'entità genitore è  
occorrenza di al più una delle entità figlie,  
altrimenti è **sovraposta**
- ▶ consideriamo (senza perdita di generalità) solo  
generalizzazioni esclusive e distinguiamo fra totali  
e parziali

Basi di Dati corso A

76



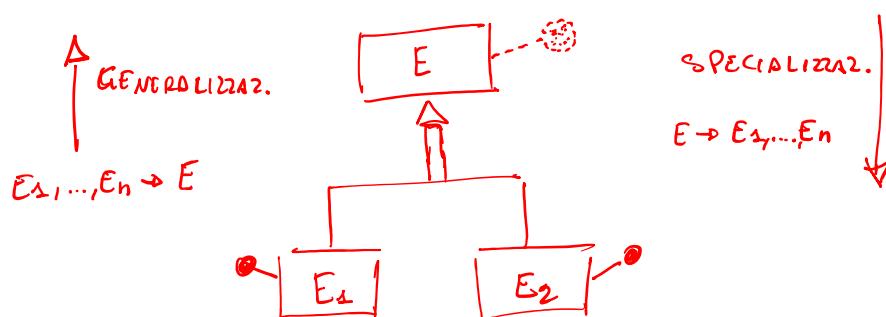
77

## Altre proprietà

- ▶ possono esistere gerarchie a più livelli e multiple generalizzazioni allo stesso livello
- ▶ un'entità può essere inclusa in più gerarchie, come genitore e/o come figlia
- ▶ se una generalizzazione ha solo un'entità figlia si parla di **sottoinsieme**
- ▶ alcune configurazioni non hanno senso
- ▶ il genitore di una generalizzazione totale può non avere identificatore, purché ...

Basi di Dati corso A

78

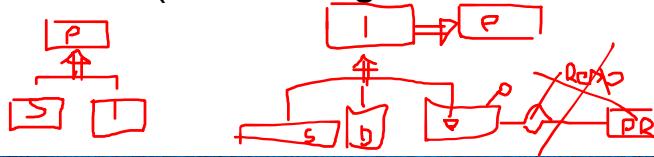


Basi di Dati corso A

79

## Esercizio

► Le persone hanno CF, cognome ed età; gli uomini anche la posizione militare; gli impiegati hanno lo stipendio e possono essere segretari, direttori o progettisti (un progettista può essere anche responsabile di progetto); gli studenti (che non possono essere impiegati) un numero di matricola; esistono persone che non sono né impiegati né studenti (ma i dettagli non ci interessano)



Basi di Dati corso A

80

## Esercizio

► Le persone hanno CF, cognome ed età; gli uomini anche la posizione militare; gli impiegati hanno lo stipendio e possono essere segretari, direttori o progettisti (un progettista può essere anche responsabile di progetto); gli studenti (che non possono essere impiegati) un numero di matricola; esistono persone che non sono né impiegati né studenti (ma i dettagli non ci interessano)



Basi di Dati corso A

81

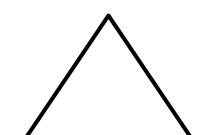
## Analisi requisiti e generazione del diagramma ER

Basi di Dati corso A

82



Totale



Parziale

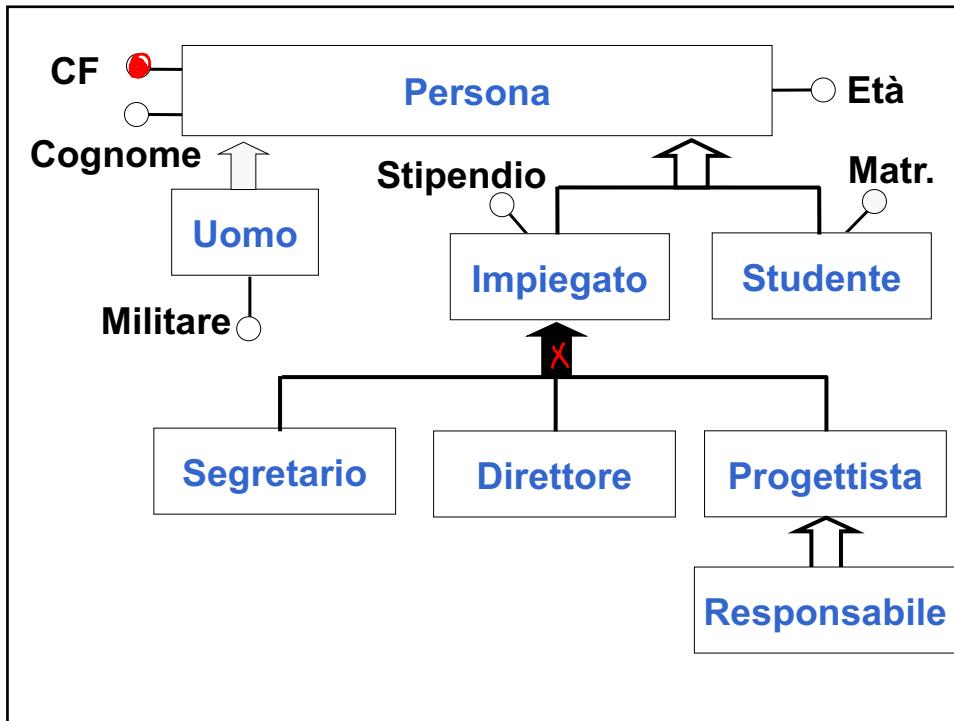
Esclusiva



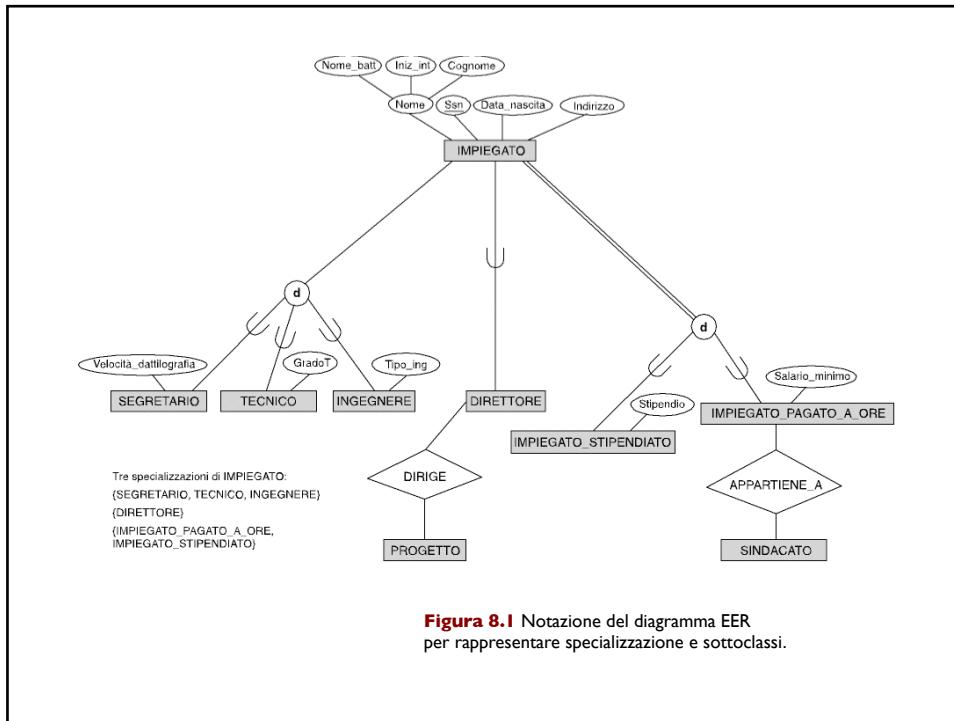
Sovrapposta



83

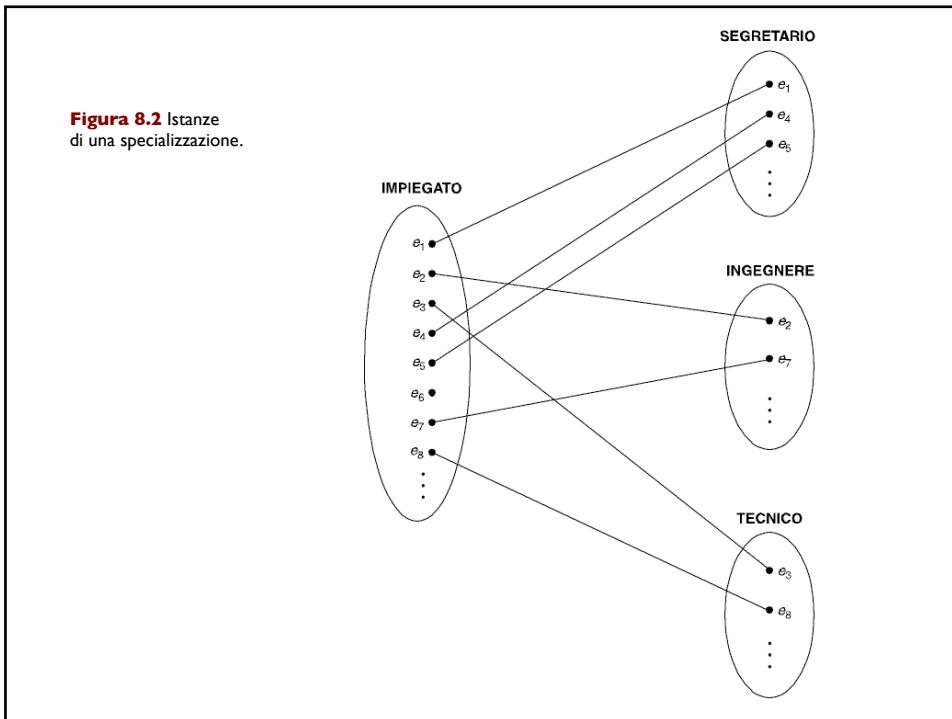


84

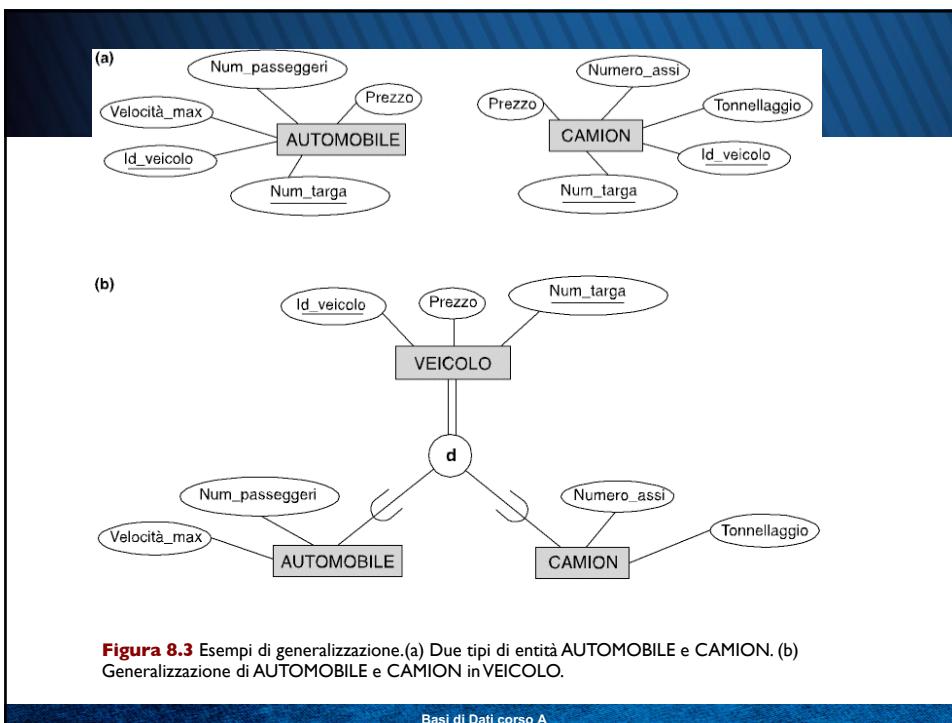


**Figura 8.1** Notazione del diagramma EER per rappresentare specializzazione e sottoclassi.

85

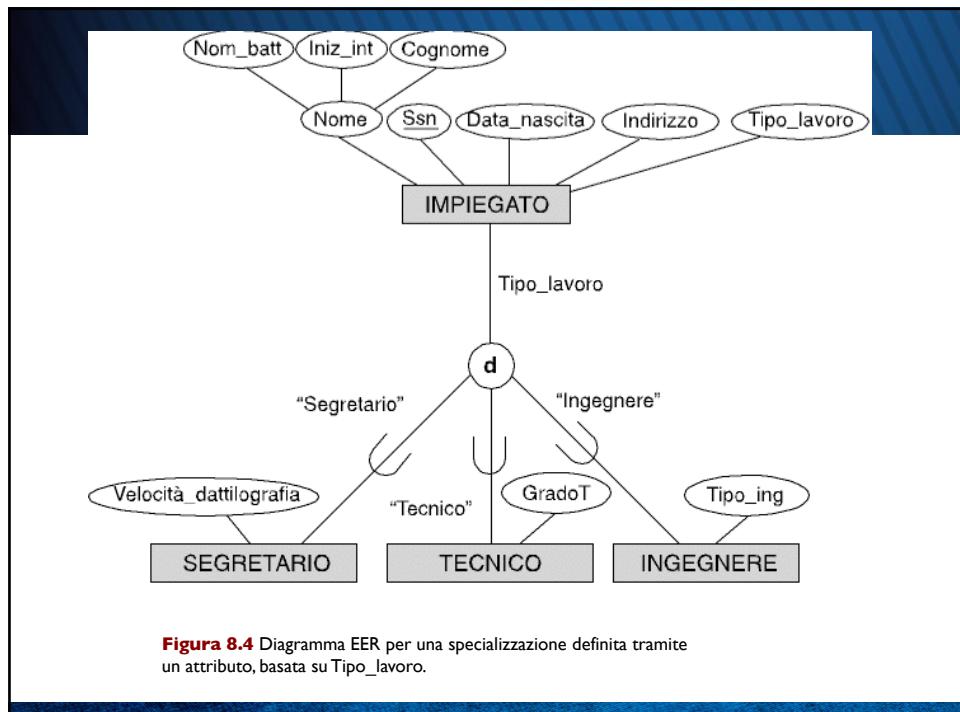


86

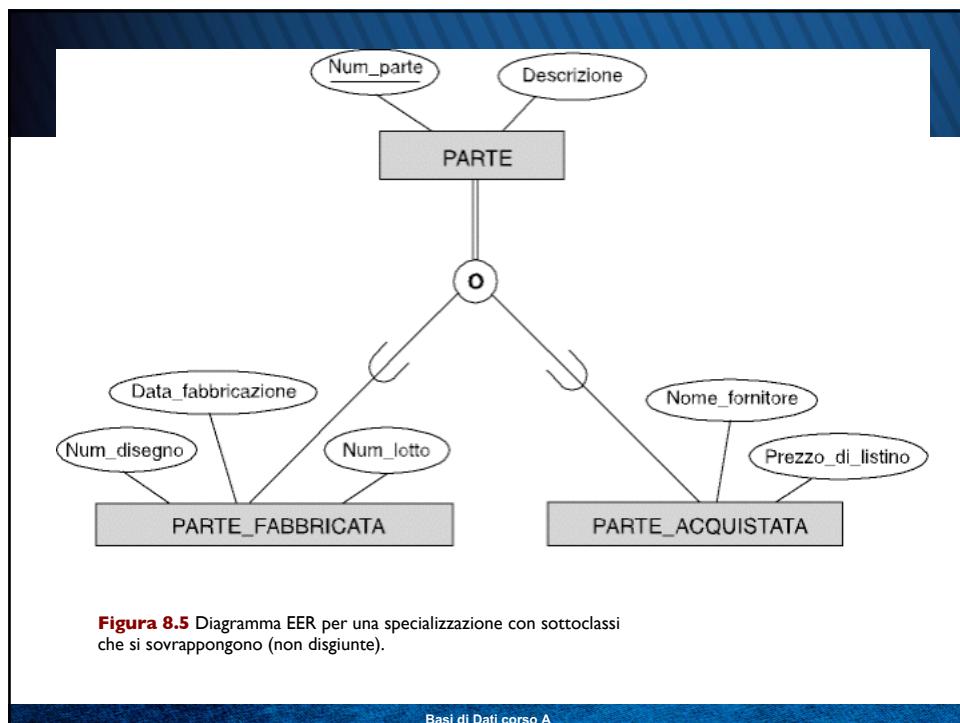


Basi di Dati corso A

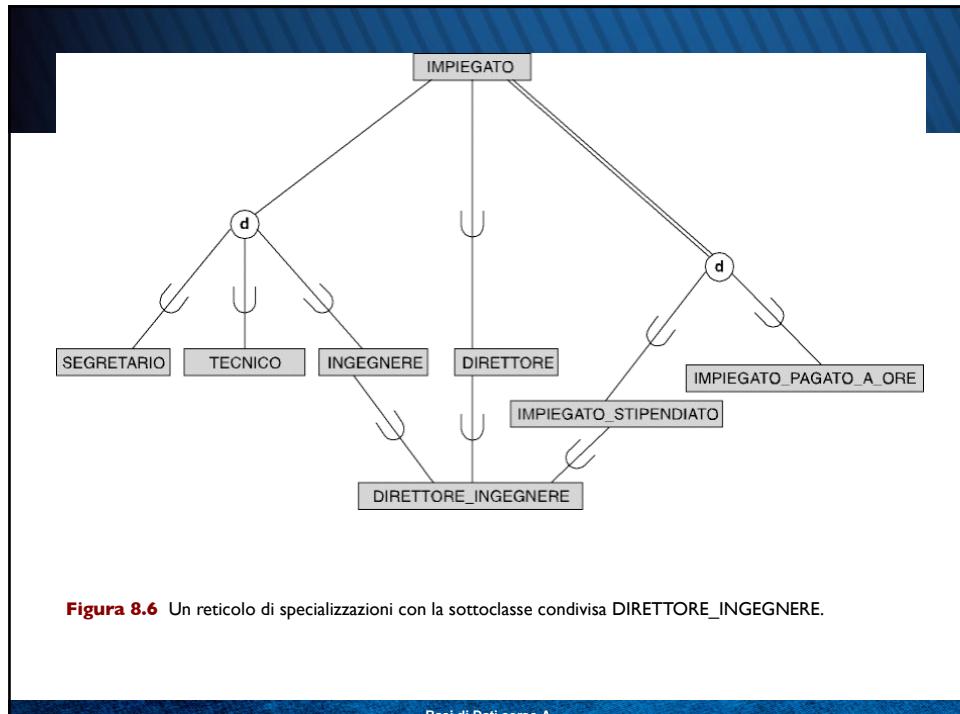
87



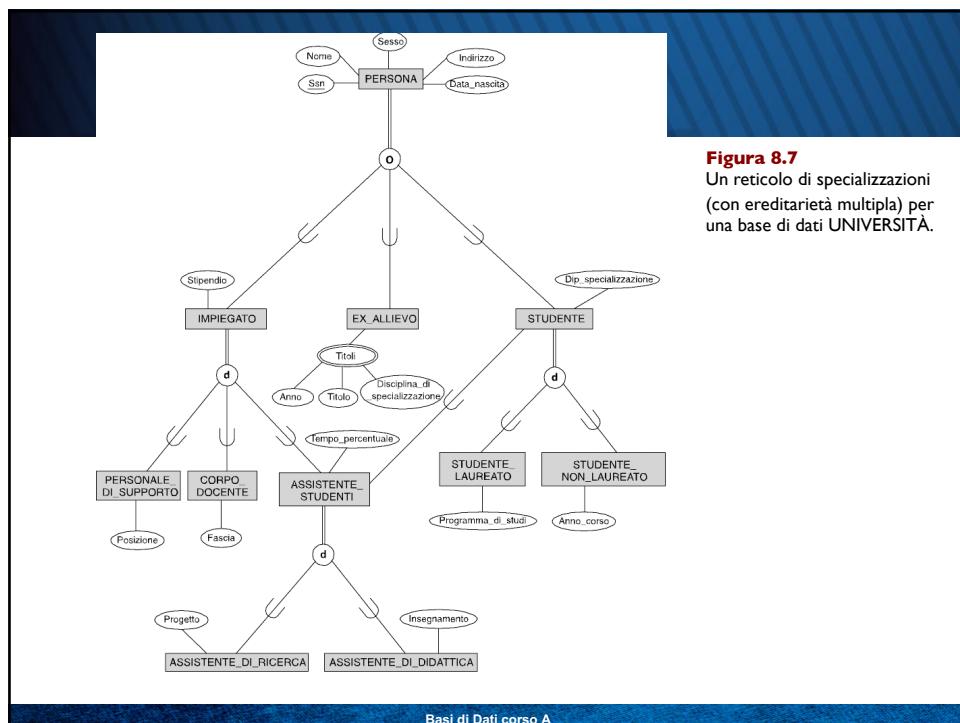
88



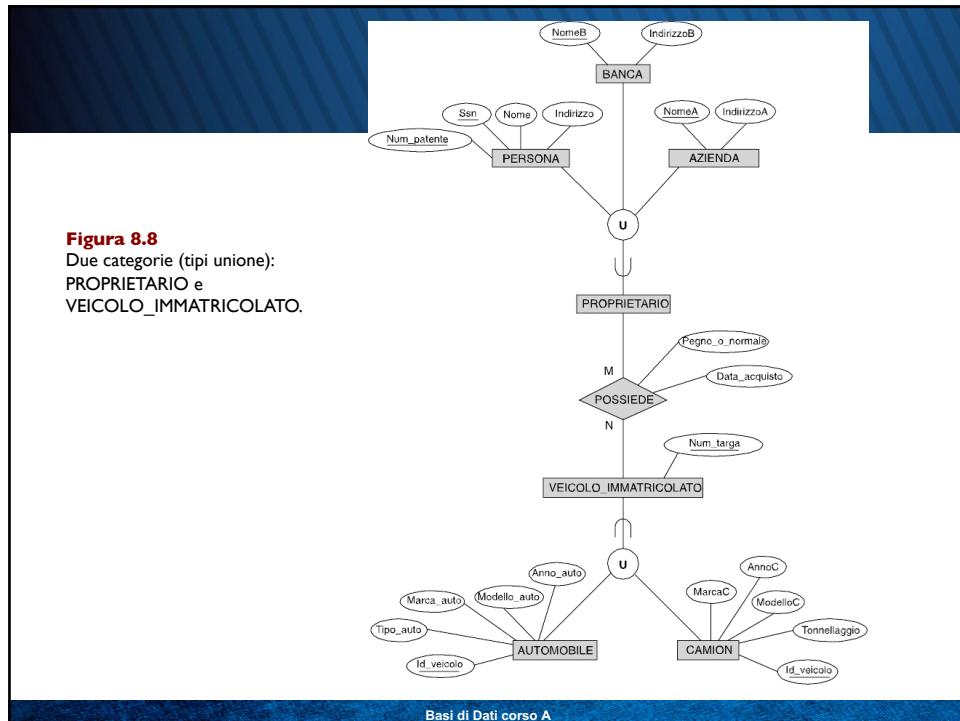
89



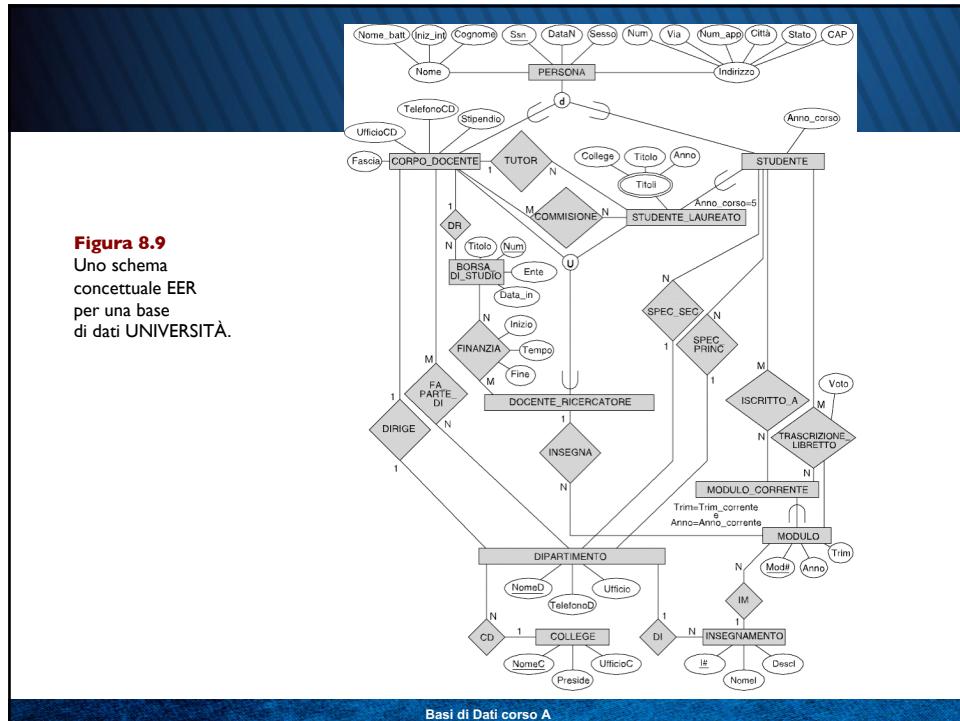
90



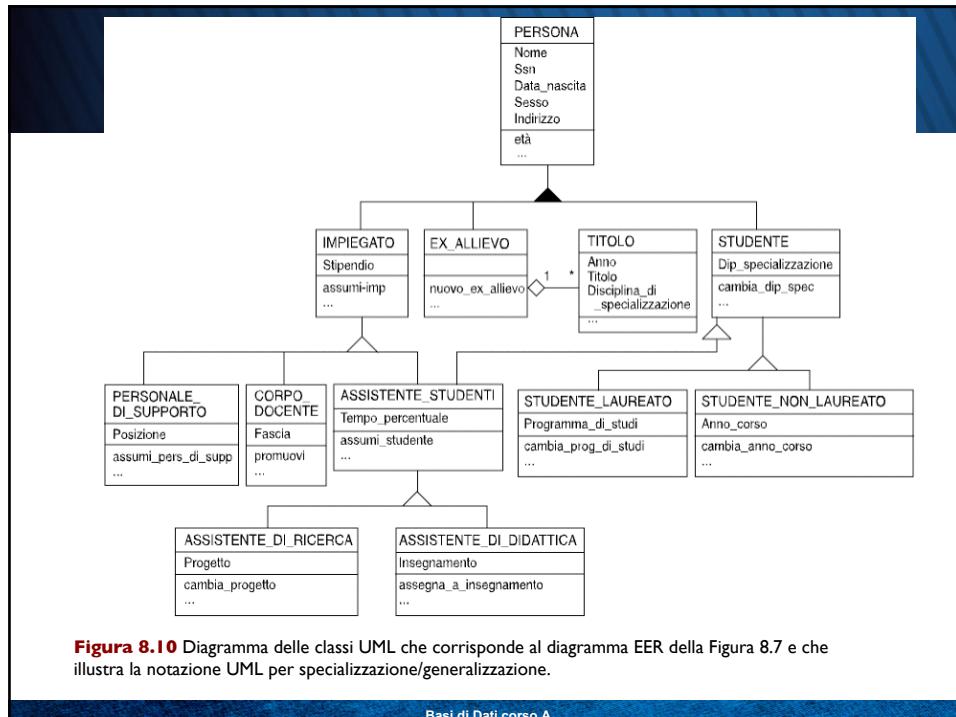
91



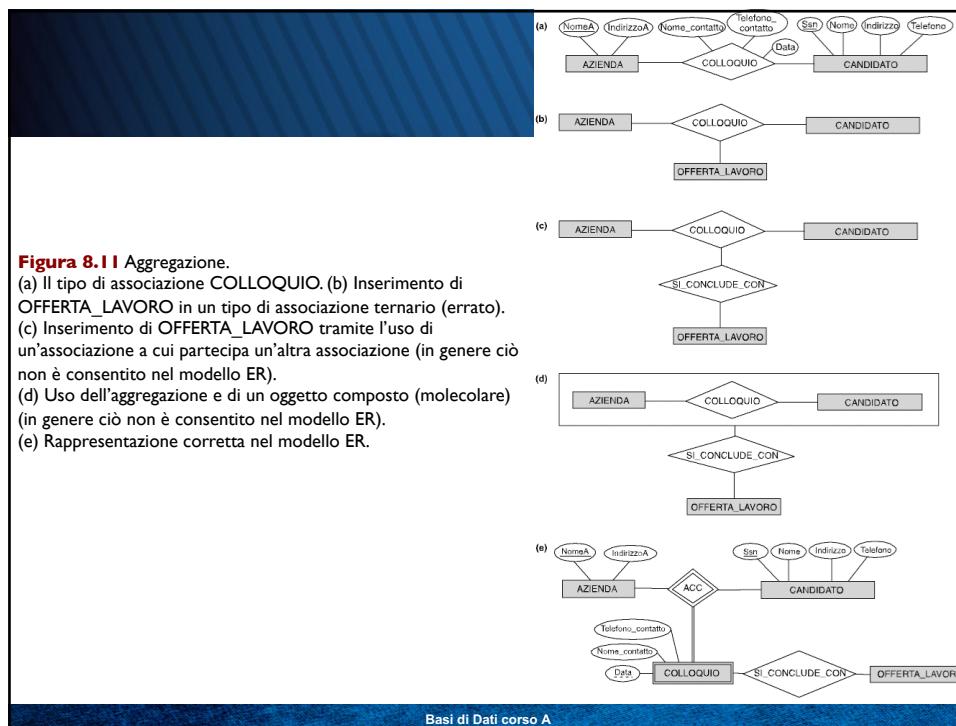
92



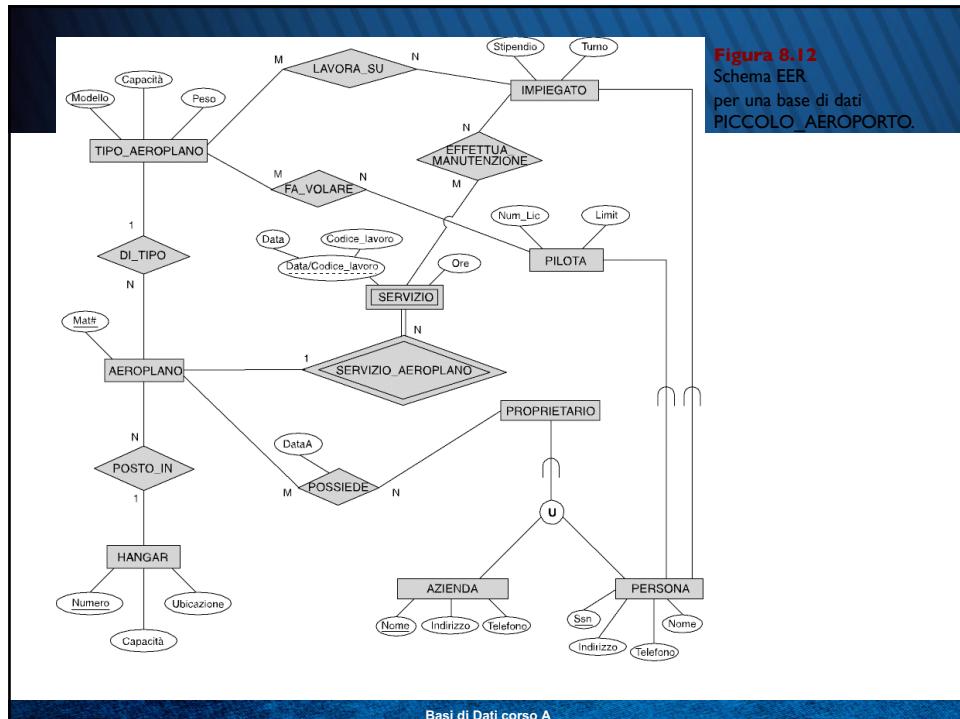
93



94



95



**Figura 8.12**  
Schema EER  
per una base di dati  
PICCOLO\_AEROPORTO.

Basi di Dati corso A

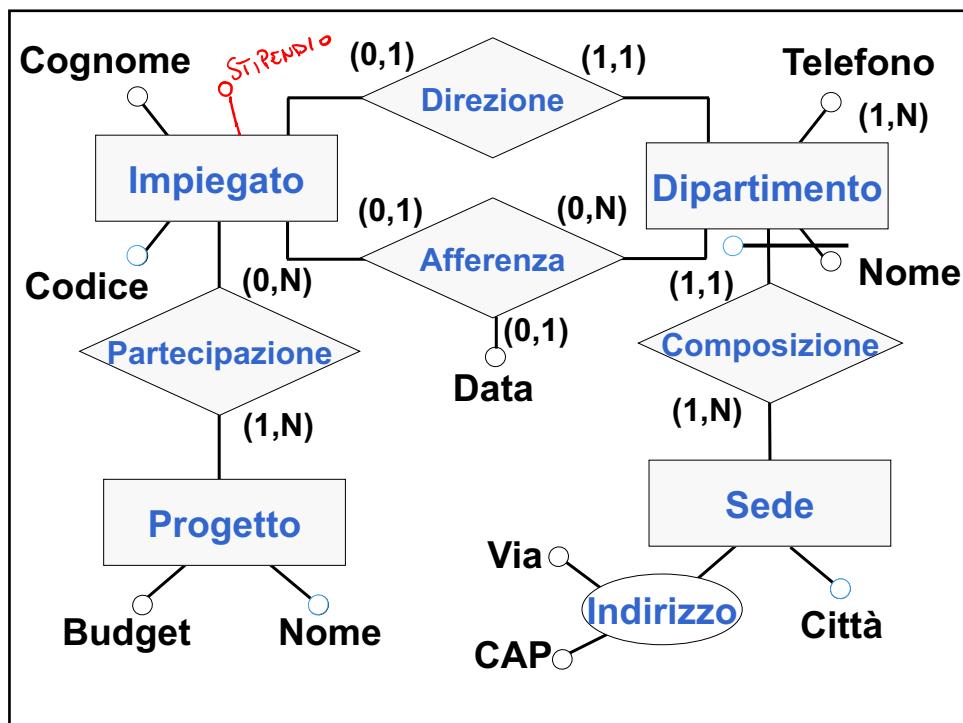
96

## Documentazione associata agli schemi concettuali

- dizionario dei dati
- entità
- relationship
- vincoli non esprimibili

Basi di Dati corso A

97



98

Dizionario dei dati (entità)			
Entità	Descrizione	Attributi	Identificatore
Impiegato	Dipendente dell'azienda	Codice, Cognome, Stipendio	Codice
Progetto	Progetti aziendali	Nome, Budget	Nome
Dipartimento	Struttura aziendale	Nome, Telefono	Nome, Sede
Sede	Sede dell'azienda	Città, Indirizzo	Città

Basi di Dati corso A

99

## Dizionario dei dati (relationship)

Relazioni	Descrizione	Componenti	Attributi
Direzione	Direzione di un dipartimento	Impiegato, Dipartimento	
Afferenza	Afferenza a un dipartimento	Impiegato, Dipartimento	Data
Partecipazione	Partecipazione a un progetto	Impiegato, Progetto	
Composizione	Composizione dell'azienda	Dipartimento, Sede	

Basi di Dati corso A

100

## Vincoli non esprimibili

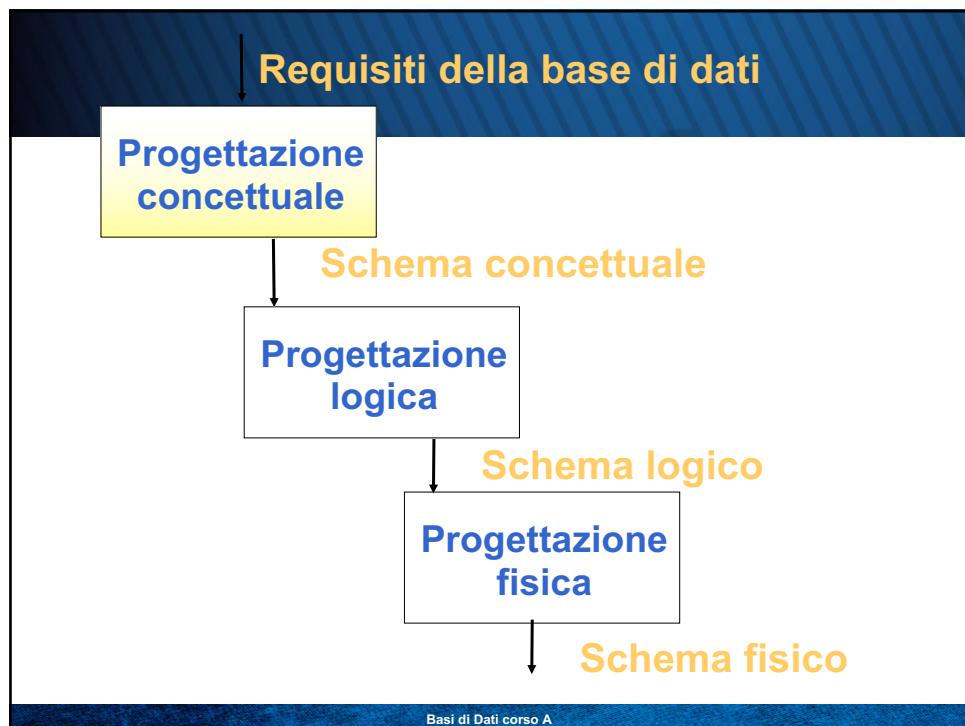
Vincoli di integrità sui dati
(1) Il direttore di un dipartimento deve afferire a tale dipartimento
(2) Un impiegato non deve avere uno stipendio maggiore del direttore del dipartimento al quale afferisce
(3) Un dipartimento con sede a Roma deve essere diretto da un impiegato con più di dieci anni di anzianità
(4) Un impiegato che non afferisce a nessun dipartimento non deve partecipare a nessun un progetto

Basi di Dati corso A

101

# Progettazione concettuale

102



103

## Analisi dei requisiti e progettazione concettuale ("Analisi dei dati")

- ▶ Comprende attività (interconnesse) di
  - ▶ acquisizione dei requisiti
  - ▶ analisi dei requisiti
  - ▶ costruzione dello schema concettuale
  - ▶ costruzione del glossario

Basi di Dati corso A

104

## Requisiti

- ▶ Possibili fonti:
  - ▶ **utenti**, attraverso:
    - ▶ interviste
    - ▶ documentazione
  - ▶ **documentazione esistente**:
    - ▶ normative (leggi, regolamenti di settore)
    - ▶ regolamenti interni, procedure aziendali
    - ▶ realizzazioni preesistenti
  - ▶ **modulistica**

Basi di Dati corso A

105

## Acquisizione e analisi dei requisiti

- ▶ Il reperimento dei requisiti è un'attività difficile e non facilmente standardizzabile (automatizzabile)
- ▶ L'attività di analisi inizia con i primi requisiti raccolti e spesso indirizza verso altre acquisizioni

Basi di Dati corso A

106

## Acquisizione per interviste

- ▶ utenti diversi possono fornire informazioni diverse
- ▶ utenti a livello più alto hanno spesso una visione più ampia ma meno dettagliata
- ▶ le interviste portano spesso ad una acquisizione dei requisiti “per raffinamenti successivi”

Basi di Dati corso A

107

## Interazione con gli utenti

- ▶ effettuare spesso verifiche di comprensione e coerenza
- ▶ verificare anche per mezzo di esempi (generali e relativi a casi limite)
- ▶ richiedere definizioni e classificazioni
- ▶ far evidenziare gli aspetti essenziali rispetto a quelli marginali

Basi di Dati corso A

108

## Requisiti: documentazione descrittiva

Regole generali:

- ▶ scegliere il corretto livello di astrazione
- ▶ standardizzare la struttura delle frasi
- ▶ suddividere le frasi articolate
- ▶ separare le frasi sui dati da quelle sulle funzioni

Basi di Dati corso A

109

## Requisiti: organizzazione di termini e concetti

### Regole generali

- ▶ costruire un glossario dei termini
- ▶ individuare omonimi e sinonimi e unificare i termini
- ▶ rendere esplicito il riferimento fra termini
- ▶ riorganizzare le frasi per concetti

Basi di Dati corso A

110

## Requisiti, un esempio

### Base di dati bibliografica

**Si vogliono organizzare i dati di interesse  
per automatizzare la gestione dei  
riferimenti bibliografici**

Basi di Dati corso A

111

**Base di dati bibliografica**

**Si vogliono organizzare i dati di interesse per automatizzare la gestione dei riferimenti bibliografici, con tutte le informazioni da riportarsi in una bibliografia.**

**Per ogni pubblicazione deve esistere un codice identificante costituito da sette caratteri, indicanti le iniziali degli autori, l'anno di pubblicazione e un carattere aggiuntivo per la discriminazione delle collisioni.**

Basi di Dati corso A

112

**Base di dati bibliografica**

**Si vogliono organizzare i dati di interesse per automatizzare la gestione dei riferimenti bibliografici, con tutte le informazioni da riportarsi in una bibliografia. Le pubblicazioni sono di due tipi, monografie (per le quali interessano editore, data e luogo di pubblicazione) e articoli su rivista (con nome della rivista, volume, numero, pagine e anno di pubblicazione); per entrambi i tipi si debbono ovviamente riportare i nomi degli autori. Per ogni pubblicazione deve esistere un codice identificante .....**

Basi di Dati corso A

113

## Altro esempio

### Società di formazione (1)

Si vuole realizzare una base di dati per una società che eroga corsi di cui vogliamo rappresentare i dati dei partecipanti ai corsi e dei docenti. Per gli studenti (circa 5000), identificati da un codice, si vuole memorizzare il codice fiscale, il cognome, l'età, il sesso, il luogo di nascita, il nome dei loro attuali datori di lavoro, i posti dove hanno lavorato in precedenza insieme al periodo, l'indirizzo e il numero di telefono, i corsi che hanno frequentato (i corsi sono in tutto circa 200) e il giudizio finale.





Basi di Dati corso A

114

### Società di formazione (2)

Rappresentiamo anche i seminari che stanno attualmente frequentando e, per ogni giorno, i luoghi e le ore dove sono tenute le lezioni. I corsi hanno un codice, un titolo e possono avere varie edizioni con date di inizio e fine e numero di partecipanti. Se gli studenti sono liberi professionisti, vogliamo conoscere l'area di interesse e, se lo possiedono, il titolo. Per quelli che lavorano alle dipendenze di altri, vogliamo conoscere invece il loro livello e la posizione ricoperta.

Basi di Dati corso A

115

**Società di formazione (3)**

Per gli insegnanti (circa 300), rappresentiamo il cognome, l'età, il posto dove sono nati, il nome del corso che insegnano, quelli che hanno insegnato nel passato e quelli che possono insegnare. Rappresentiamo anche tutti i loro recapiti telefonici. I docenti possono essere dipendenti interni della società o collaboratori esterni.

(INSEGNANTE) (C,N) O NOME TEL. ESCLUSIVA

Basi di Dati corso A

116

## Altro esempio

**Società di formazione (1)**

Si vuole realizzare una base di dati per una società che eroga corsi, di cui vogliamo rappresentare i dati dei partecipanti ai corsi e dei docenti. Per gli studenti (circa 5000), identificati da un codice, si vuole memorizzare il codice fiscale, il cognome, l'età, il sesso, il luogo di nascita, il nome dei loro attuali datori di lavoro, i posti dove hanno lavorato in precedenza insieme al periodo, l'indirizzo e il numero di telefono, i corsi che hanno frequentato (i corsi sono in tutto circa 200) e il giudizio finale.

Basi di Dati corso A

117

**Società di formazione (2)**

Rappresentiamo anche i seminari che stanno attualmente frequentando e, per ogni giorno, i luoghi e le ore dove sono tenute le lezioni. I corsi hanno un codice, un titolo e possono avere varie edizioni con date di inizio e fine e numero di partecipanti. Se gli studenti sono liberi professionisti, vogliamo conoscere l'area di interesse e, se lo possiedono, il titolo. Per quelli che lavorano alle dipendenze di altri, vogliamo conoscere invece il loro livello e la posizione ricoperta.

Basi di Dati corso A

118

**Società di formazione (3)**

Per gli insegnanti (circa 300), rappresentiamo il cognome, l'età, il posto dove sono nati, il nome del corso che insegnano, quelli che hanno insegnato nel passato e quelli che possono insegnare. Rappresentiamo anche tutti i loro recapiti telefonici. I docenti possono essere dipendenti interni della società o collaboratori esterni.

Basi di Dati corso A

119

## Glossario dei termini

<b>Termine</b>	<b>Descrizione</b>	<b>Sinonimi</b>	<b>Collegamenti</b>
Partecipante	Persona che partecipa ai corsi	Studente	Corso, Società
Docente	Persona che insegna nei corsi. Può essere esterno	Insegnante	Corso
Corso	Corso organizzato dalla società. Può avere più edizioni.	Seminario	Docente Partecipante
Società	Ente presso cui i partecipanti lavorano o hanno lavorato	Posti	Partecipante

Basi di Dati corso A

120

## Strutturazione dei requisiti in gruppi di frasi omogenee

Basi di Dati corso A

121

**Frasi di carattere generale**

**Si vuole realizzare una base di dati per una società che eroga corsi, di cui vogliamo rappresentare i dati dei partecipanti ai corsi e dei docenti.**

Basi di Dati corso A

122

**Frasi relative ai partecipanti**

**Per i partecipanti (circa 5000), identificati da un codice, rappresentiamo il codice fiscale, il cognome, l'età, il sesso, la città di nascita, i nomi dei loro attuali datori di lavoro e di quelli precedenti (insieme alle date di inizio e fine rapporto), le edizioni dei corsi che stanno attualmente frequentando e quelli che hanno frequentato nel passato, con la relativa votazione finale in decimi.**

Basi di Dati corso A

123

**Frasi relative ai datori di lavoro**

**Relativamente ai datori di lavoro presenti e passati dei partecipanti, rappresentiamo il nome, l'indirizzo e il numero di telefono.**

**Frasi relative ai corsi**

**Per i corsi (circa 200), rappresentiamo il titolo e il codice, le varie edizioni con date di inizio e fine e, per ogni edizione, rappresentiamo il numero di partecipanti e il giorno della settimana, le aule e le ore dove sono tenute le lezioni.**

Basi di Dati corso A

124

**Frasi relative a tipi specifici di partecipanti**

**Per i partecipanti che sono liberi professionisti, rappresentiamo l'area di interesse e, se lo possiedono, il titolo professionale. Per i partecipanti che sono dipendenti, rappresentiamo invece il loro livello e la posizione ricoperta.**

Basi di Dati corso A

125

**Frasi relative ai docenti**

**Per i docenti (circa 300), rappresentiamo il cognome, l'età, la città di nascita, tutti i numeri di telefono, il titolo del corso che insegnano, di quelli che hanno insegnato nel passato e di quelli che possono insegnare. I docenti possono essere dipendenti interni della società di formazione o collaboratori esterni.**

Basi di Dati corso A

126

Quale costrutto E-R va utilizzato per rappresentare un concetto presente nelle specifiche?

► Bisogna basarsi sulle definizioni dei costrutti del modello E-R

Basi di Dati corso A

127

- ▶ se ha proprietà significative e descrive oggetti con esistenza autonoma
  - ▶ entità
- ▶ se è semplice e non ha proprietà
  - ▶ attributo
- ▶ se correla due o più concetti
  - ▶ relazione
- ▶ se è caso particolare di un altro
  - ▶ generalizzazione

Basi di Dati corso A

128

## Strategie di progetto

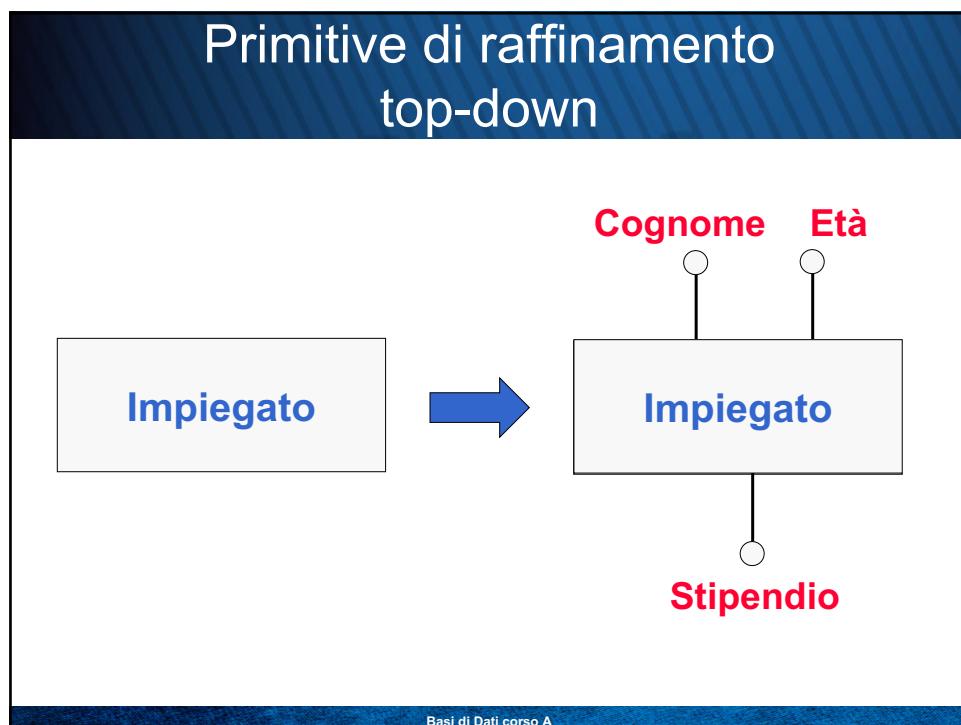
- ▶ top-down
- ▶ bottom-up
- ▶ inside-out

Basi di Dati corso A

129

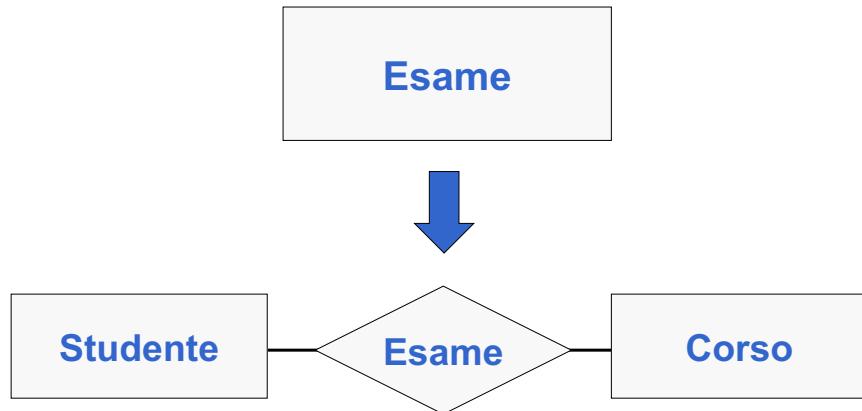


130



131

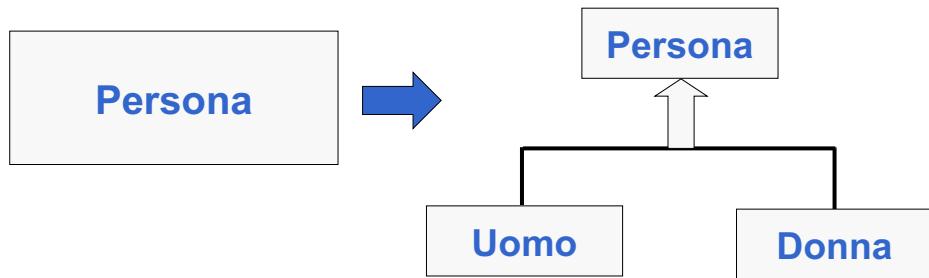
## Primitive di raffinamento top-down



Basi di Dati corso A

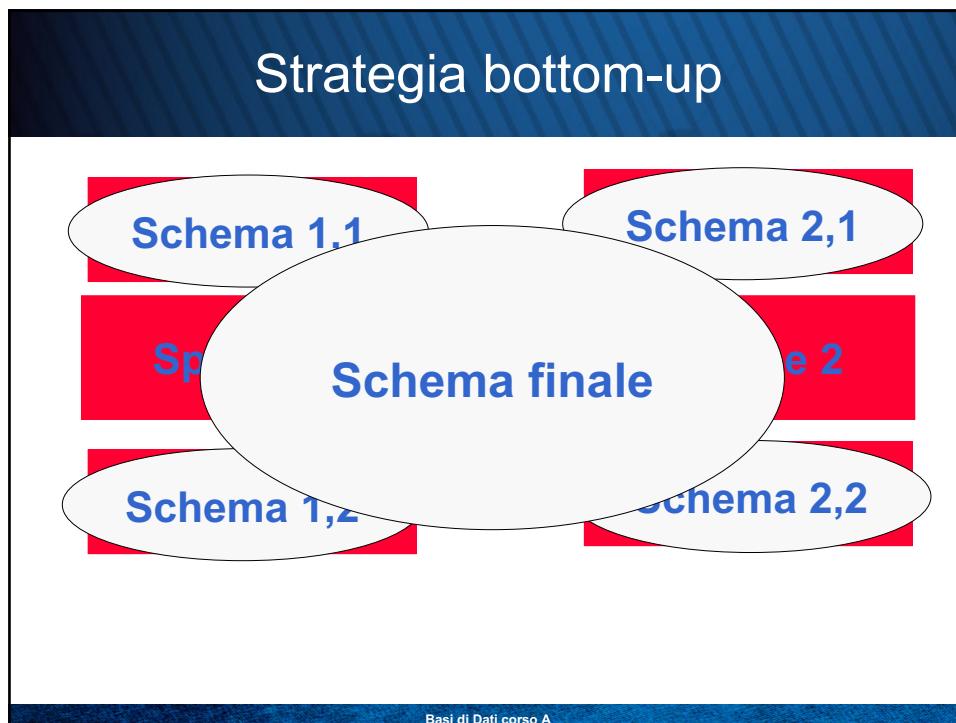
132

## Primitive di raffinamento top-down



Basi di Dati corso A

133

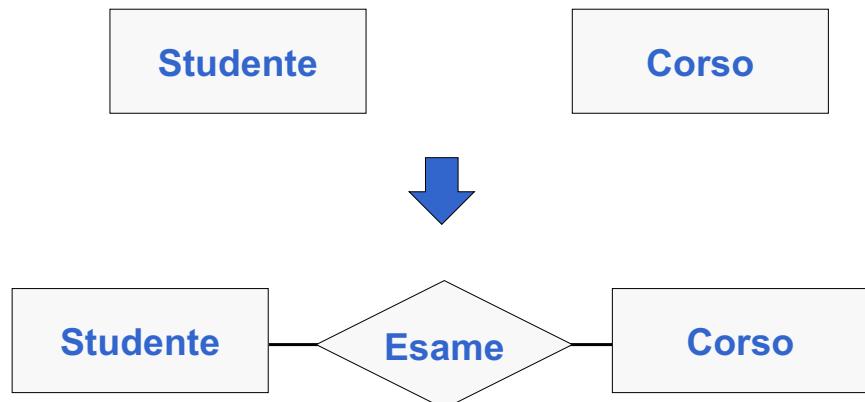


134



135

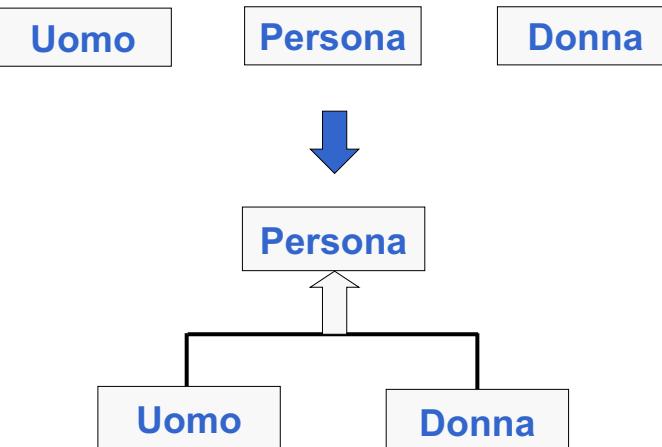
## Primitive di raffinamento Bottom-up



Basi di Dati corso A

136

## Primitive di raffinamento Bottom-up



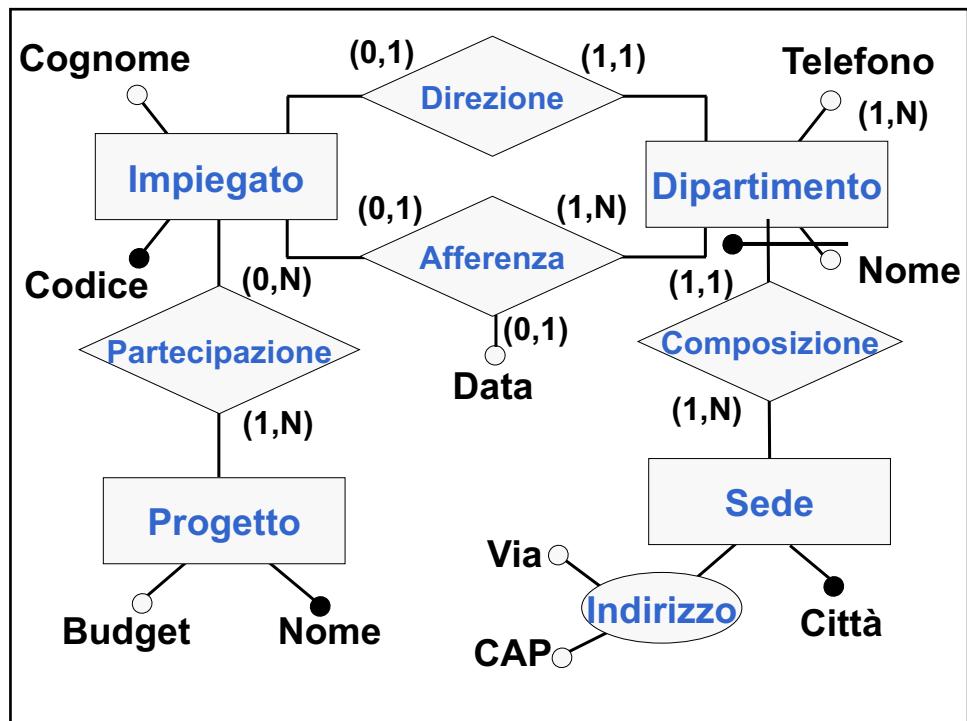
Basi di Dati corso A

137

## Strategia inside-out: un esempio

Basi di Dati corso A

138



139

## In pratica

- ▶ si procede di solito con una strategia ibrida (mista):
  - ▶ si individuano i concetti principali e si realizza uno **schema scheletro**
  - ▶ sulla base di questo si può decomporre
  - ▶ poi si raffina, si espande, si integra

Basi di Dati corso A

140

## Definizione dello schema scheletro

(iniziale)

- ▶ Si individuano i concetti più importanti, ad esempio perché più citati o perché indicati esplicitamente come cruciali e li si organizza in un semplice schema concettuale

Basi di Dati corso A

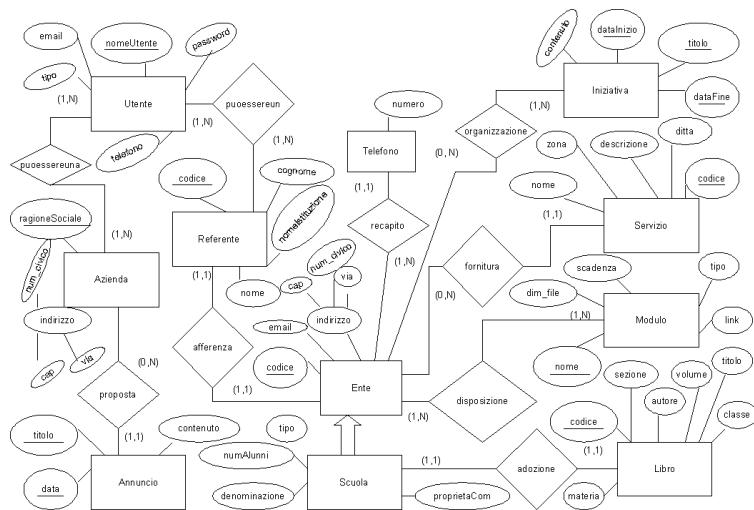
141

## Qualità di uno schema concettuale

- ▶ correttezza
- ▶ completezza
- ▶ leggibilità
- ▶ minimalità

Basi di Dati corso A

142



Basi di Dati corso A

143

## Esempio di progettazione concettuale

- ▶ Società di formazione

Basi di Dati corso A

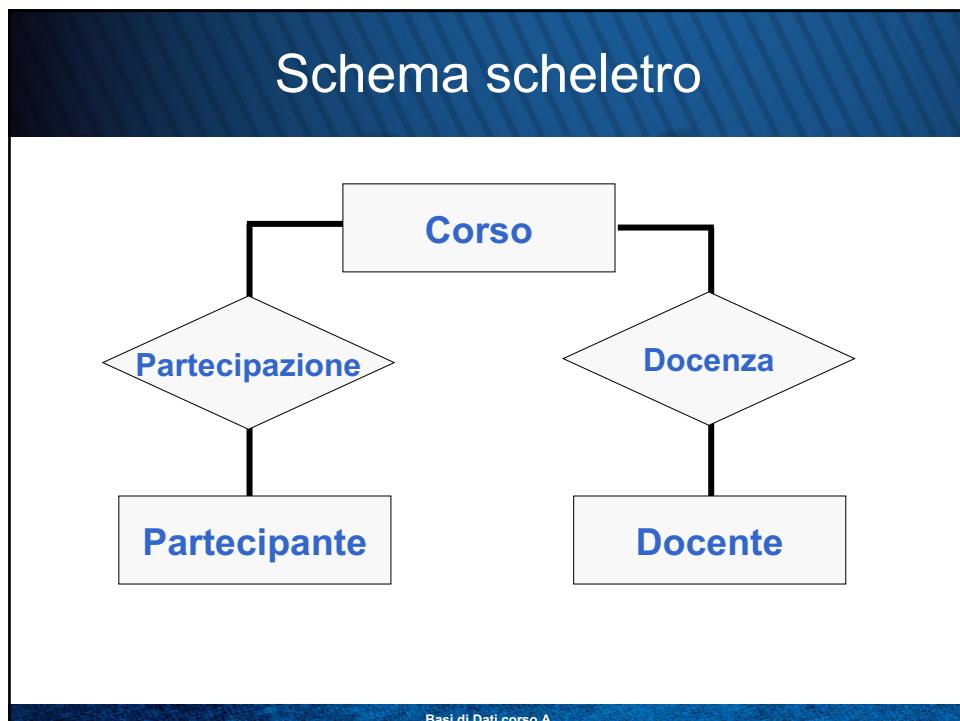
144

### Frasi di carattere generale

Si vuole realizzare una base di dati per  
una società che eroga corsi, di cui  
vogliamo rappresentare i dati dei  
partecipanti ai corsi e dei docenti.

Basi di Dati corso A

145



146

<p><b>Frasi relative ai partecipanti</b></p> <p>Per i partecipanti (circa 5000), identificati da un codice, rappresentiamo il codice fiscale, il cognome, l'età, il sesso, la città di nascita, i nomi dei loro attuali datori di lavoro e di quelli precedenti (insieme alle date di inizio e fine rapporto), le edizioni dei corsi che stanno attualmente frequentando e quelli che hanno frequentato nel passato, con la relativa votazione finale in decimi.</p>
--

Basi di Dati corso A

147

**Frasi relative ai datori di lavoro**

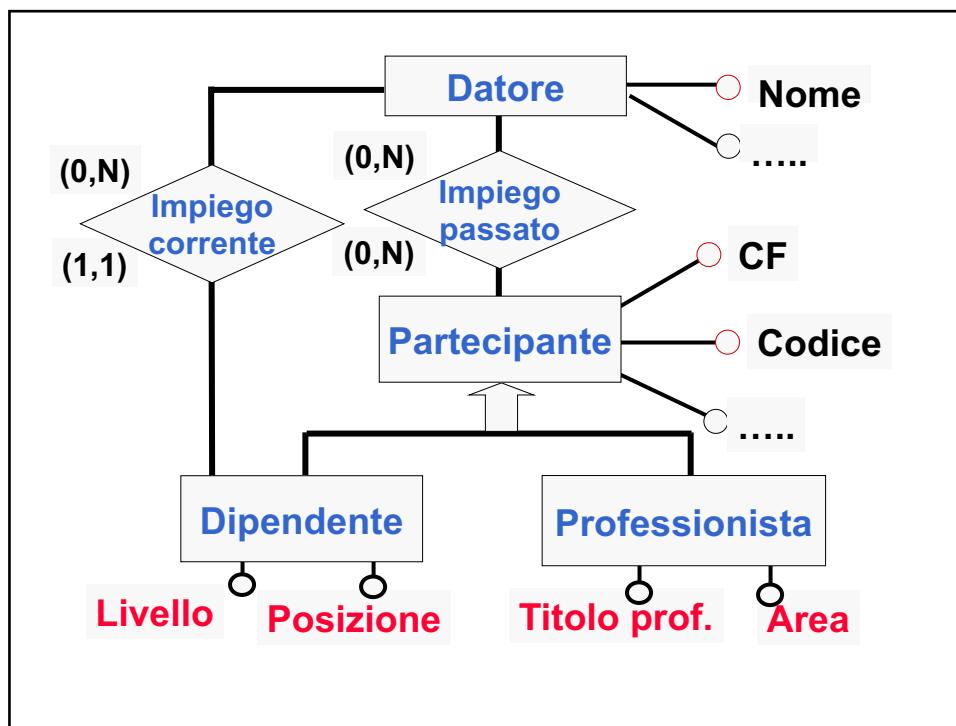
Relativamente ai datori di lavoro presenti e passati dei partecipanti, rappresentiamo il nome, l'indirizzo e il numero di telefono.

**Frasi relative a tipi specifici di partecipanti**

Per i partecipanti che sono **liberi professionisti**, rappresentiamo l'area di interesse e, se lo possiedono, il titolo professionale. Per i partecipanti che sono **dipendenti**, rappresentiamo invece il loro livello e la posizione ricoperta.

Basi di Dati corso A

148



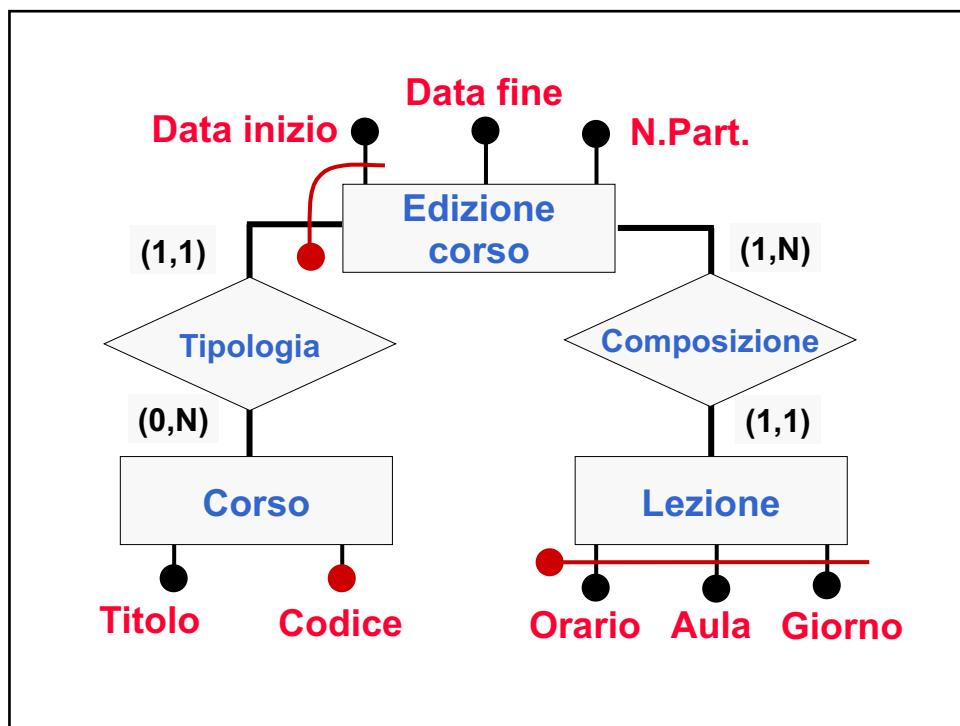
149

**Frasi relative ai corsi**

**Per i corsi (circa 200), rappresentiamo il titolo e il codice, le varie edizioni con date di inizio e fine e, per ogni edizione, rappresentiamo il numero di partecipanti e il giorno della settimana, le aule e le ore dove sono tenute le lezioni.**

Basi di Dati corso A

150



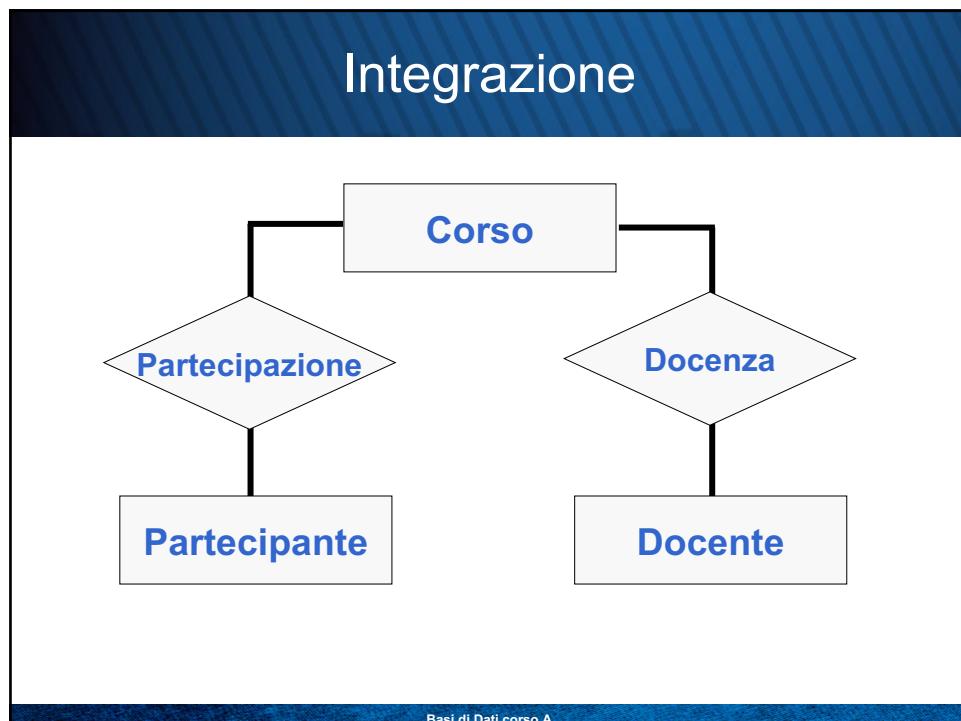
151

**Frasi relative ai docenti**

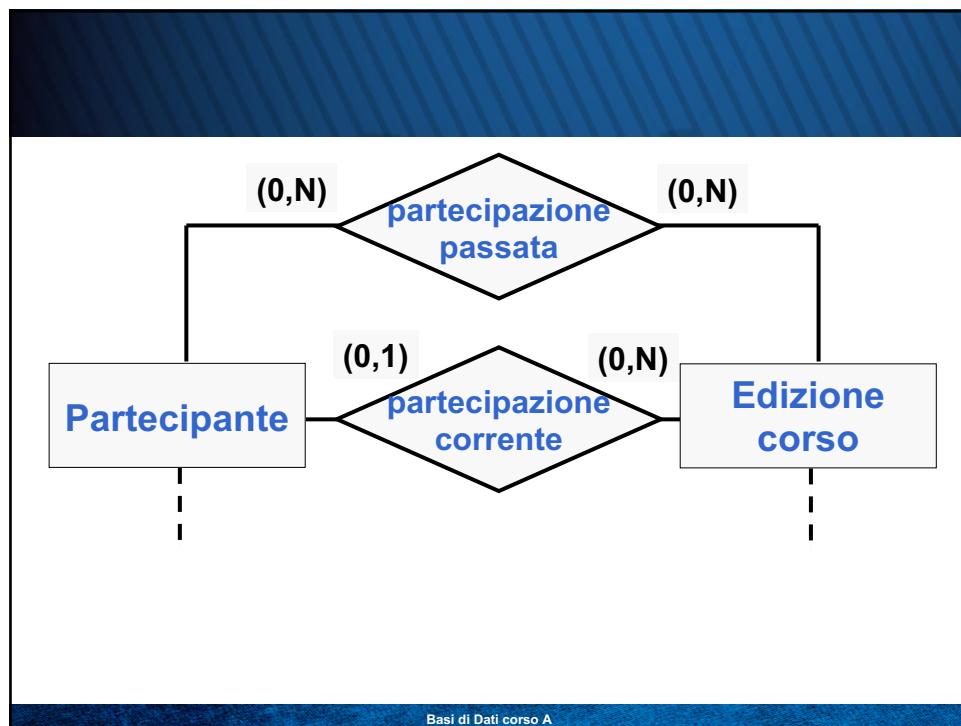
Per i docenti (circa 300), rappresentiamo il cognome, l'età, la città di nascita, tutti i numeri di telefono, il titolo del corso che insegnano, di quelli che hanno insegnato nel passato e di quelli che possono insegnare. I docenti possono essere dipendenti interni della società di formazione o collaboratori esterni.

152

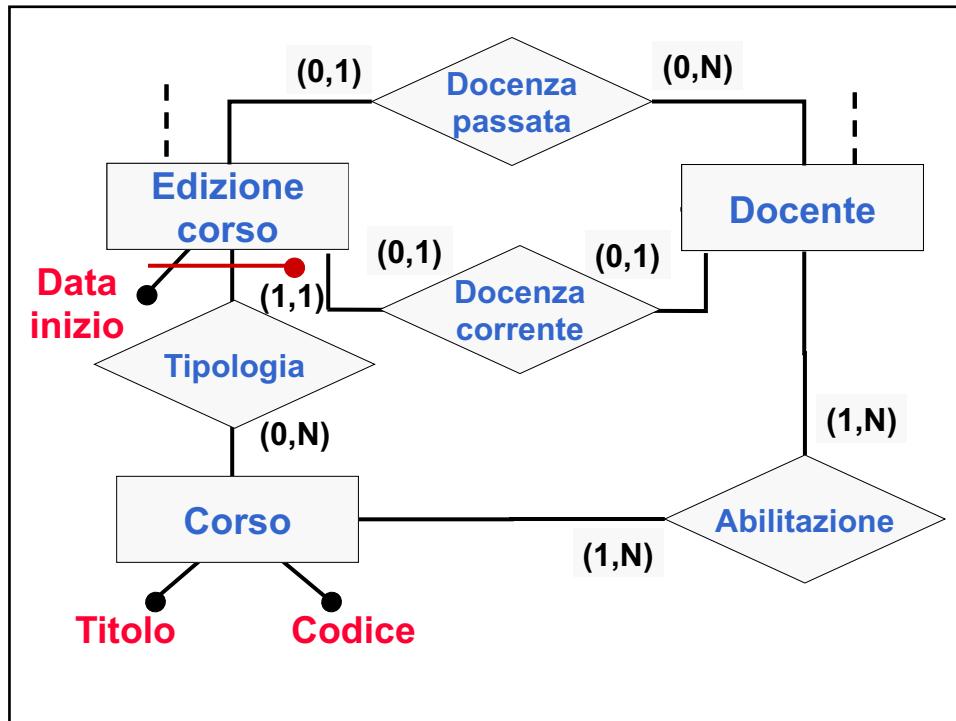
153



154



155



156

## Una metodologia

- ▶ **Analisi dei requisiti**
  - ▶ Analizzare i requisiti ed eliminare le ambiguità
  - ▶ Costruire un glossario dei termini
  - ▶ Raggruppare i requisiti in insiemi omogenei
- ▶ **Passo base**
  - ▶ Definire uno schema scheletro con i concetti più rilevanti
- ▶ **Passo iterativo**  
(da ripetere finché non si è soddisfatti)
  - ▶ Raffinare i concetti presenti sulla base delle loro specifiche
  - ▶ Aggiungere concetti per descrivere specifiche non descritte
- ▶ **Analisi di qualità**  
(ripetuta e distribuita)
  - ▶ Verificare le qualità dello schema e modificarlo

Basi di Dati corso A

157

## Una metodologia con integrazione

- ▶ **Analisi dei requisiti**
- ▶ **Passo base**
- ▶ **Decomposizione**
  - ▶ decomporre i requisiti con riferimento ai concetti nello schema scheletro
- ▶ **Passo iterativo**, per ogni sottoschema
- ▶ **Integrazione**
  - ▶ integrare i vari sottoschemi in uno schema complessivo, facendo riferimento allo schema scheletro
- ▶ **Analisi di qualità**

Basi di Dati corso A

158

## Altra metodologia con integrazione

- ▶ **Analisi dei requisiti**
- ▶ **Decomposizione dei requisiti**
  - ▶ individuazione di settore di interesse e suddivisione dei requisiti (o addirittura acquisizione separata)
- ▶ **Per ciascun settore**
  - ▶ **Passo base**
  - ▶ **Passo iterativo**
- ▶ **Integrazione**
- ▶ **Analisi di qualità**

Basi di Dati corso A

159

## Esempio Testate giornistiche

- ▶ Il candidato progetti la base di dati per conto delle inserzioni gratuite in un insieme di testate giornistiche, tenendo conto delle informazioni riportate nel seguito. Ciascuna delle testate trattate nella base di dati è identificata attraverso il nome (ad esempio Secondamano, ilCercaTrova, ecc..). Inoltre è noto il comitato di redazione della testata stessa, la città in cui questa viene pubblicata, ed eventualmente l'indirizzo della pagina web, se disponibile. Il contenuto dei giornali è organizzato in categorie. Ciascuna categoria è identificata da un nome. Per ogni categoria sono note le sottocategorie in cui questa è strutturata (ad esempio affitti e vendite per la categoria case). Si noti che ogni sottocategoria può a sua volta essere ulteriormente suddivisa (ad esempio, vendita case nell'area nord-ovest di Torino).

Basi di Dati corso A

160

## Esempio Testate giornistiche

- ▶ Il numero finale di livelli in cui è organizzata la gerarchia di categorie non è noto a priori. Inoltre per ogni categoria sono note le testate in cui essa è presente.
- ▶ Ciascuna inserzione può essere un annuncio di un privato o un messaggio pubblicitario ed è identificata da un codice numerico e caratterizzata da un testo descrittivo. Per ogni inserzione è nota la categoria di appartenenza. Per ogni privato sono noti il nome, i recapiti telefonici disponibili, il domicilio ed eventualmente l'indirizzo di e-mail. Per le aziende sono noti il nome dell'azienda, il nome di un referente ed un recapito telefonico.
- ▶ La base di dati tiene traccia di tutte le pubblicazioni fatte per le inserzioni presenti all'interno della base stessa. Ogni annuncio può essere pubblicato su testate diverse e in corrispondenza di più edizioni della stessa testata.

Basi di Dati corso A

161

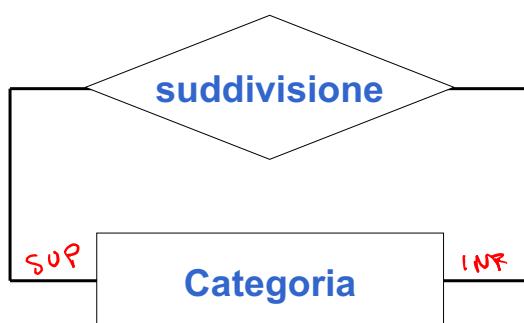
## Esempio Testate giornalistiche

- ▶ Con riferimento alla descrizione precedente:
- ▶ Si progetti la base di dati per l'applicazione descritta, indicando esplicitamente tutte le scelte di progetto effettuate e le motivazioni corrispondenti. Il progetto deve essere corredata da adeguata documentazione (ad esempio, devono essere indicati eventuali vincoli di integrità sui dati).

Basi di Dati corso A

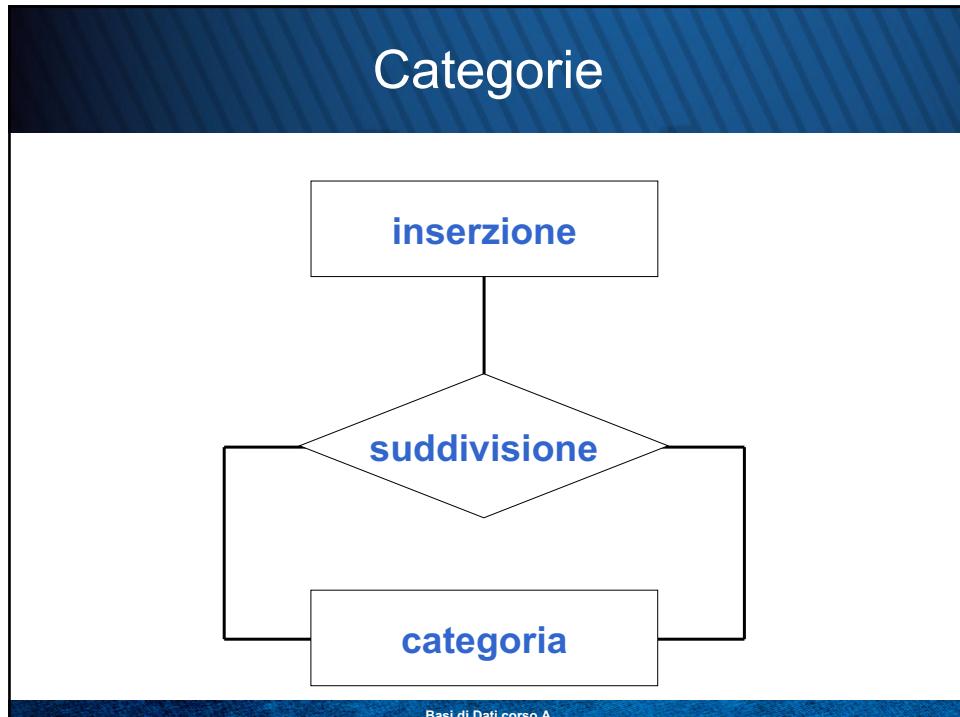
162

## Categorie

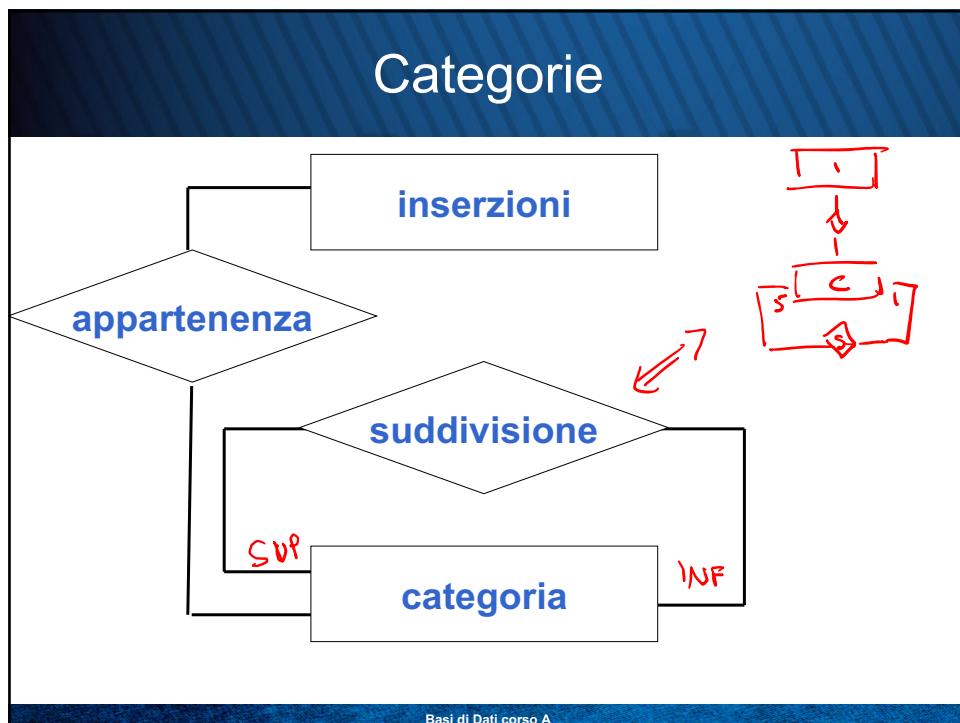


Basi di Dati corso A

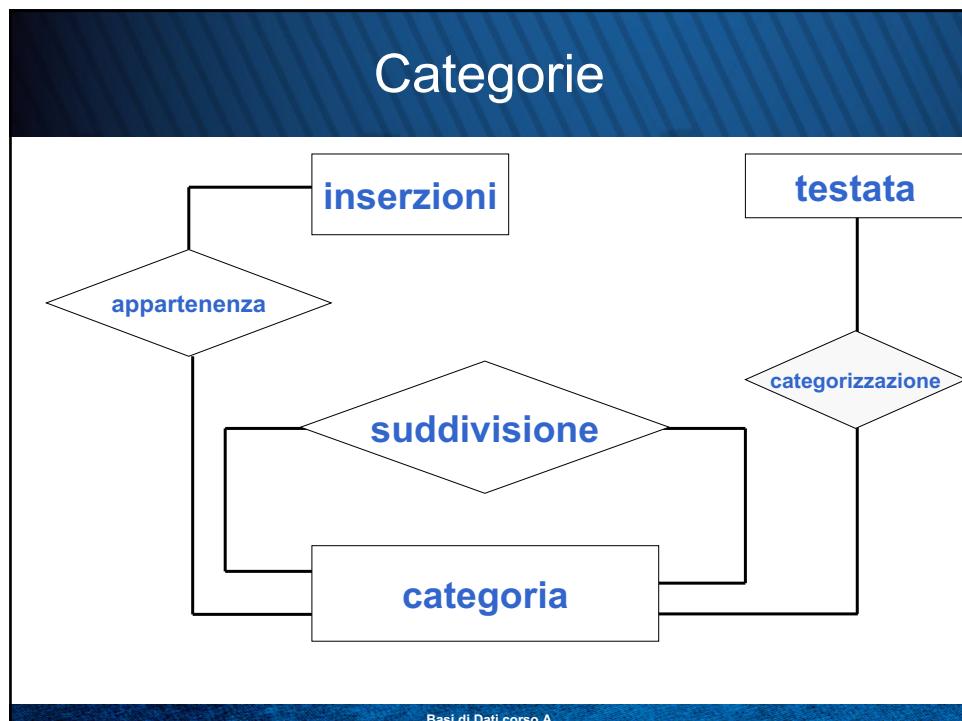
163



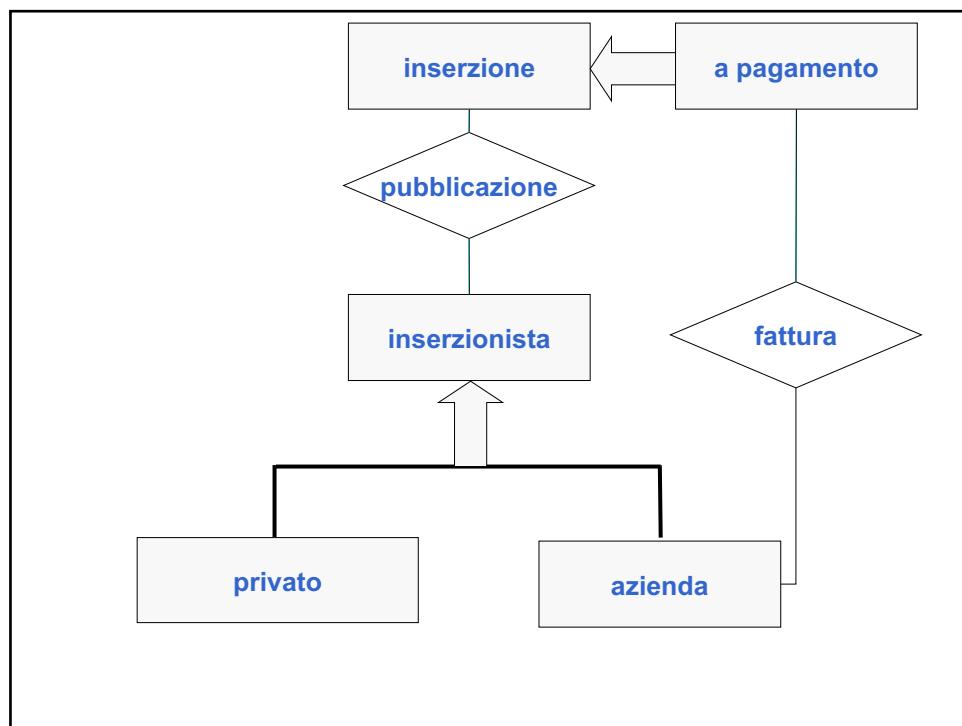
164



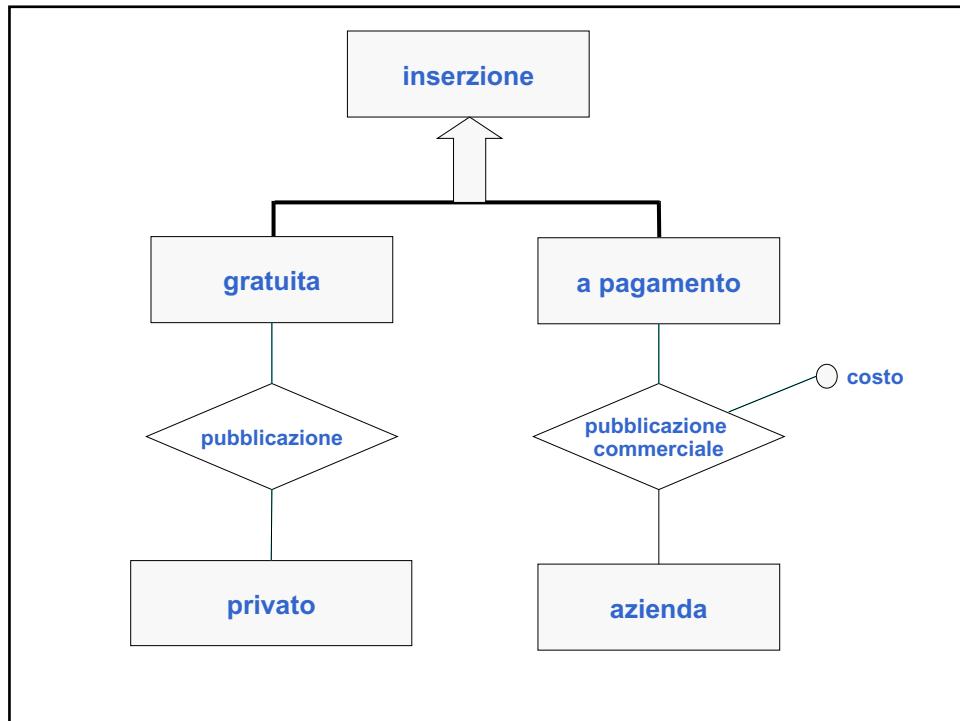
165



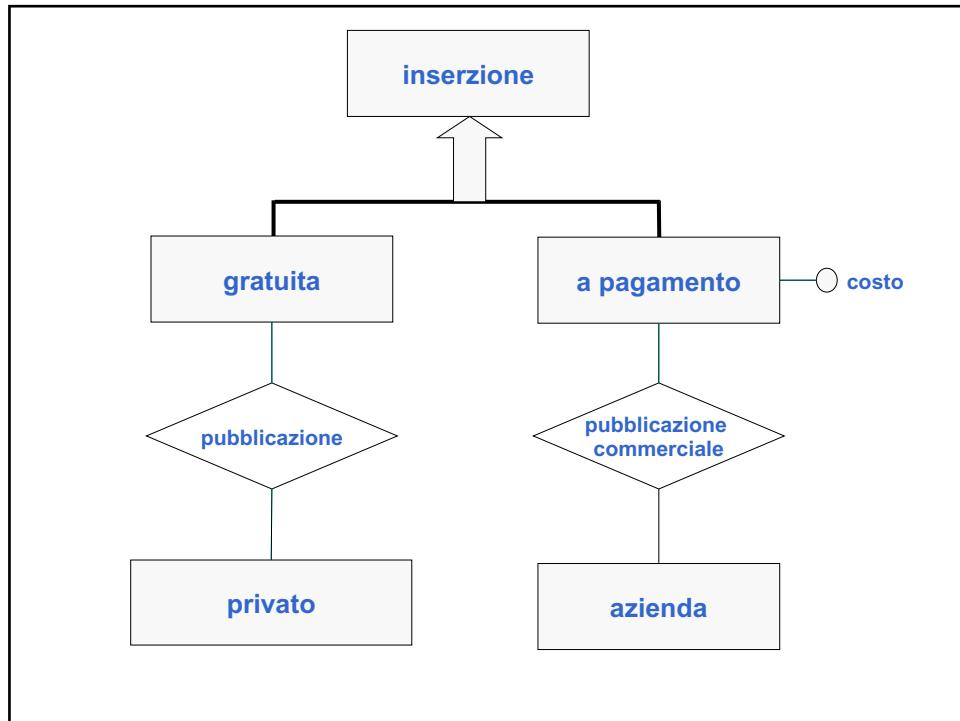
166



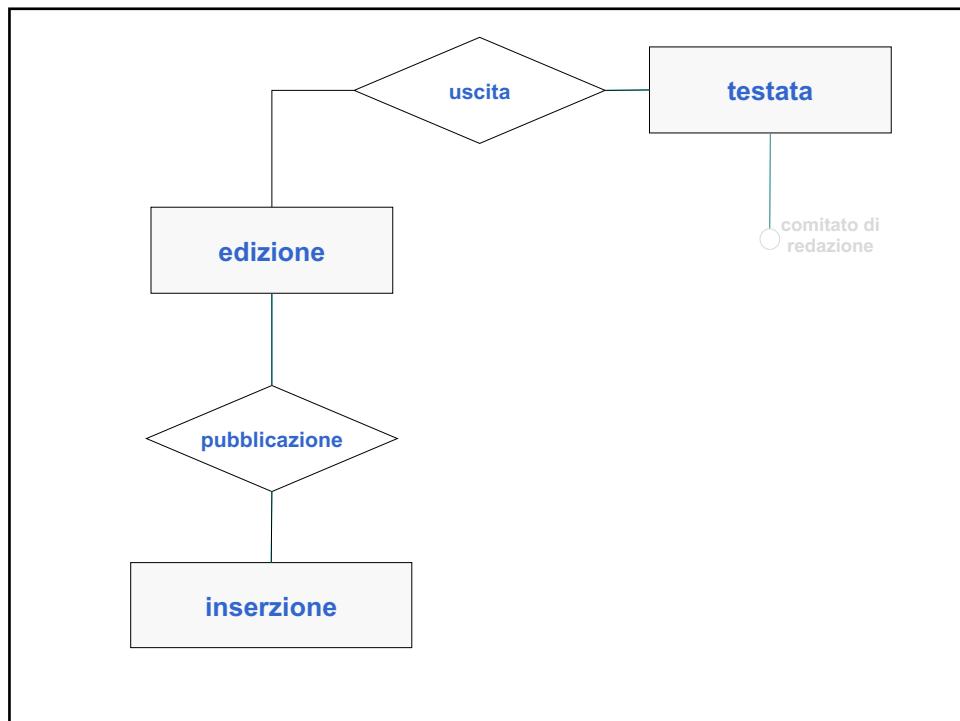
167



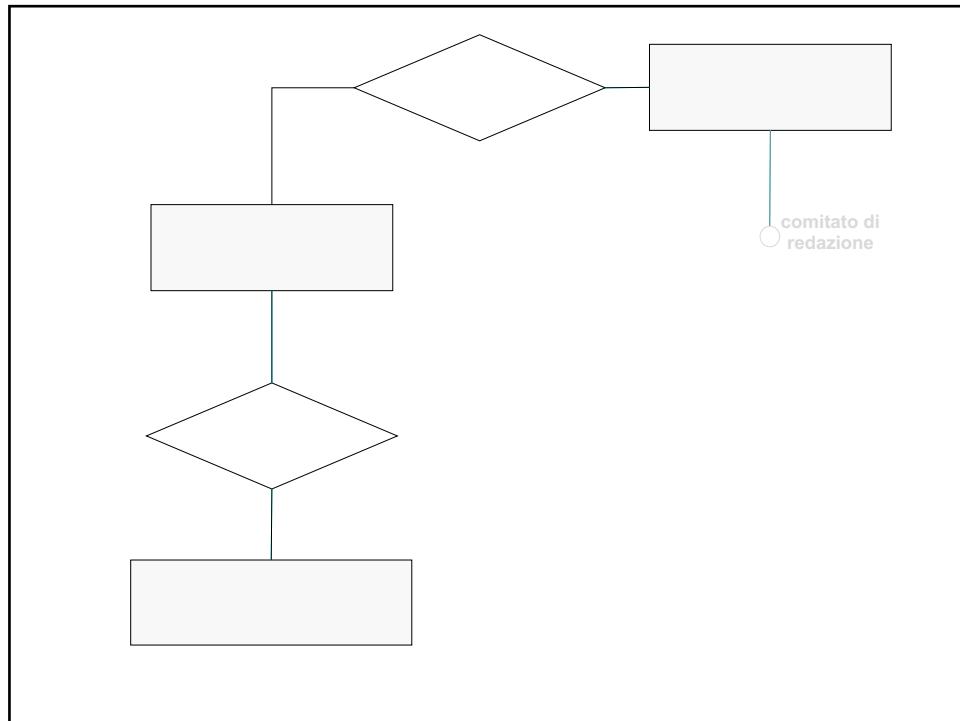
168



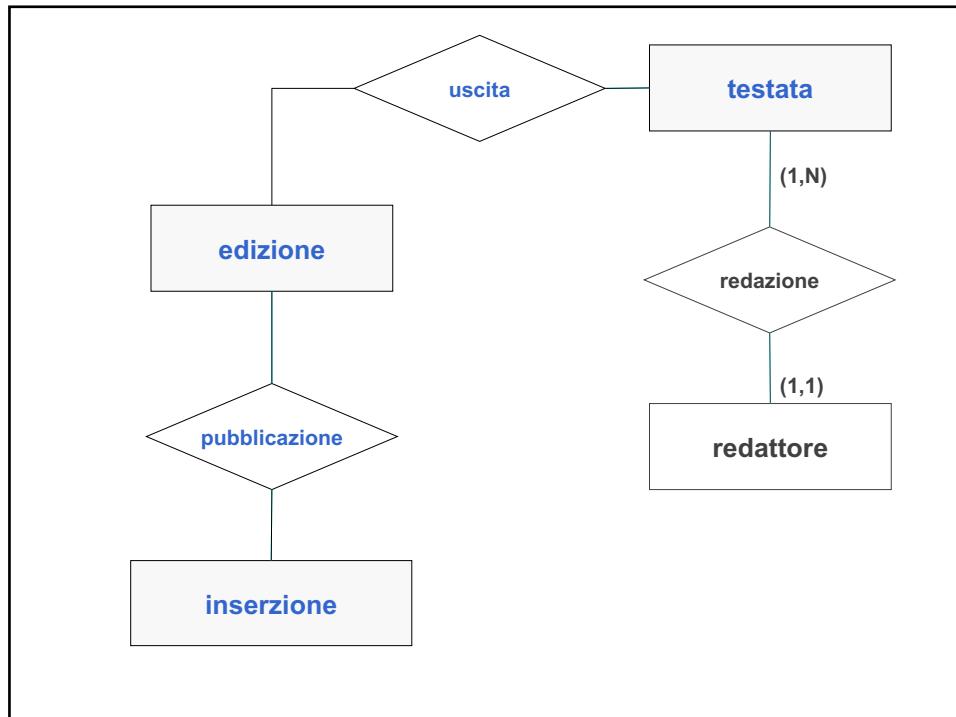
169



170



171



172

## Esempio: Catena di librerie di Harry

Harry vuole memorizzare i propri dati in un database e spera di ricavare grossi vantaggi; in particolare, vuole essere certo che i dati siano sempre aggiornati e precisi. Harry spera di produrre anche diversi prospetti importanti. Spera anche di poter formulare domande sui dati e ottenere risposte in modo agevole e rapidamente. Nel gestire la catena di librerie, Harry raccoglie e organizza informazioni su editori, autori e libri. In libreria sono registrate anche il numero di copie del libro che sono in scorta nelle filiali della catena.

Basi di Dati corso A

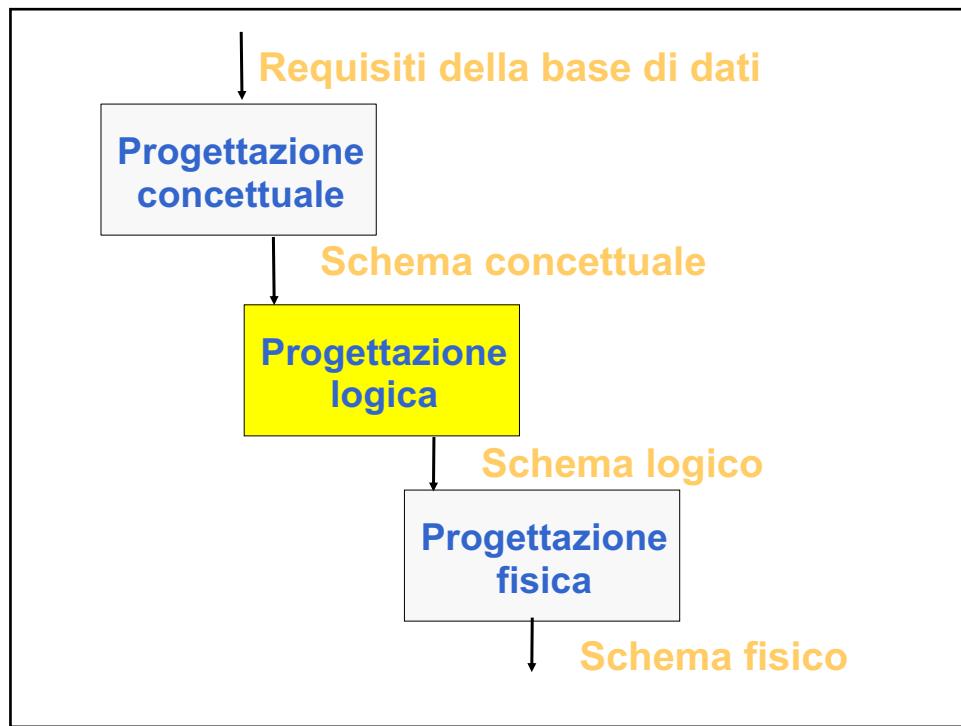
173

# Progettazione logica

Traduzione dal modello ER al modello logico

Basi di Dati corso A

174



175

## Obiettivo della progettazione logica

- ▶ "tradurre" lo schema concettuale in uno schema logico che rappresenti gli stessi dati in maniera corretta ed efficiente

Basi di Dati corso A

176

## Dati di ingresso e uscita

- ▶ **Ingresso:**
  - ▶ schema concettuale
  - ▶ informazioni sul carico applicativo
  - ▶ modello logico
- ▶ **Uscita:**
  - ▶ schema logico
  - ▶ documentazione associata

Basi di Dati corso A

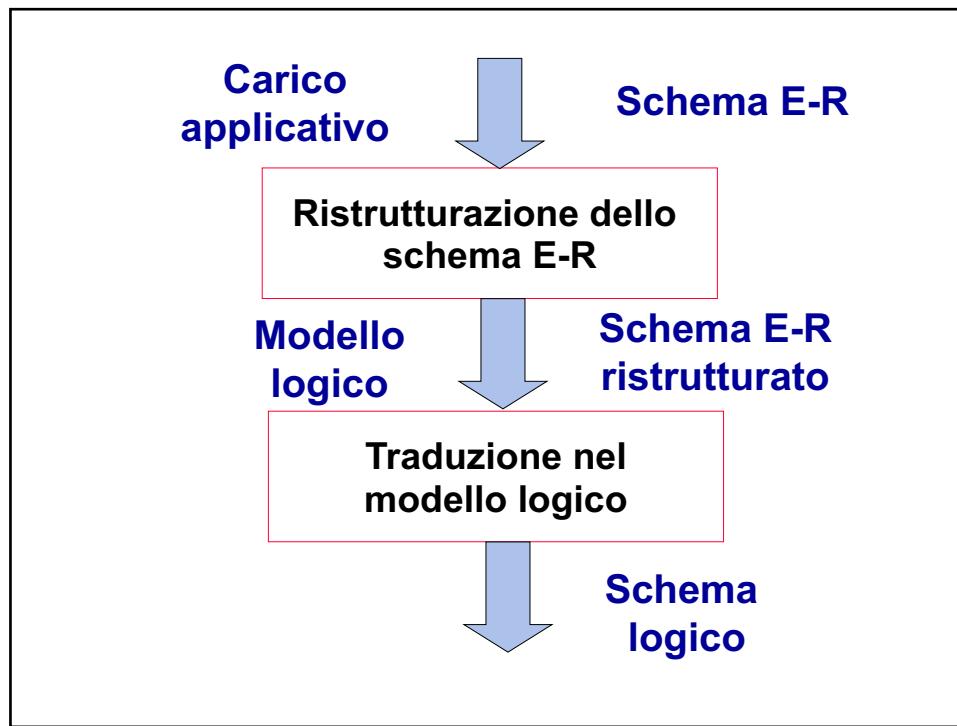
177

## Non si tratta di una pura e semplice traduzione

- ▶ alcuni aspetti non sono direttamente rappresentabili
- ▶ è necessario considerare le prestazioni

Basi di Dati corso A

178



179

## Ristrutturazione schema E-R

► Motivazioni:

- ▶ semplificare la traduzione
- ▶ "ottimizzare" le prestazioni

► Osservazione:

- ▶ uno schema E-R ristrutturato non è (più) uno schema concettuale nel senso stretto del termine

Basi di Dati corso A

180

Per ottimizzare il risultato analizzare le prestazioni a questo livello

- Le prestazioni non sono valutabili con precisione su uno schema concettuale

Basi di Dati corso A

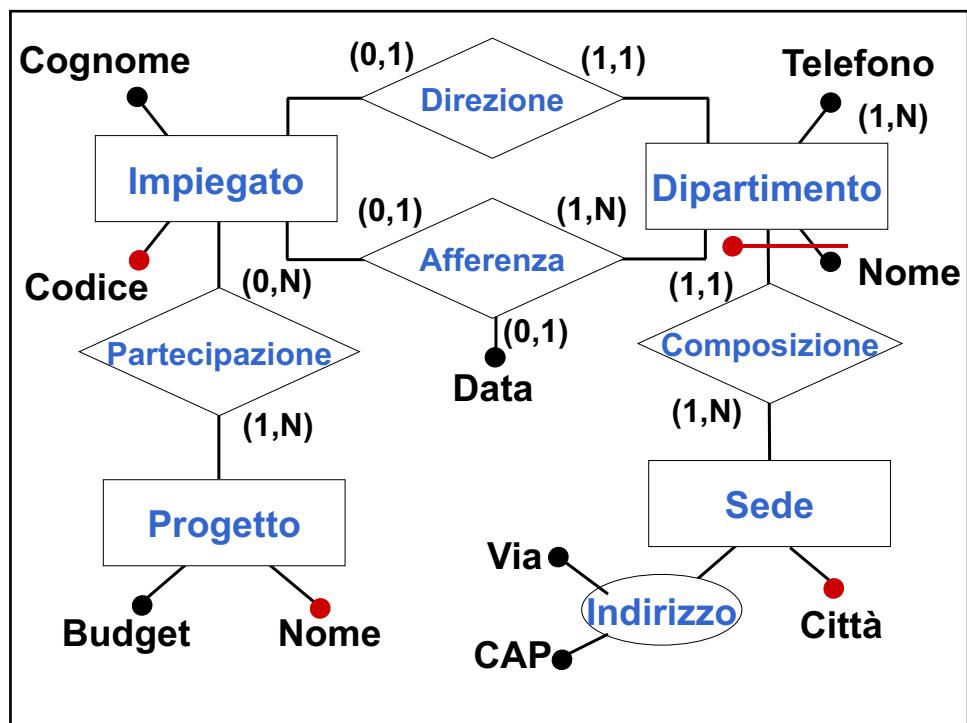
181

## Indicatori di prestazioni

- ▶ spazio: numero di occorrenze previste
- ▶ tempo: numero di occorrenze (di entità e relationship) visitate durante un'operazione

Basi di Dati corso A

182



183

## Tavola dei volumi

Concetto	Tipo	Volume
Sede	E	10
Dipartimento	E	80
Impiegato	E	2000
Progetto	E	500
Composizione	R	80
Afferenza	R	1900
Direzione	R	80
Partecipazione	R	6000

Basi di Dati corso A

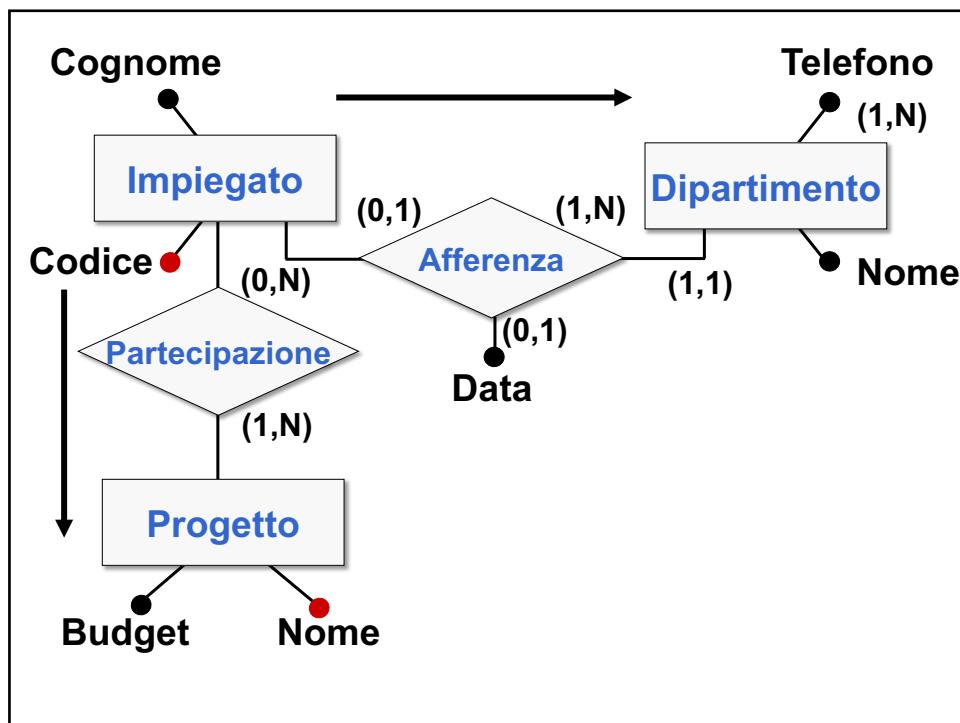
184

## Esempio di valutazione di costo

- ▶ Operazione:
  - ▶ trova tutti i dati di un impiegato, del dipartimento nel quale lavora e dei progetti ai quali partecipa
  - ▶ Si costruisce una **tavola degli accessi** basata su uno **schema di navigazione**

Basi di Dati corso A

185



186

Tavola degli accessi			
Concetto	Costrutto	Accessi	Tipo
Impiegato	Entità	1	L
Afferenza	Relazione	1	L
Dipartimento	Entità	1	L
Partecipazione	Relazione	3	L
Progetto	Entità	3	L

Basi di Dati corso A

187

## Attività della ristrutturazione

- ▶ Analisi delle ridondanze
- ▶ Eliminazione delle generalizzazioni
- ▶ Partizionamento/accorpamento di entità e relationship
- ▶ Scelta degli identificatori primari

Basi di Dati corso A

188

## Analisi delle ridondanze

- ▶ Una ridondanza in uno schema E-R è una informazione significativa ma derivabile da altre
- ▶ in questa fase si decide se eliminare le ridondanze eventualmente presenti o mantenerle

Basi di Dati corso A

189

## Ridondanze

### ► Vantaggi

- semplificazione delle interrogazioni

### ► Svantaggi

- appesantimento degli aggiornamenti
- maggiore occupazione di spazio

Basi di Dati corso A

190

## Forme di ridondanza in uno schema E-R

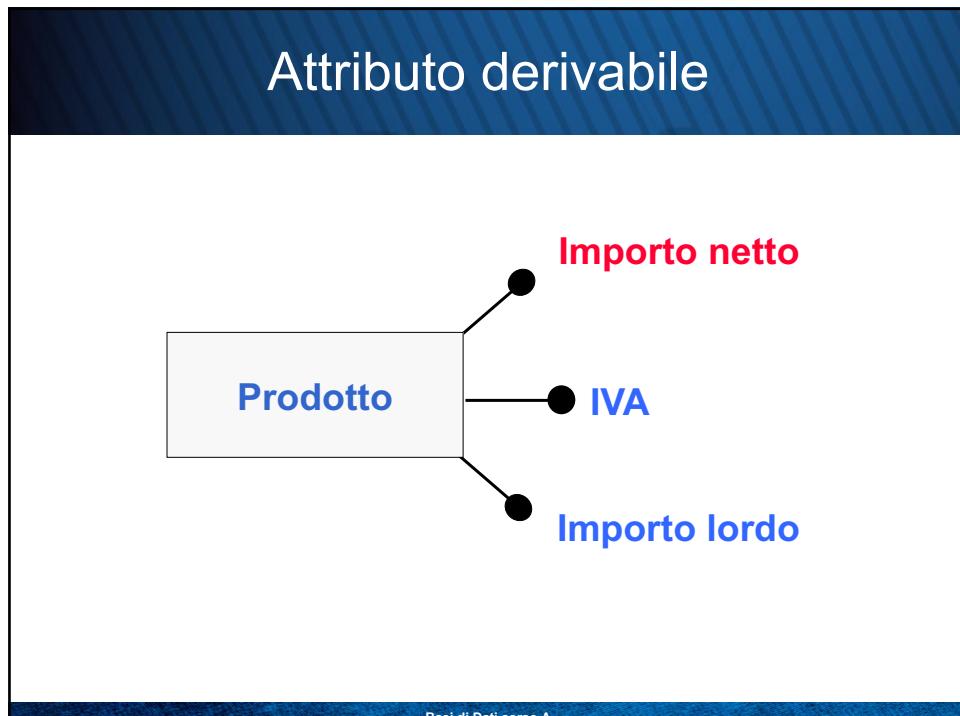
### ► attributi derivabili:

- da altri attributi della stessa entità (o relazione)
- da attributi di altre entità (o relazioni)

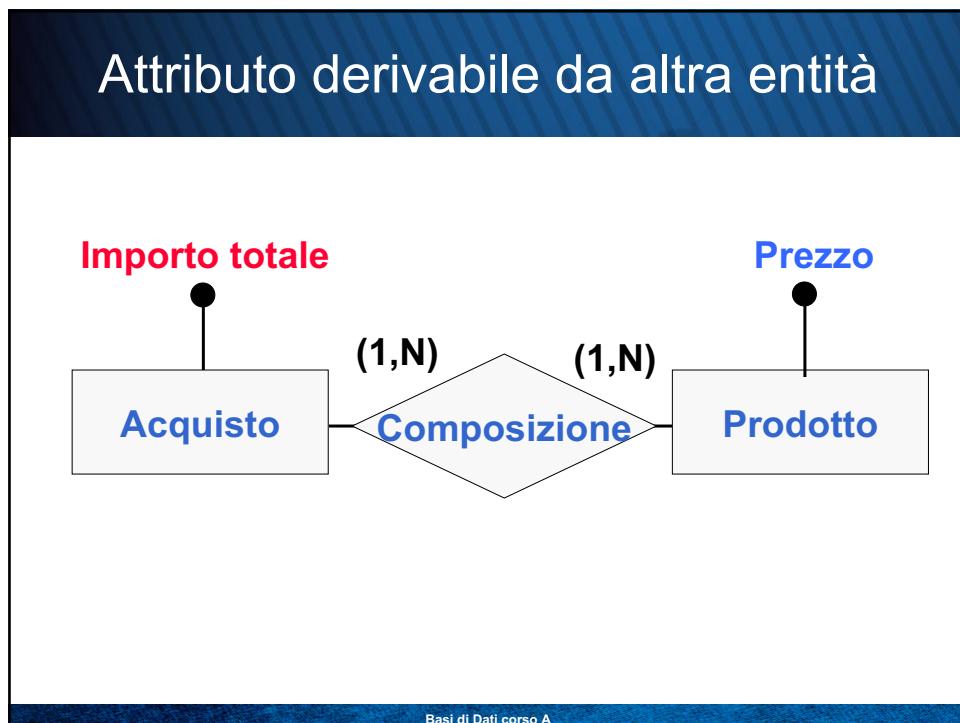
### ► relazioni derivabili dalla composizione di altre relazioni in presenza di cicli

Basi di Dati corso A

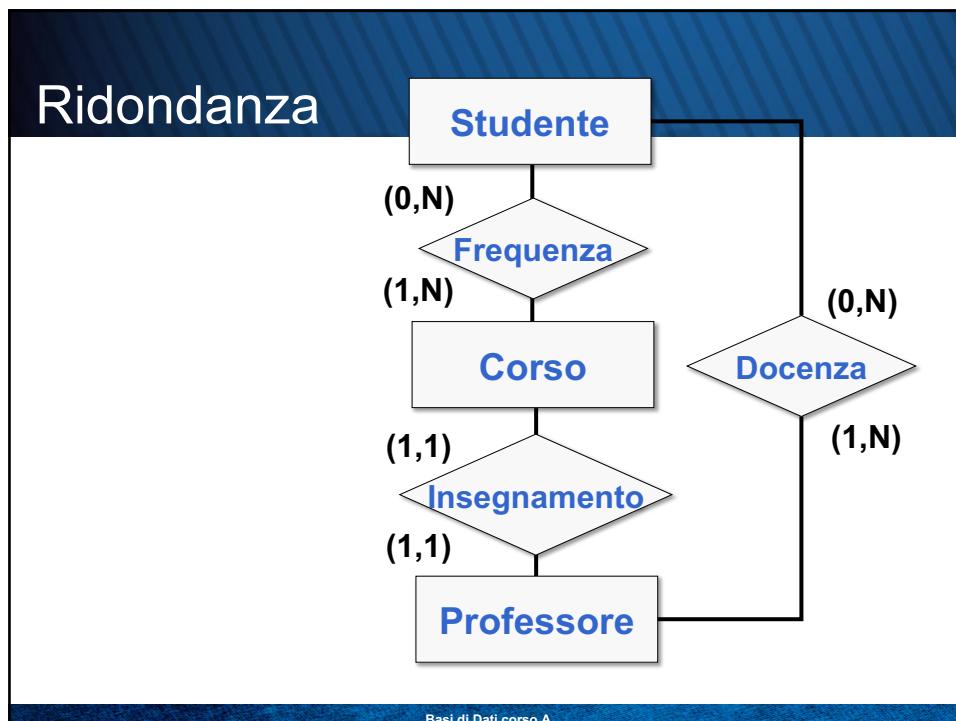
191



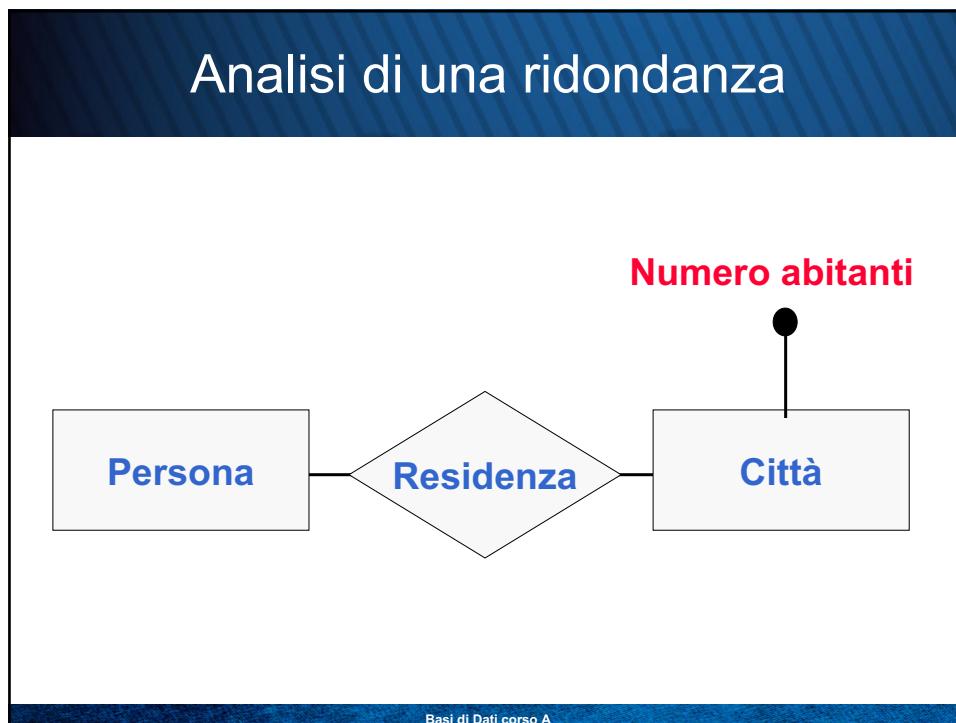
192



193



194



195

<b>Concetto</b>	<b>Tipo</b>	<b>Volume</b>
Città	E	200
Persona	E	1.000.000
Residenza	R	1.000.000

► **Operazione 1:** memorizza una nuova persona con la relativa città di residenza (500 volte al giorno)

► **Operazione 2:** stampa tutti i dati di una città (incluso il numero di abitanti) (2 volte al giorno)

Basi di Dati corso A

196

## Presenza di ridondanza

**Operazione 1**

<b>Concetto</b>	<b>Costrutto</b>	<b>Accessi</b>	<b>Tipo</b>
Persona	Entità	1	S
Residenza	Relazione	1	S
Città	Entità	1	L
Città	Entità	1	S

**Operazione 2**

<b>Concetto</b>	<b>Costrutto</b>	<b>Accessi</b>	<b>Tipo</b>
Città	Entità	1	L

Basi di Dati corso A

197

## Assenza di ridondanza

### Operazione 1

Concetto	Costrutto	Accessi	Tipo
Persona	Entità	1	S
Residenza	Relazione	1	S

### Operazione 2

Concetto	Costrutto	Accessi	Tipo
Città	Entità	1	L
Residenza	Relazione	5.000	L

Basi di Dati corso A

198

## Presenza di ridondanza

### ► Costi

- Operazione 1: 1.500 accessi in scrittura e 500 accessi in lettura al giorno
- Operazione 2: trascurabile, perché = 2
- Contiamo doppi gli accessi in scrittura

Totale di 3.502 accessi al giorno

Basi di Dati corso A

199

## Assenza di ridondanza



### ► Costi

- ▶ Operazione 1: 1.000 accessi in scrittura
- ▶ Operazione 2: 10.002 accessi in lettura al giorno

### ► Contiamo doppi gli accessi in scrittura

**Totale di 12.002 accessi al giorno**

Basi di Dati corso A

200

## Attività della ristrutturazione

- ▶ Analisi delle ridondanze
- ▶ **Eliminazione delle generalizzazioni**
- ▶ Partizionamento/accorpamento di entità e relazioni
- ▶ Scelta degli identificatori primari

Basi di Dati corso A

201

## Eliminazione delle gerarchie

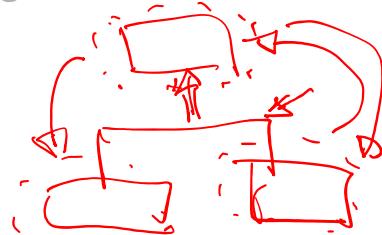
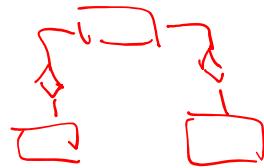
- ▶ il modello relazionale non può rappresentare direttamente le generalizzazioni
- ▶ entità e relazioni sono invece direttamente rappresentabili
- ▶ si eliminano perciò le gerarchie, sostituendole con entità e relazioni

Basi di Dati corso A

202

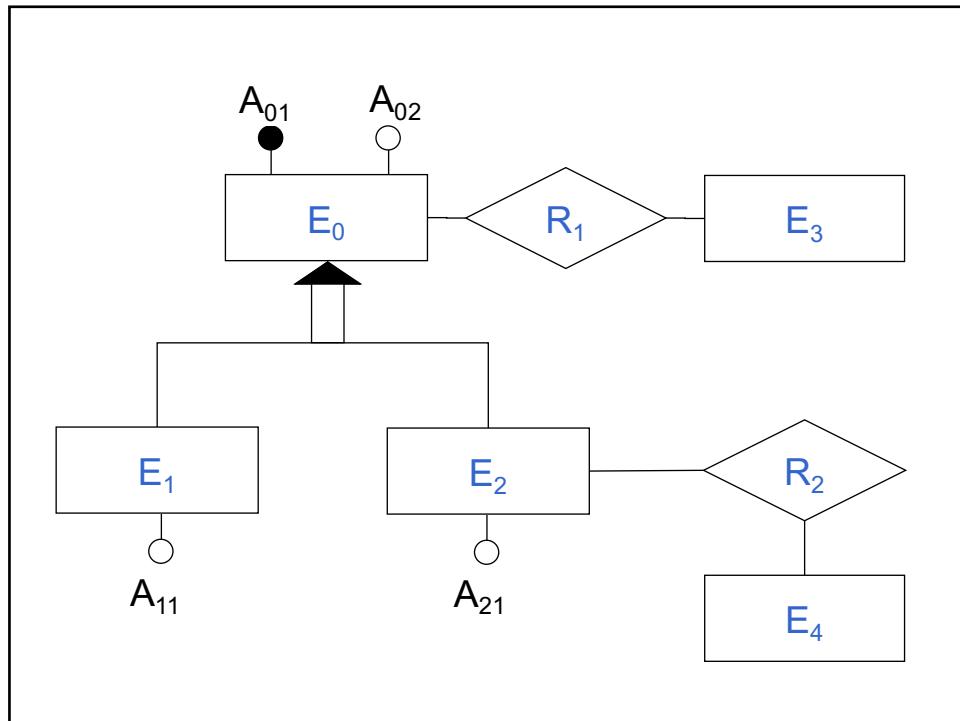
## Tre possibilità

- ▶ accorpamento delle figlie della generalizzazione nel genitore
- ▶ accorpamento del genitore della generalizzazione nelle figlie
- ▶ sostituzione della generalizzazione con relazioni

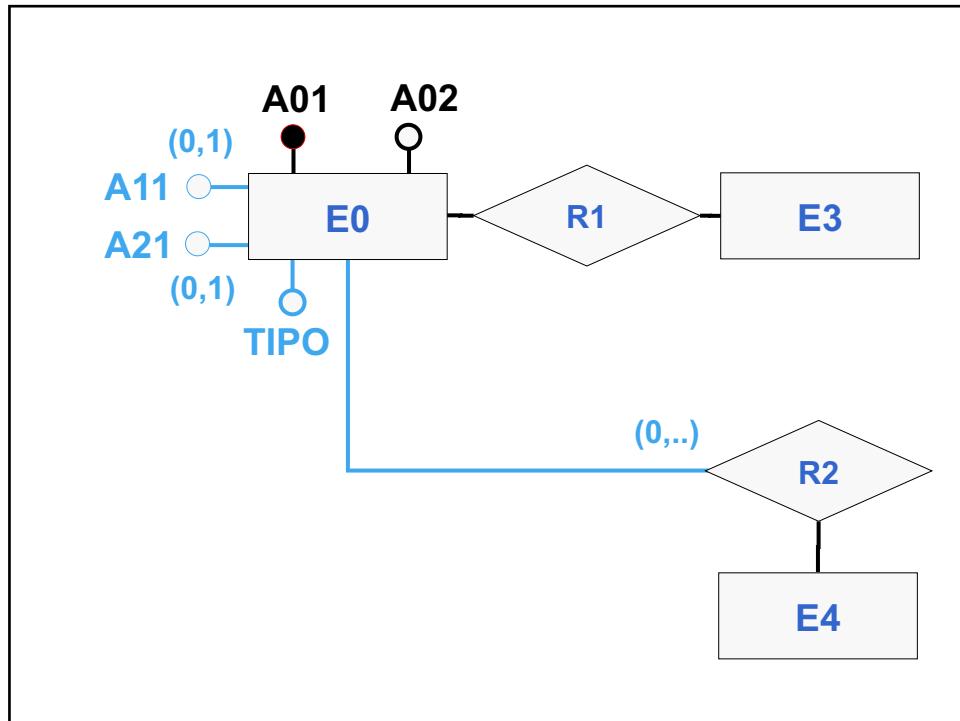


Basi di Dati corso A

203



204



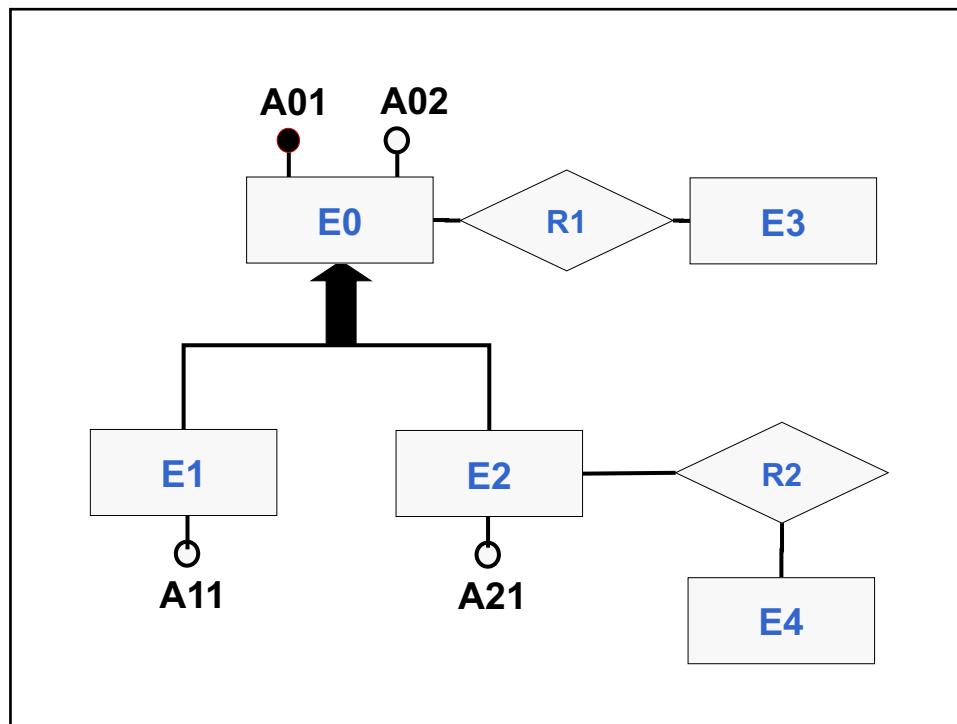
205

## Tre possibilità

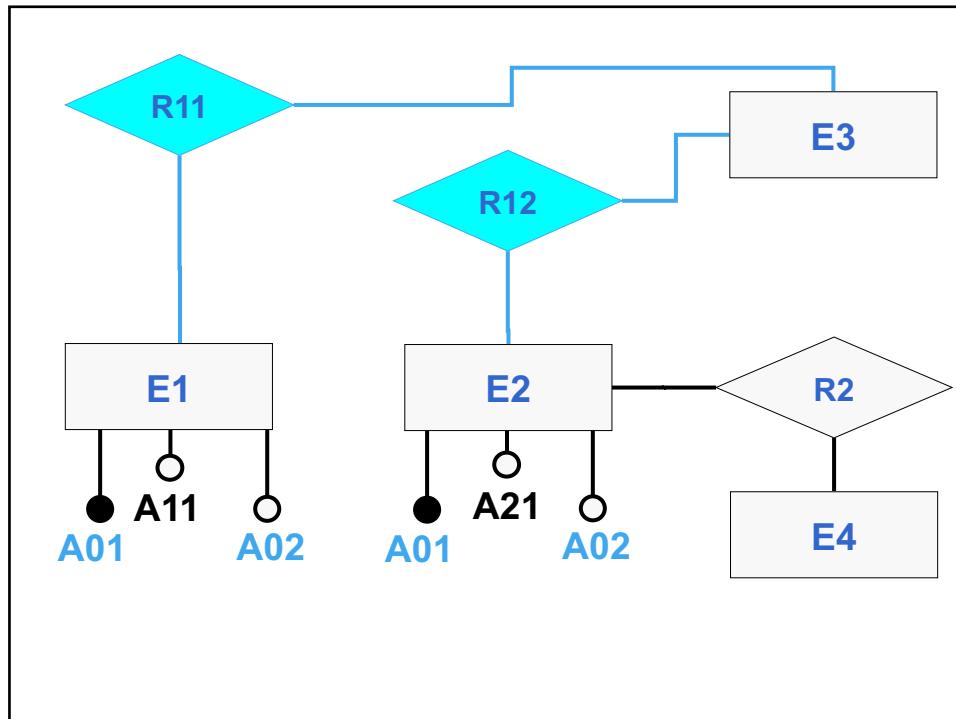
- ▶ accorpamento delle figlie della generalizzazione nel genitore
- ▶ accorpamento del genitore della generalizzazione nelle figlie
- ▶ sostituzione della generalizzazione con relazioni

Basi di Dati corso A

206



207



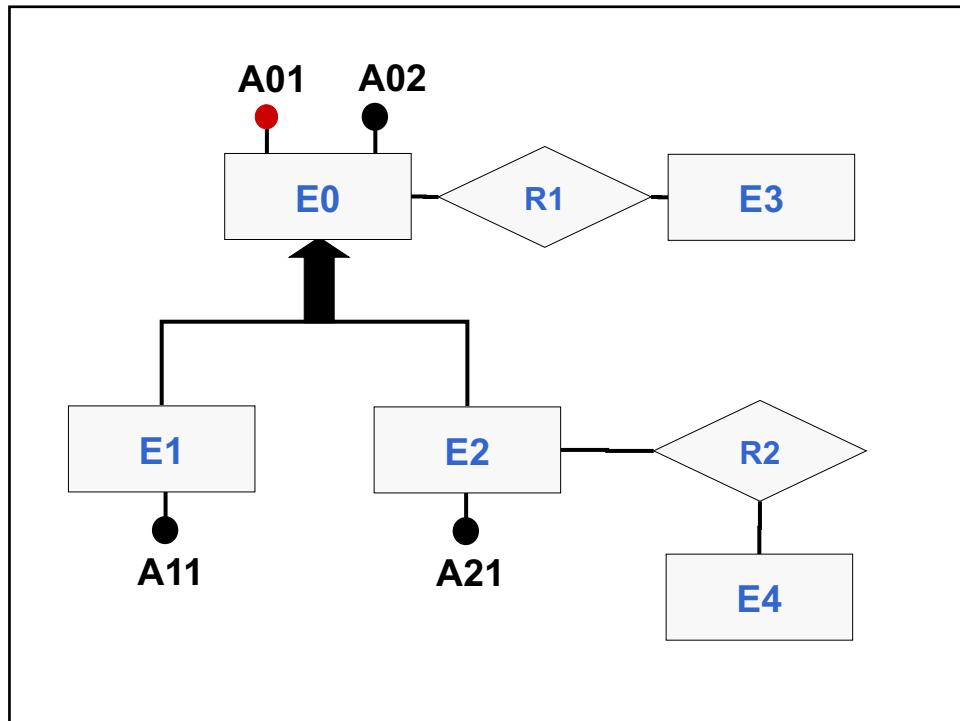
208

## Tre possibilità

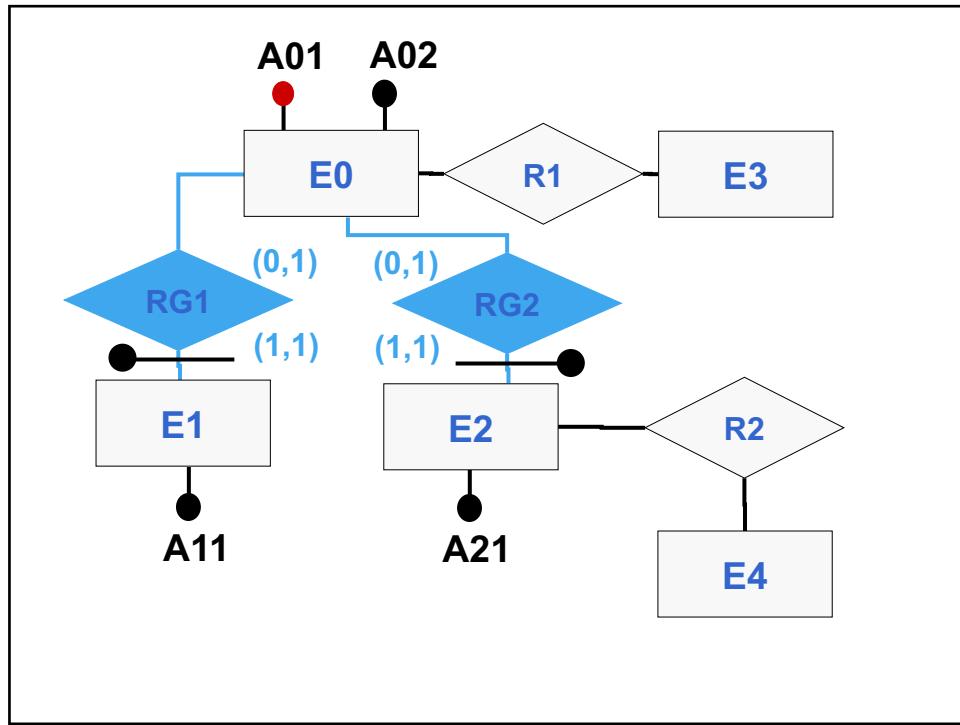
- ▶ accorpamento delle figlie della generalizzazione nel genitore
- ▶ accorpamento del genitore della generalizzazione nelle figlie
- ▶ sostituzione della generalizzazione con relazioni

Basi di Dati corso A

209



210



211

- ▶ la scelta fra le alternative si può fare con metodo simile a quello visto per l'analisi delle ridondanze (però non basato solo sul numero degli accessi)
- ▶ è possibile seguire alcune semplici regole generali

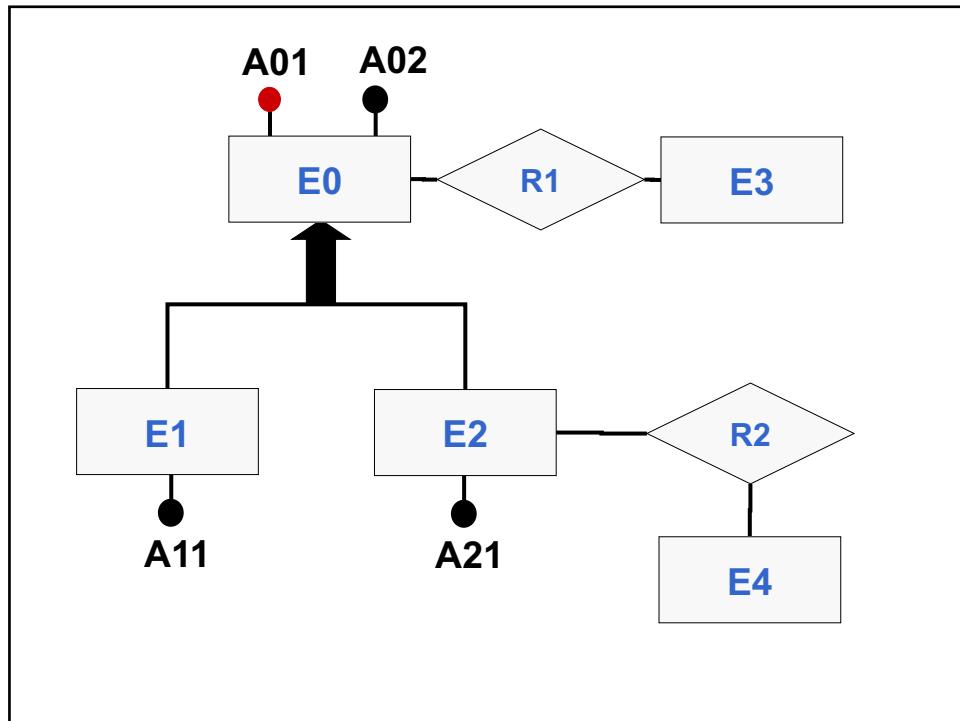
Basi di Dati corso A

212

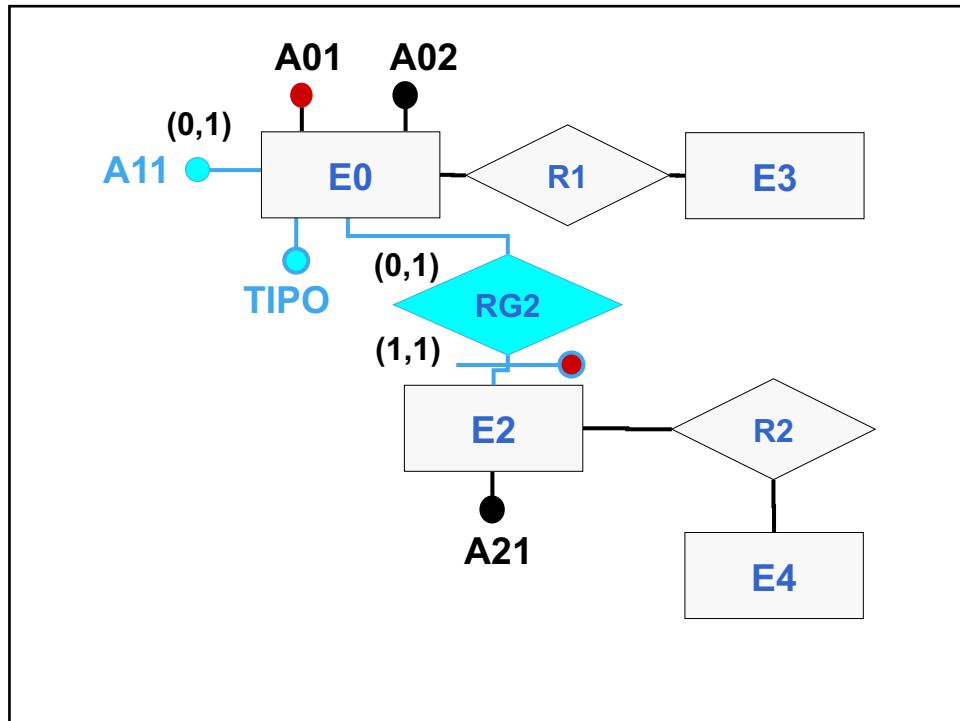
- ▶ gli accessi al genitore e alle figlie sono contestuali?
- ▶ gli accessi alle figlie sono distinti?
- ▶ gli accessi alle entità figlie sono separati dagli accessi al genitore?
- ▶ sono anche possibili soluzioni “ibride”, soprattutto in gerarchie a più livelli

Basi di Dati corso A

213



214



215

## Attività della ristrutturazione

- ▶ Analisi delle ridondanze
- ▶ Eliminazione delle generalizzazioni
- ▶ Partizionamento/accorpamento di entità e relazioni
- ▶ Scelta degli identificatori primari

Basi di Dati corso A

216

- ▶ Ristrutturazioni effettuate per rendere più efficienti le operazioni in base a un semplice principio
- ▶ Gli accessi si riducono
  - ▶ separando attributi di un concetto che vengono acceduti separatamente
  - ▶ raggruppando attributi di concetti diversi acceduti insieme

Basi di Dati corso A

217

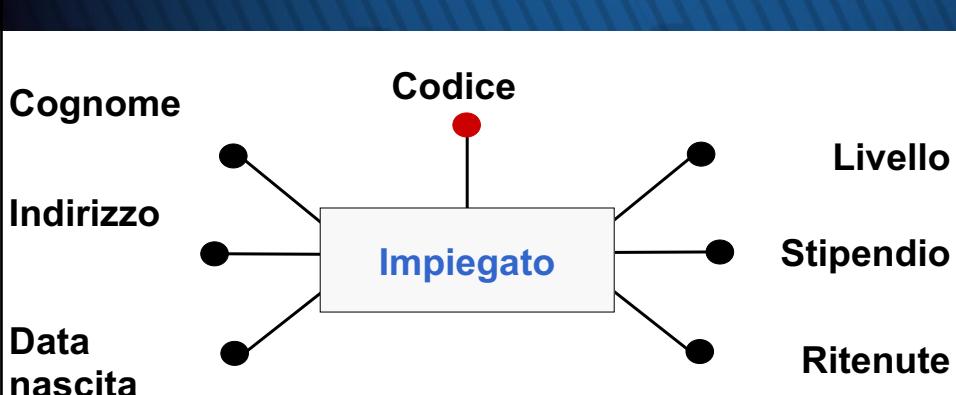
## Ristrutturazioni, casi principali

- ▶ partizionamento verticale di entità
- ▶ partizionamento orizzontale di relationship
- ▶ eliminazione di attributi multivалore
- ▶ accorpamento di entità/ relationship

(v. 7.2.2)

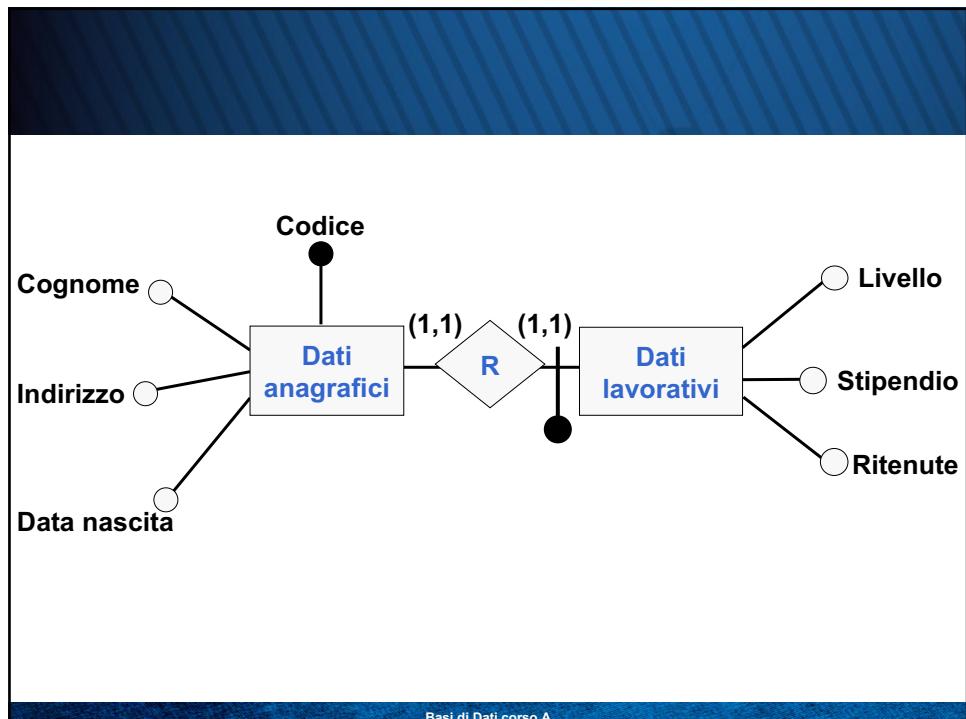
Basi di Dati corso A

218

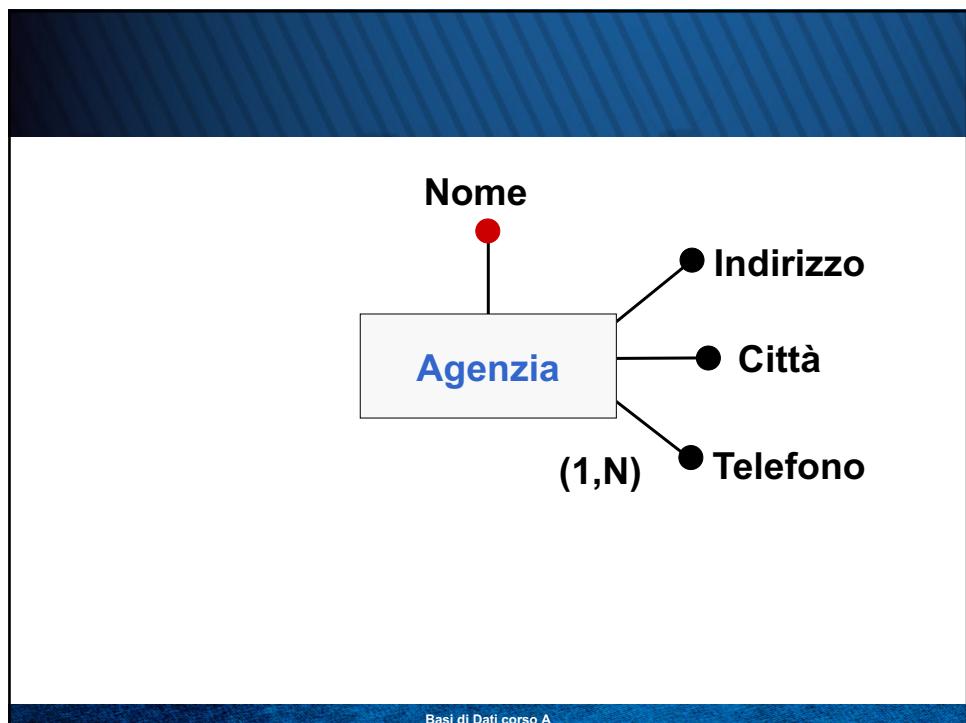


Basi di Dati corso A

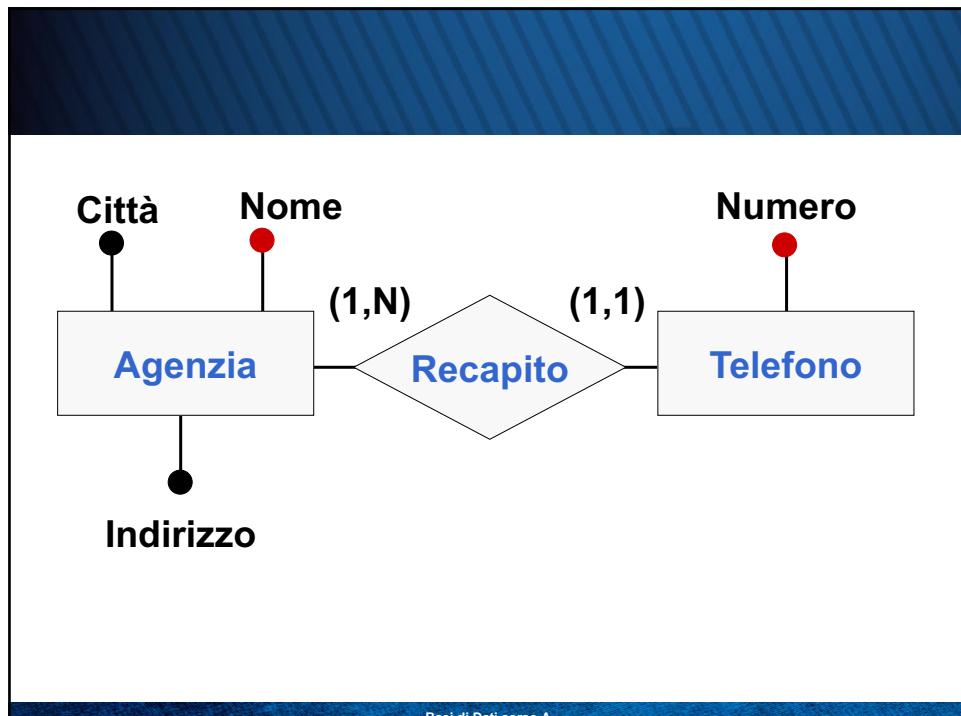
219



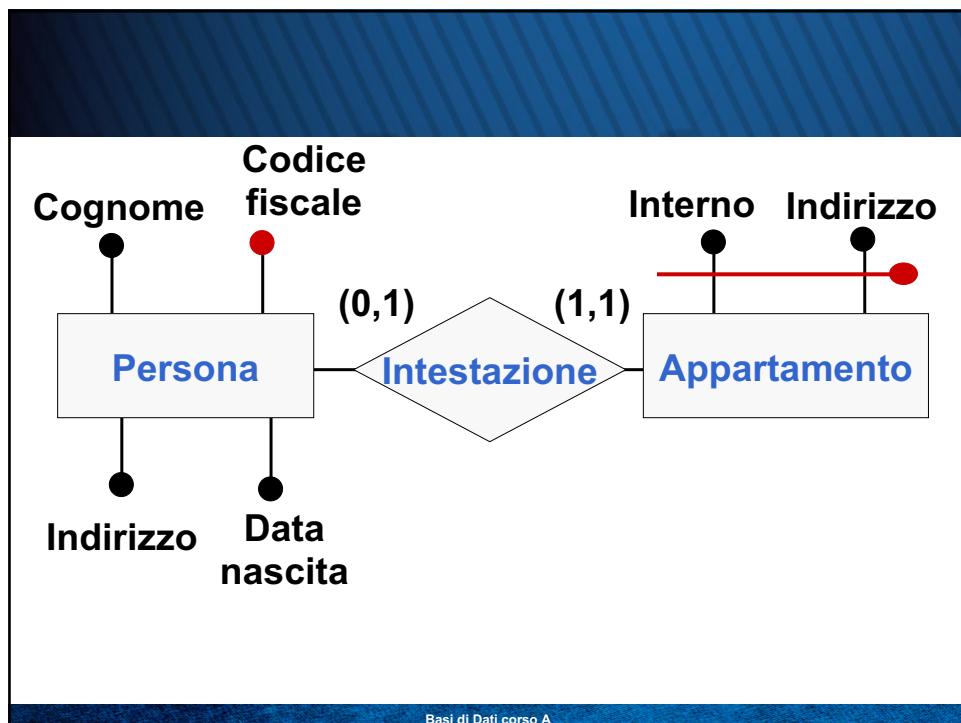
220



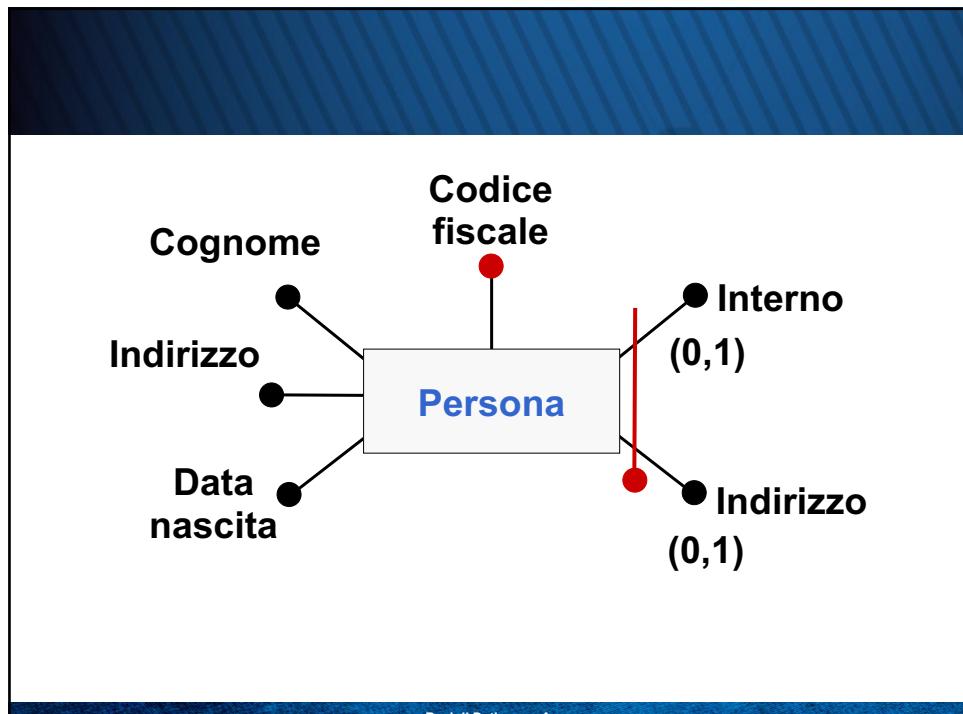
221



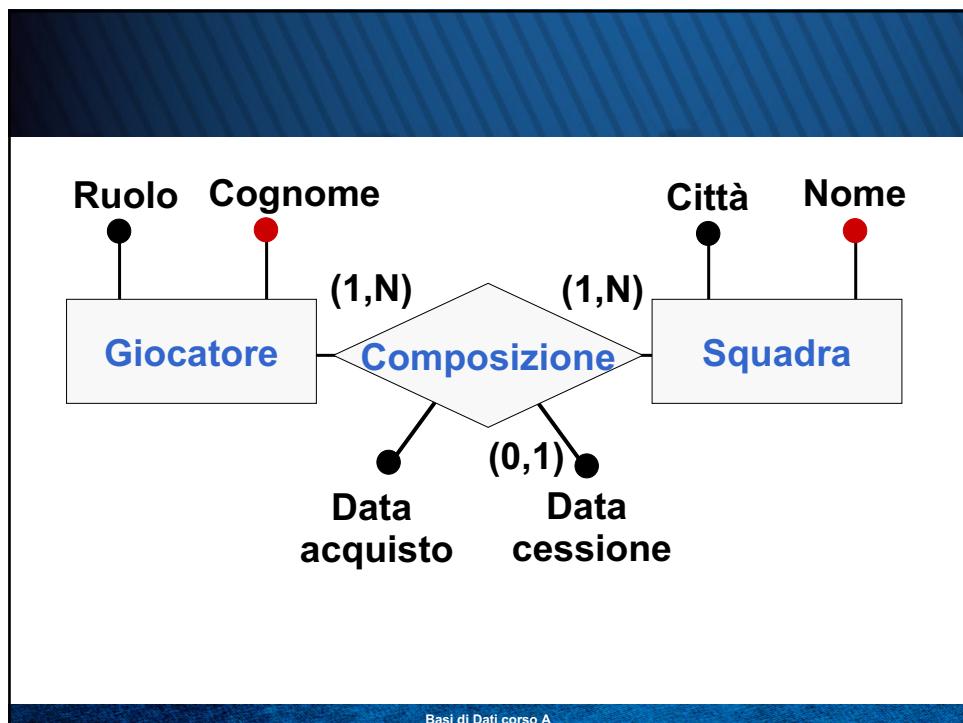
222



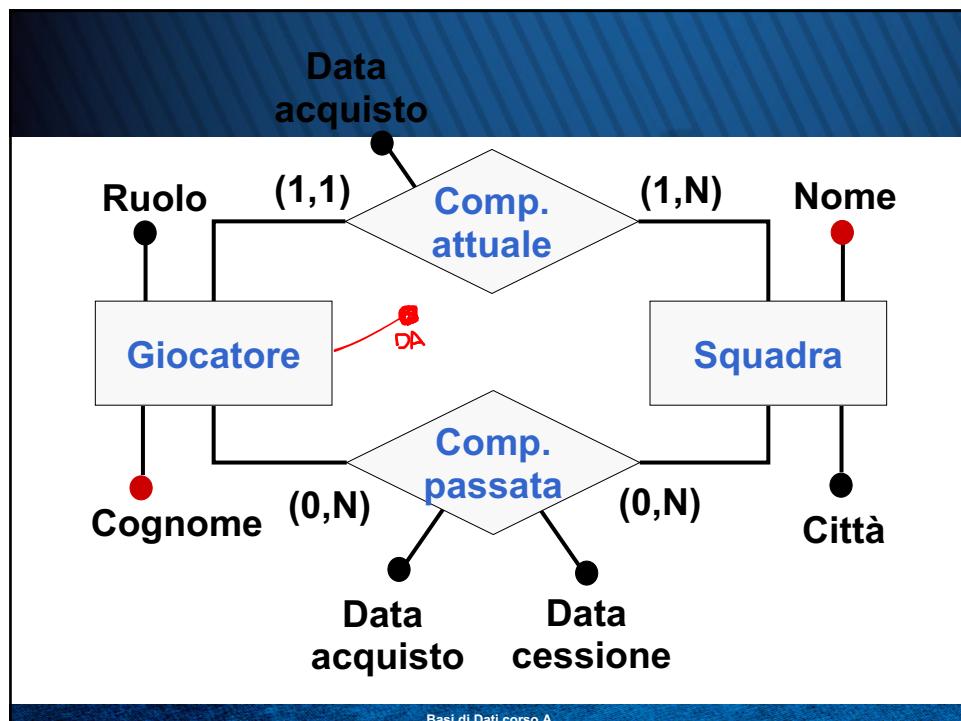
223



224



225



226

## Attività della ristrutturazione

- ▶ Analisi delle ridondanze
- ▶ Eliminazione delle generalizzazioni
- ▶ Partizionamento/accorpamento di entità e relazioni
- ▶ Scelta degli identificatori primari

Basi di Dati corso A

227

## Scelta degli identificatori principali

- ▶ operazione indispensabile per la traduzione nel modello relazionale
- ▶ Criteri
  - ▶ assenza di opzionalità
  - ▶ semplicità
  - ▶ utilizzo nelle operazioni più frequenti o importanti

Basi di Dati corso A

228

## Se nessuno degli identificatori soddisfa i requisiti?

- ▶ Si introducono nuovi attributi (**codici**) contenenti valori speciali generati appositamente per questo scopo

Basi di Dati corso A

229

## MODELLO LOGICO

Basi di Dati corso A

230

## Traduzione verso il modello relazionale

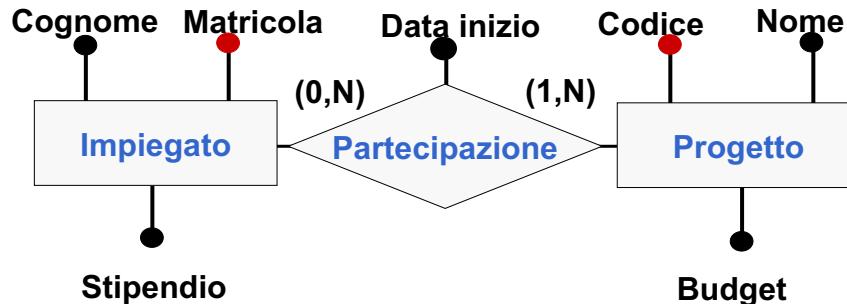
### ► idea di base:

- le entità diventano relazioni sugli stessi attributi
- le associazioni (ovvero le relazioni E-R) diventano relazioni sugli identificatori delle entità coinvolte (più gli attributi propri)

Basi di Dati corso A

231

## Entità e relationship molti a molti



**Impiegato(Matricola, Cognome, Stipendio)**

**Progetto(Codice, Nome, Budget)**

**Partecipazione(Matricola, Codice, DataInizio)**

Basi di Dati corso A

232

## Entità e relationship molti a molti

**Impiegato(Matricola, Cognome, Stipendio)**

**Progetto(Codice, Nome, Budget)**

**Partecipazione(Matricola, Codice, DataInizio)**

- ▶ con vincoli di integrità referenziale fra
  - ◀ Matricola in Partecipazione e (la chiave di) Impiegato
  - ◀ Codice in Partecipazione e (la chiave di) Progetto

Basi di Dati corso A

233

**Nomi più espressivi per gli attributi  
della chiave della relazione che  
rappresenta la relationship**

**Impiegato(Matricola, Cognome, Stipendio)**

**Progetto(Codice, Nome, Budget)**

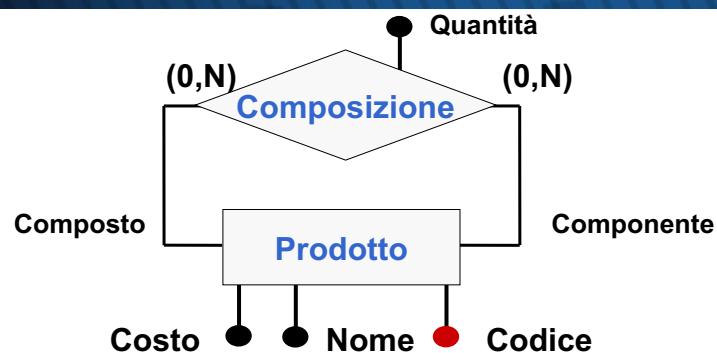
**Partecipazione(Matricola, Codice, DataInizio)**

**Partecipazione(Impiegato, Progetto, DataInizio)**

Basi di Dati corso A

234

## Relationship ricorsive

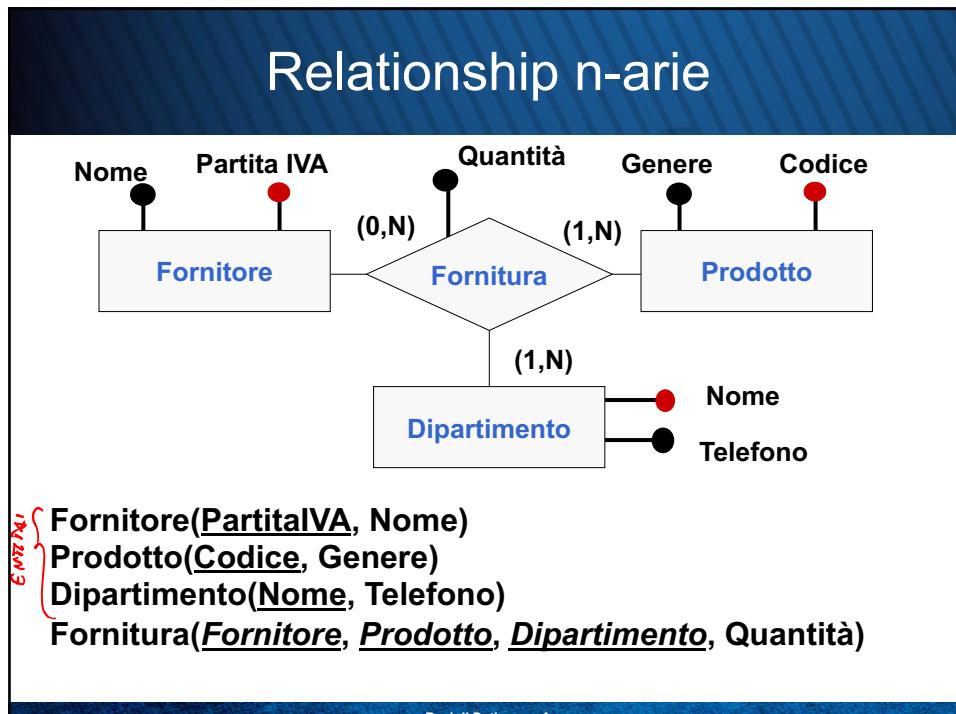


**Prodotto(Codice, Nome, Costo)**

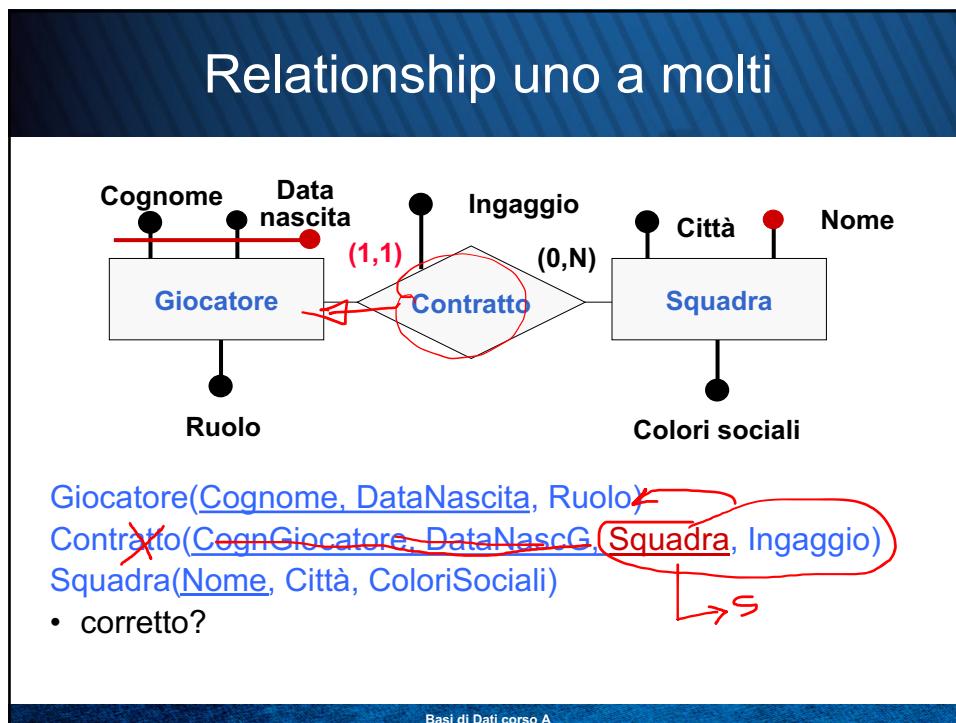
**Composizione(Composto, Componente, Quantità)**

Basi di Dati corso A

235



236



237

## Soluzione più compatta

**Giocatore(Cognome, DataNascita, Ruolo)**  
**Contratto(CognGiocatore, DataNascG, Squadra, Ingaggio)**  
**Squadra(Nome, Città, ColoriSociali)**

**Giocatore(Cognome, DataNasc, Ruolo, Squadra, Ingaggio)**  
**Squadra(Nome, Città, ColoriSociali)**

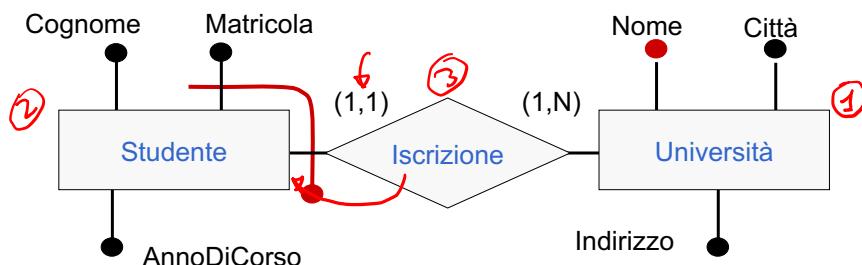
$\left\{ \begin{array}{l} \text{not null} \leftarrow (1,1) \\ \text{not null} \leftarrow (0,1) \end{array} \right.$

- con vincolo di integrità referenziale fra **Squadra** in **Giocatore** e la chiave di **Squadra**
- se la cardinalità minima della relationship è 0, allora **Squadra** in **Giocatore** deve ammettere valore nullo

Basi di Dati corso A

238

## Entità con identificazione esterna



**Studente(Matricola, Università, Cognome, AnnoDiCorso)**

università

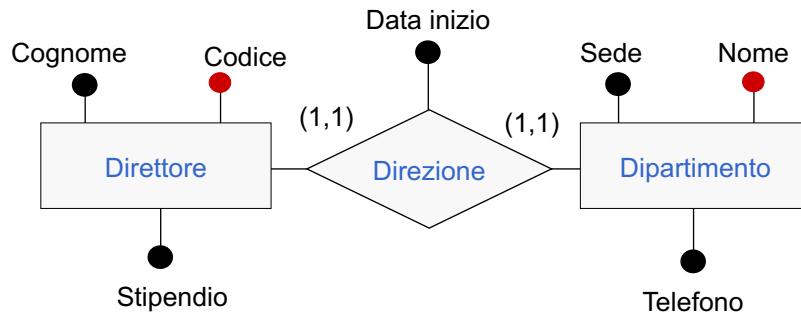
**Università(Nome, Città, Indirizzo)**

- con vincolo ...

Basi di Dati corso A

239

## Relationship uno a uno

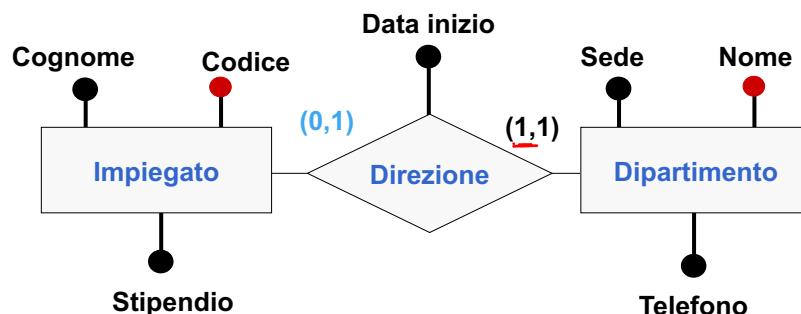


- varie possibilità
  - fondere da una parte o dall'altra
  - fondere tutto?

Basi di Dati corso A

240

## Una possibilità privilegiata



**Impiegato** (Codice, Cognome, Stipendio)

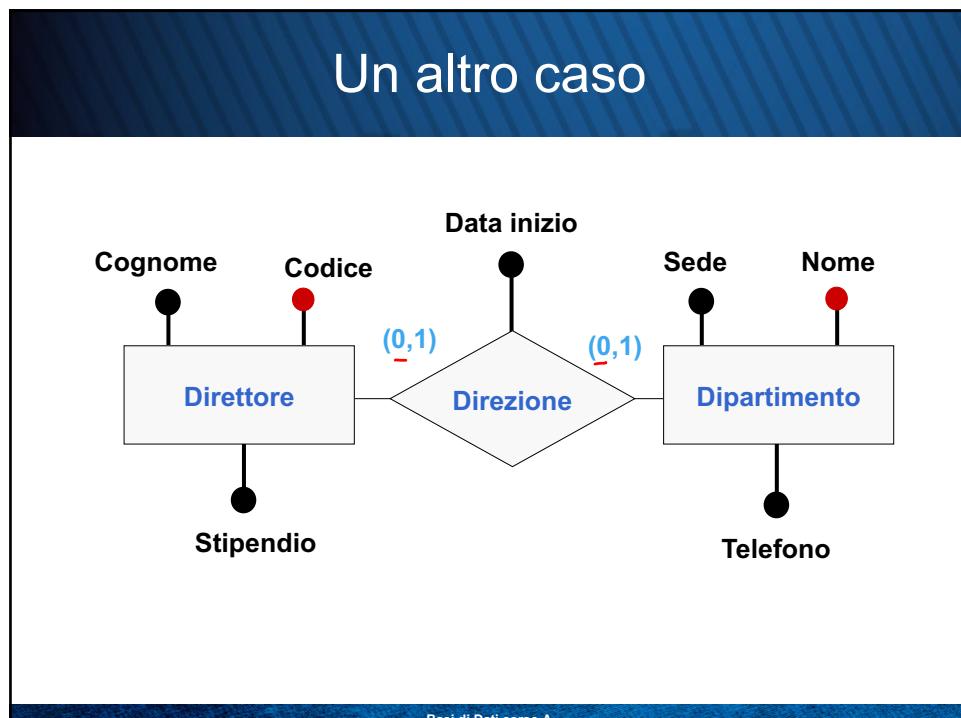
**Dipartimento** (Nome, Sede, Telefono, Direttore, InizioD)

- con vincolo di integrità referenziale, senza valori nulli

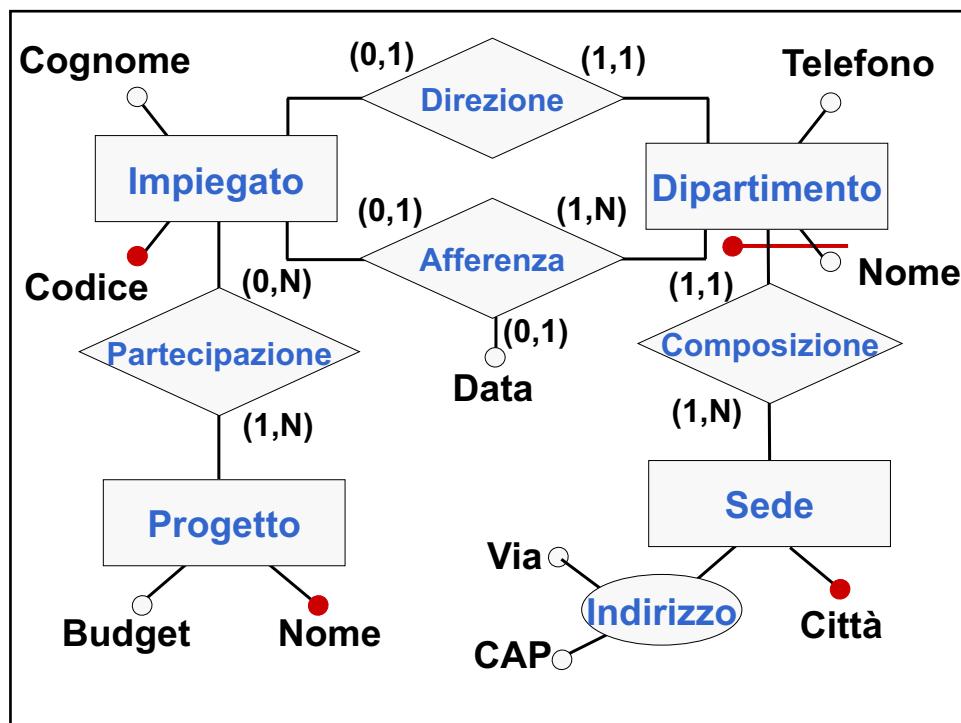
↳ Imp.

Basi di Dati corso A

241



242



243

## Schema finale

Impiegato(Codice, Cognome, *Dipartimento*\*, Data\*)

Dipartimento(Nome, Sede, Telefono, *Direttore*)

Sede(Città, Via, CAP)

Progetto(Nome, Budget)

Partecipazione(Impiegato, Progetto)

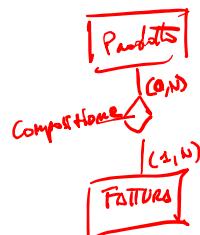
\*: opzionalità

Basi di Dati corso A

244

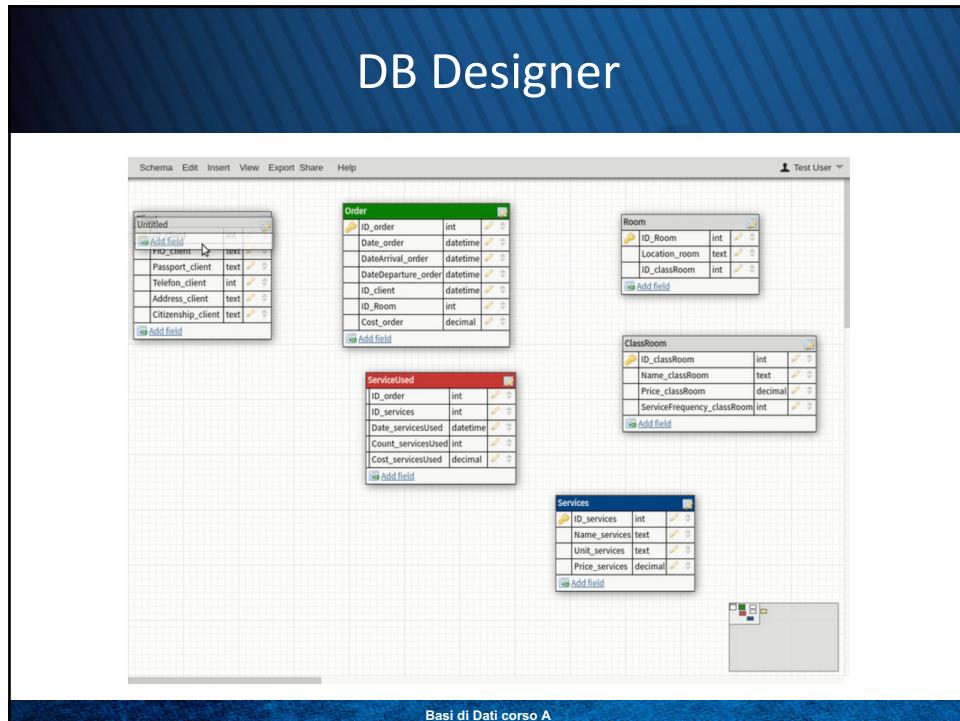
## Strumenti di supporto

- ▶ Esistono sul mercato prodotti CASE che forniscono un supporto a tutte le fasi della progettazione di basi di dati

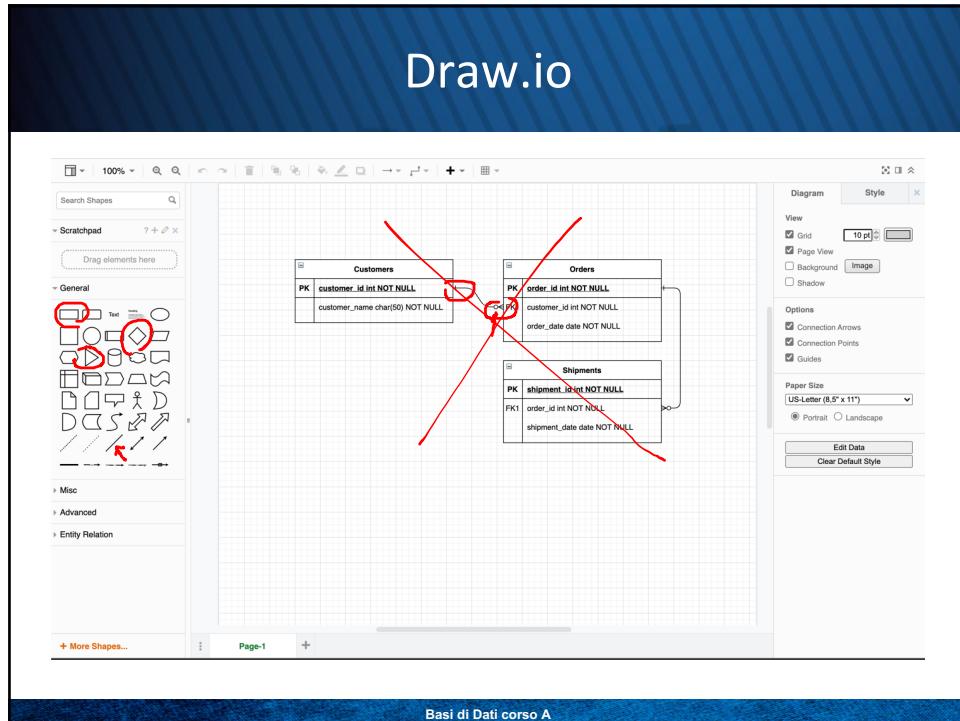


Basi di Dati corso A

245



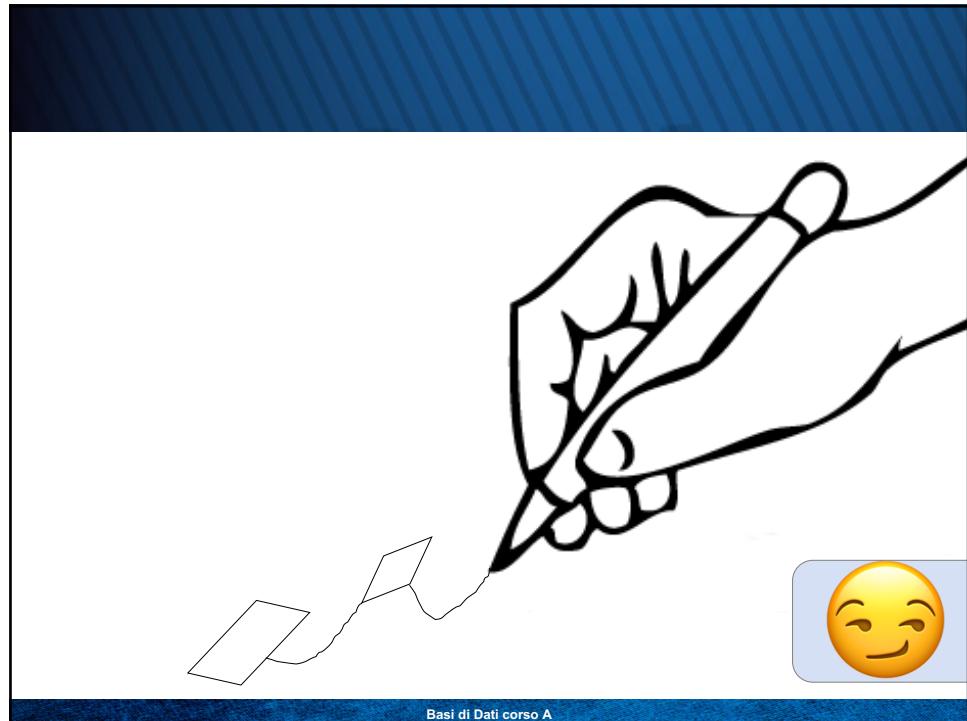
246



247

Qual è lo strumento migliore?

248



Basi di Dati corso A

249

## Schema EER → Schema relazionale

- Step 1: Traduzione di tipi di entità
- Step 2: Traduzione di tipi di entità deboli
- Step 3: Traduzione associazioni binarie 1:1
- Step 4: Traduzione associazioni binarie 1:N
- Step 5: Traduzione associazioni binaria M:N
- Step 6: Traduzione attributi multivalore
- Step 7: Traduzione associazioni N-arie
- Step 8: Traduzione specializzazioni/generalizzazioni
  - superclasse e sottoclasse
  - solo sottoclassi
  - singola relazione con un attributo tipo
  - singola relazione con molti attributi tipo
- Step 9: Traduzione tipi di unione

Basi di Dati corso A

250

## Forme normali

251

## Forme normali

- ▶ Una forma normale è una proprietà di una base di dati relazionale che ne garantisce la “qualità”, cioè l'assenza di determinati difetti
- ▶ Quando una relazione non è normalizzata:
  - ▶ presenta ridondanze
  - ▶ si presta a comportamenti poco desiderabili durante gli aggiornamenti
- ▶ Le forme normali sono di solito definite sul modello relazionale, ma hanno anche senso in altri contesti, es. nel modello E-R

Basi di Dati corso A

252

## Normalizzazione

- ▶ Procedura che permette di trasformare schemi non normalizzati in schemi che soddisfano una forma normale
- ▶ Da usare come tecnica di **verifica** dei risultati della progettazione di una base di dati
- ▶ Non costituisce una metodologia di progettazione

Basi di Dati corso A

253

## Prova del nove

- ▶ Consuetudine delle elementari.
- ▶ Esempio:  $1902 \times 1964 = 3\ 735\ 528$



- $1902 \rightarrow 1+9+0+2 = 12 \rightarrow 1+2 = 3$
- $1964 \rightarrow 1+9+6+4 = 20 \rightarrow 2+0 = 2$
- $3\ 735\ 528 \rightarrow 3+7+3+5+5+2+8 = 33 \rightarrow 3+3 = 6$

$$\begin{array}{c|c} 3 & 2 \\ \hline 6 & ? \end{array}$$

$$2 \times 3 = 6$$

1.Se i due numeri sono *diversi* allora il risultato è senz'altro *errato*  
2.Se i due numeri sono *uguali* allora il risultato può essere corretto

$$\begin{array}{c|c} 3 & 2 \\ \hline 6 & 6 \end{array}$$

Basi di Dati corso A

254

## linee guida (non ortogonali)

- ▶ progettare schemi semplici da spiegare
- ▶ progettare schemi esenti da anomalie di inserimento, modifica, cancellazione
- ▶ evitare nelle relazioni attributi con valori nulli frequenti
- ▶ progettare schemi ricostruibili (usando equijoin tra chiavi esterne e chiavi primarie)

Basi di Dati corso A

255

## Esempio di relazione

Impiegato	Stipendio	Progetto	Bilancio	Funzione
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

Basi di Dati corso A

256

## Esempio di relazione



Impiegato	Stipendio	Progetto	Bilancio	Funzione
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
{ 1 Mori 2 Mori}	48 → 55	Marte	2	direttore
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	progettista

Basi di Dati corso A

257

## Anomalie

- ▶ Lo stipendio di ciascun impiegato è ripetuto in tutte le ennuple relative
  - ▶ ridondanza
- ▶ Se lo stipendio di un impiegato varia, è necessario modificare diverse ennuple
  - ▶ anomalia di aggiornamento
- ▶ Se un impiegato interrompe la partecipazione a tutti i progetti, deve essere cancellato
  - ▶ anomalia di cancellazione
- ▶ Un nuovo impiegato senza progetto non può essere inserito
  - ▶ anomalia di inserimento

Basi di Dati corso A

258

## Fenomeni indesiderabili

- ▶ abbiamo usato un'unica relazione per rappresentare informazioni eterogenee
  - ▶ gli impiegati con i relativi stipendi
  - ▶ i progetti con i relativi bilanci
  - ▶ le partecipazioni degli impiegati ai progetti con le relative funzioni
- ▶ SQL consente aggiornamento con una sola istruzione

Basi di Dati corso A

259

## Dipendenza funzionale

- ▶ Vincolo di integrità che consente di studiare in maniera sistematica tali aspetti

### ▶ Proprietà

- ▶ Ogni impiegato ha un solo stipendio (anche se partecipa a più progetti)
- ▶ Ogni progetto ha un bilancio
- ▶ Ogni impiegato in ciascun progetto ha una sola funzione (ma può avere funzioni diverse in progetti diversi)

Basi di Dati corso A

260

## Dipendenza funzionale

- ▶ vincolo di integrità
- ▶ relazione  $r$  su  $R(X)$
- ▶ due sottoinsiemi non vuoti  $Y$  e  $Z$  di  $X$
- ▶ esiste in  $r$  una dipendenza funzionale (FD) da  $Y$  a  $Z$  se

per ogni coppia di ennuple  $t_1$  e  $t_2$  di  $r$  con gli stessi valori su  $Y$  ( $t_1[Y] = t_2[Y]$ )

risulta che  $t_1$  e  $t_2$  hanno gli stessi valori anche su  $Z$  ( $t_1[Z] = t_2[Z]$ )

*Impiegato → Stipendio       $Y \rightarrow Z$*

Basi di Dati corso A

261

## Notazione

$Y \rightarrow Z$

► Esempi:

Impiegato → Stipendio

Progetto → Bilancio

Impiegato Progetto → Funzione

$Y$   $Z$

Se  $Z = \{A_1, \dots, A_k\}$  allora  $Y \rightarrow Z$  se e solo se  
 $Y \rightarrow A_i$ , è soddisfatta per ogni  $i = 1, \dots, k$

Basi di Dati corso A

262

## FD banale

► Impiegato Progetto → Progetto

$Y$   $Z$   
 $|A|=1$

►  $Y \rightarrow A$  è non banale se  $A$  non appartiene a  $Y$

►  $Y \rightarrow Z$  è non banale se nessun attributo in  $Z$   
 appartiene a  $Y$

$|Z| > 1$

Basi di Dati corso A

263

## Legame con il vincolo di chiave

- ▶ Sia K chiave per una relazione r
- ▶ Esiste una FD tra K e ogni altro attributo di r (dalla definizione di chiave e di FD)
- ▶ Considerato che Impiegato e Progetto formano una chiave allora vale la FD

Impiegato Progetto → Funzione

- ▶ Quindi il vincolo di FD generalizza il vincolo di chiave, in particolare, una FD  $Y \rightarrow Z$  su  $R(X)$  degenera in vincolo di chiave se  $Y \cup Z = X$

Basi di Dati corso A

264

## Anomalie legate ad alcune FD (ridondanza)

- ▶ gli impiegati hanno un unico stipendio

Impiegato → Stipendio

- ▶ i progetti hanno un unico bilancio

Progetto → Bilancio

Basi di Dati corso A

265

## Non tutte le FD causano anomalie

- ▶ In ciascun progetto, un impiegato svolge una sola funzione

Impiegato Progetto → Funzione

Basi di Dati corso A

266

## Una differenza fra FD

Impiegato → Stipendio  
Progetto → Bilancio

- ▶ possono causare anomalie  
Impiegato Progetto → Funzione
- ▶ non causa anomalie

Basi di Dati corso A

267

Impiegato	Stipendio	Progetto	Bilancio	Funzione
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

**Impiegato → Stipendio**

**Progetto → Bilancio**

**Impiegato Progetto → Funzione**

Basi di Dati corso A

268

## FD e anomalie

- ▶ La terza FD corrisponde ad una superchiave e non causa anomalie
- ▶ Le prime due FD non corrispondono a chiavi e sono cause di anomalie
- ▶ La relazione contiene alcune informazioni legate alla chiave e altre ad attributi che non formano una chiave

Basi di Dati corso A

269

## motivazione

- ▶ si è usato un'unica relazione per rappresentare informazioni eterogenee
  - ▶ gli impiegati con i relativi stipendi
  - ▶ i progetti con i relativi bilanci
  - ▶ le partecipazioni degli impiegati ai progetti con le relative funzioni

Basi di Dati corso A

270

Impiegato → Stipendio

Progetto → Bilancio

Impiegato Progetto → Funzione

Impiegato Progetto è chiave

Impiegato da solo no

Progetto da solo no

Le anomalie sono causate dalla presenza di concetti eterogenei:

- ▶ proprietà degli impiegati (lo stipendio)
- ▶ proprietà di progetti (il bilancio)
- ▶ proprietà della chiave Impiegato Progetto

271

## Forma normale di Boyce e Codd (BCNF)

- ▶  $r$  è in forma normale di Boyce e Codd se per ogni dipendenza funzionale (non banale)  $X \rightarrow A$  definita su di essa,  
 $X$  contiene una chiave  $K$  di  $r$  ( $X$  superchiave per  $r$ )
- ▶ La forma normale richiede che i concetti in una relazione siano omogenei (solo proprietà direttamente associate alla chiave)
- ▶ BCNF non presenta anomalie e ridondanze

Basi di Dati corso A

273

## Decomposizione in BCNF

- ▶ Se una relazione non è in BCNF la si sostituisce con altre relazioni che soddisfano la BCNF decomponendo sulla base delle dipendenze funzionali, al fine di separare i concetti

Basi di Dati corso A

274

<u>Impiegato</u>	<u>Stipendio</u>	<u>Progetto</u>	<u>Bilancio</u>	<u>Funzione</u>
<u>Impiegato</u>	<u>Stipendio</u>	<u>Progetto</u>	<u>Bilancio</u>	<u>Funzione</u>
Rossi Verdi	20 35	Marte Progetto	2	tecnico progettista progettista
Impiegato Rossi Verdi Neri Mori Bianchi	Stipendio 20 35 55 48 48	Impiegato Rossi Verdi Verdi Neri Neri Neri Mori Mori Bianchi Bianchi	Marte Giove Venere Venere Giove Marte Marte Venere Venere Giove	progettista progettista direttore consulente consulente direttore progettista progettista direttore
Mori Bianchi	Mori Bianchi	40	Giove	15
				2 15 15
				direttore progettista progettista direttore

275

## Decomposizione naturale

- ▶ Spesso si decomponete in base alle FD aventi primo membro differente
- ▶ Ma può non essere possibile o complicato da scegliere

276

## esempio

Impiegato	Progetto	Sede
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Venere	Milano
Neri	Saturno	Milano
Neri	Venere	Milano

**Impiegato → Sede**  
**Progetto → Sede**  
**Come decomponiamo?**

Basi di Dati corso A

277

## Decomponiamo sulla base delle dipendenze

Impiegato	Progetto	Sede
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Venere	Milano
Neri	Saturno	Milano
Neri	Venere	Milano

Impiegato	Sede
Rossi	Roma
Verdi	Milano
Neri	Milano

Progetto	Sede
Marte	Roma
Giove	Milano
Venere	Milano
Saturno	Milano

Basi di Dati corso A

278

Proviamo a ricostruire con join naturale			
Impiegato	Sede	Progetto	Sede
Rossi	Roma	Marte	Roma
Verdi	Milano	Giove	Milano
Neri	Milano	Saturno	Milano
		Venere	Milano
Impiegato	Progetto	Sede	
Rossi	Marte	Roma	
Verdi	Giove	Milano	
Verdi	Venere	Milano	
Neri	Saturno	Milano	
Neri	Venere	Milano	
Verdi	Saturno	Milano	
Neri	Giove	Milano	

**Diversa dalla relazione di partenza!**

Basi di Dati corso A

279

## PROPRIETÀ DELLE DECOMPOSIZIONI

Basi di Dati corso A

280

## Decomposizione senza perdita

- ▶ Una relazione  $r$  si **decomponе senza perdita** su  $X_1$  e  $X_2$  se il join delle proiezioni di  $r$  su  $X_1$  e  $X_2$  è uguale a  $r$  stessa (cioè non contiene ennuple spurie)
- ▶ La decomposizione senza perdita è garantita se gli **attributi comuni** contengono una chiave per almeno una delle relazioni composte

Basi di Dati corso A

281

## Decomposizione senza perdita

Impiegato	Progetto	Sede
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Venere	Milano
Neri	Saturno	Milano
Neri	Venere	Milano

Impiegato	Sede
Rossi	Roma
Verdi	Milano
Neri	Milano

Impiegato	Progetto
Rossi	Marte
Verdi	Giove
Verdi	Venere
Neri	Saturno
Neri	Venere

**Impiegato → Sede**  
**Progetto → Sede**

Basi di Dati corso A

282

## Un altro problema

- Inserimento di una nuova ennupla che specifica la partecipazione dell'impiegato Neri, che opera a Milano, al progetto Marte

Impiegato	Sede	Impiegato	Progetto
Rossi	Roma	Rossi	Marte
Verdi	Milano	Verdi	Giove
Neri	Milano	Verdi	Venere
		Neri	Saturno
		Neri	Venere

Impiegato → Sede  
Progetto → Sede

Basi di Dati corso A

283

Impiegato	Progetto	Sede
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Venere	Milano
Neri	Saturno	Milano
Neri	Venere	Milano
Neri	Marte	Milano

Basi di Dati corso A

284

Impiegato	Sede	Impiegato	Progetto
Rossi	Roma	Rossi	Marte
Verdi	Milano	Verdi	Giove
Neri	Milano	Neri	Venere
		Neri	Saturno
			Venere
		Neri	Marte

Basi di Dati corso A

285

## Conservazione delle dipendenze

- ▶ Una decomposizione conserva le dipendenze se ciascuna delle dipendenze funzionali dello schema originario coinvolge attributi che compaiono tutti insieme in uno degli schemi decomposti
- ▶  $\text{Progetto} \rightarrow \text{Sede}$  non è conservata

Impiegato	Sede
-----------	------

Impiegato	Progetto
-----------	----------

Impiegato  $\rightarrow$  Sede  
 Progetto  $\rightarrow$  Sede

Basi di Dati corso A

286

## Qualità delle decomposizioni

- ▶ Una decomposizione dovrebbe sempre soddisfare:
  - ▶ la **decomposizione senza perdita**, che garantisce la ricostruzione delle informazioni originarie
  - ▶ la **conservazione delle dipendenze**, che garantisce il mantenimento dei vincoli di integrità originari

Basi di Dati corso A

287

## Una relazione non-normalizzata *de - mormalizzate*

<u>Dirigente</u>	<u>Progetto</u>	<u>Sede</u>
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Marte	Milano
Neri	Saturno	Milano
Neri	Venere	Milano

Ogni dirigente opera presso una sede.

Ogni progetto può avere più dirigenti che sono responsabili, ma in sedi diverse

Ogni dirigente può essere responsabile di più progetti però, per ogni sede, un progetto ha un solo responsabile.

**Progetto Sede → Dirigente**  
**Dirigente → Sede**

Basi di Dati corso A

288

## La decomposizione è problematica

- ▶ Progetto Sede → Dirigente coinvolge tutti gli attributi e quindi nessuna decomposizione può preservare tale dipendenza
- ▶ quindi la BCNF può non essere raggiungibile

Basi di Dati corso A

289

## Un'altra forma normale

- ▶ È quindi necessario rilassare i vincoli di BCNF
- ▶ Una relazione  $r$  è in terza forma normale se, per ogni FD (non banale)  $X \rightarrow A$  definita su  $r$ , è verificata almeno una delle condizioni:
  - ▶  $X$  contiene una chiave  $K$  di  $r$  (BCNF)
  - ▶  $A$  appartiene ad almeno una chiave di  $r$

Basi di Dati corso A

290

## BCNF e terza forma normale

- ▶ la terza forma normale è meno restrittiva della forma normale di Boyce e Codd (e ammette relazioni con alcune anomalie)
- ▶ ha il vantaggio però di essere sempre “raggiungibile”

**Progetto Sede → Dirigente**  
**Dirigente → Sede**

parte della chiave Progetto Sede

Basi di Dati corso A

291

## Decomposizione in terza forma normale

- ▶ si crea una relazione per ogni gruppo di attributi coinvolti in una dipendenza funzionale
- ▶ si verifica che alla fine una relazione contenga una chiave della relazione originaria
- ▶ **Dipende dalle dipendenze individuate**

Basi di Dati corso A

292

## Una possibile strategia

- ▶ se la relazione non è normalizzata si decomponе in **terza forma normale (3NF)**
- ▶ alla fine si verifica se lo schema ottenuto è anche in **BCNF**
- ▶ Spesso la decomposizione in 3NF produce schemi in BCNF
- ▶ **Se una relazione ha una sola chiave allora le due forme normali coincidono**

Basi di Dati corso A

293

## Altre forme normali

- ▶ Prima: gli attributi sono definiti su valori atomici  
---- commenti?
- ▶ Seconda: se su di essa non sono definite *dipendenze parziali*  
 $R(\text{Impiegato}, \text{Categoria}, \text{stipendio})$
- ▶ Terza: se su di essa non sono definite *dipendenze transitive*

Basi di Dati corso A

295

## Uno schema non decomponibile in BCNF

► considerazione

Dirigente	Progetto	Sede
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Marte	Milano
Neri	Saturno	Milano
Neri	Venere	Milano

**Dirigente → Sede**  
**Progetto Sede → Dirigente**

Basi di Dati corso A

296

## Una possibile riorganizzazione

Dirigente	Progetto	Sede	Reparto
Rossi	Marte	Roma	1
Verdi	Giove	Milano	1
Verdi	Marte	Milano	1
Neri	Saturno	Milano	2
Neri	Venere	Milano	2

**Dirigente → Sede Reparto**  
**Sede Reparto → Dirigente**  
**Progetto Sede → Reparto**

Basi di Dati corso A

297

Decomposizione in BCNF		
$R_1$		
Dirigente	Sede	Reparto
Rossi	Roma	1
Verdi	Milano	1
Neri	Milano	2

$R_2$		
Progetto	Sede	Reparto
Marte	Roma	1
Giove	Milano	1
Marte	Milano	1
Saturno	Milano	2
Venere	Milano	2

Dirigente → Sede Reparto  
 Sede Reparto → Dirigente  
 Progetto Sede → Reparto

$$\begin{matrix} X \\ \hline Y \end{matrix} \rightarrow \begin{matrix} Y \\ \hline \end{matrix}$$

$\Rightarrow R_1, R_2$  sono in  
 forma normale  
 di BC

298

Analisi inadeguata		
<ul style="list-style-type: none"> <li>▶ Spesso la non raggiungibilità della forma normale di Boyce e Codd può dipendere da un'analisi non sufficientemente accurata</li> </ul>		

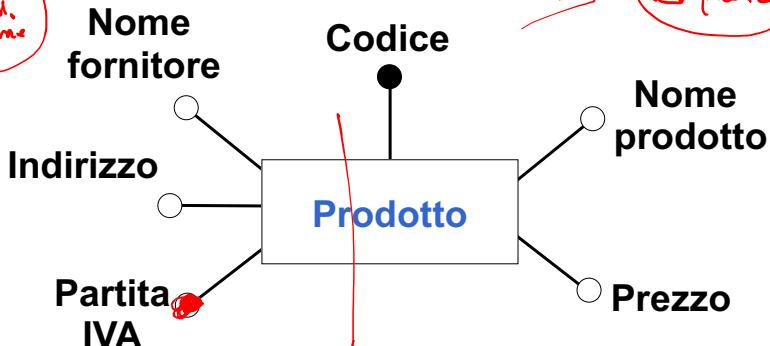
299

## Progettazione e normalizzazione

- ▶ la teoria della normalizzazione può essere usata nella progettazione logica per verificare lo schema relazionale finale
- ▶ si può usare anche durante la progettazione concettuale per verificare la qualità dello schema concettuale

Basi di Dati corso A

300



**Partita IVA → NomeFornitore Indirizzo  
codice → NomeProdotto Prezzo partita IVA**

Basi di Dati corso A

301

## Analisi dell'entità

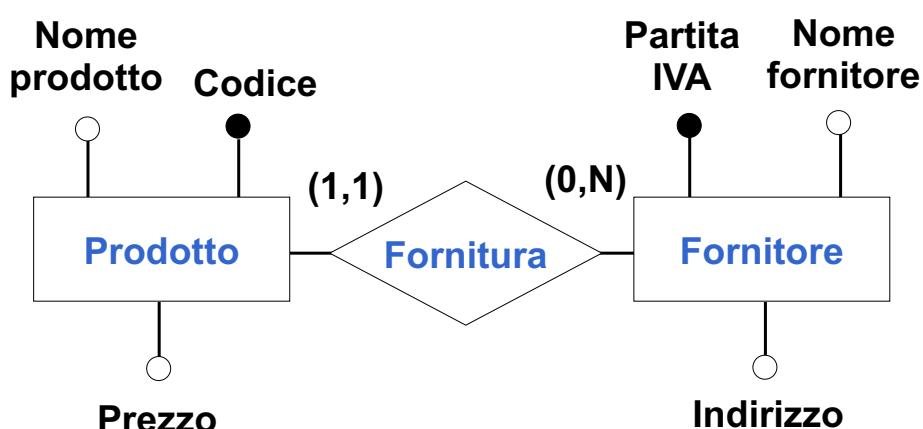
► L'entità viola la terza forma normale a causa della dipendenza:

PartitaIVA → NomeFornitore Indirizzo

► Possiamo decomporre sulla base di questa dipendenza

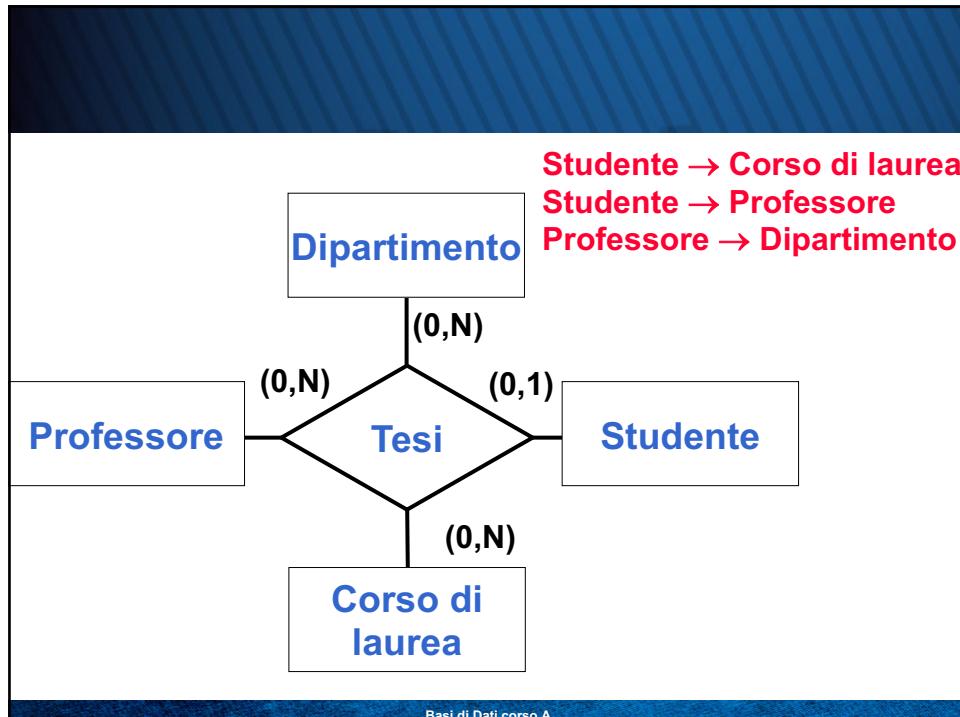
Basi di Dati corso A

302



Basi di Dati corso A

303



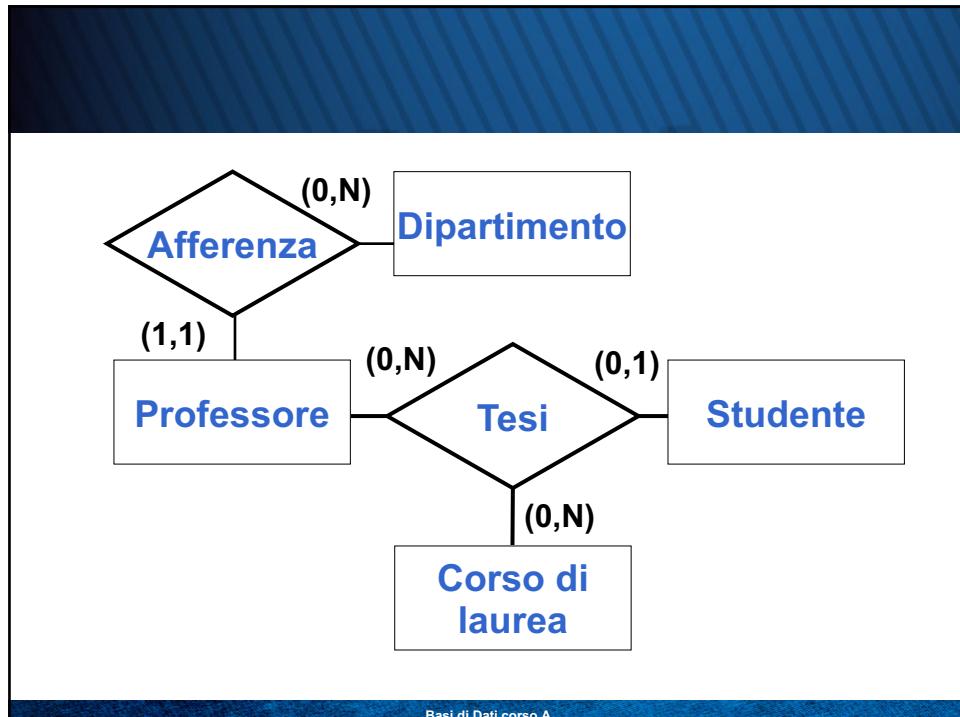
304

## Analisi della relationship

- ▶ La relationship viola la terza forma normale a causa della dipendenza:  
**Professore → Dipartimento**
- ▶ Possiamo decomporre sulla base di questa dipendenza

Basi di Dati corso A

305



Basi di Dati corso A

306

## Ulteriore analisi sulla base delle dipendenze

- ▶ La relationship **Tesi** è in BCNF sulla base delle dipendenze

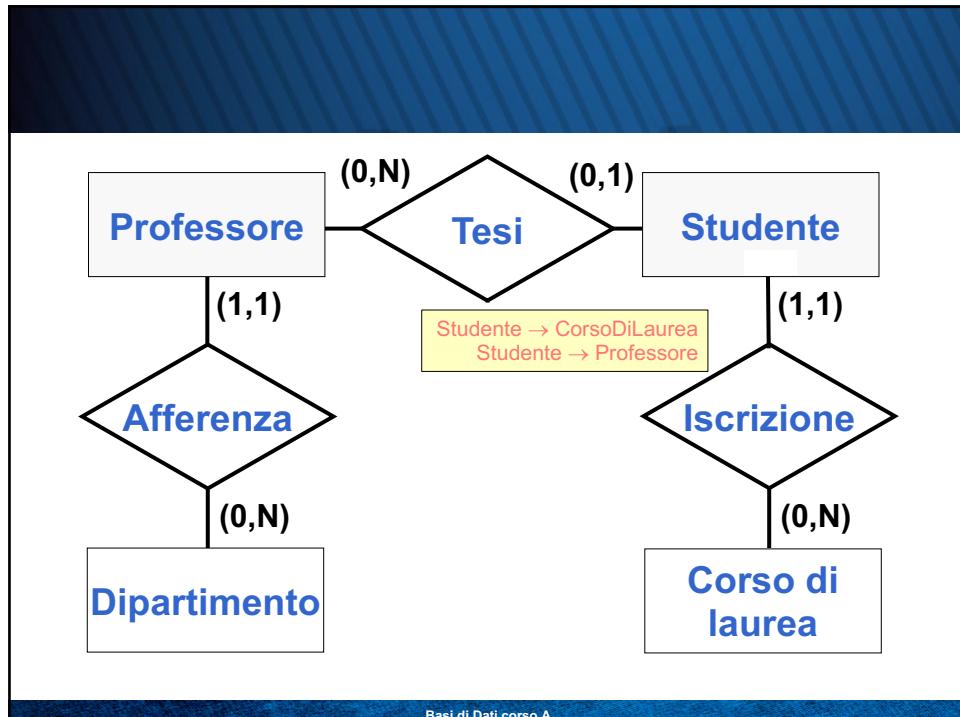
Studente → CorsoDiLaurea

Studente → Professore

- ▶ le due proprietà sono indipendenti
- ▶ questo **suggerisce** una ulteriore decomposizione

Basi di Dati corso A

307



308

## Quante sono le forme normali?

- ▶ 1NF (First Normal Form)
- ▶ 2NF (Second Normal Form)
- ▶ 3NF (Third Normal Form)
- ▶ BCNF (Boyce-Codd Normal Form)
- ancora...
- ▶ 4NF (Fourth Normal Form)
- ▶ 5NF (Fifth Normal Form)

Basi di Dati corso A

309

## repetita juvant

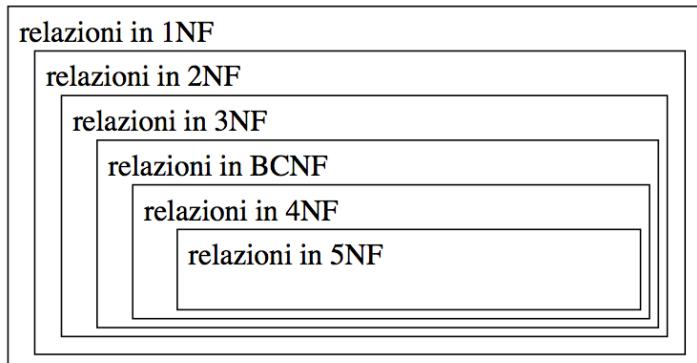
- ▶ Dipendenza funzionale:
  - ▶ è una dipendenza semantica
  - ▶ Data una relazione R, l'attributo Y di R dipende funzionalmente dall'attributo X di R se e solo se ogni valore di X in R ha associato un unico valore di Y in R e si scrive:
  - ▶  $X \rightarrow Y$
- ▶ In altri termini:
  - ▶ Y dipende funzionalmente da X
  - ▶ X è il determinante di Y

Basi di Dati corso A

310

## normalizzazione

- ▶ Insieme di condizioni da rispettare per garantire la qualità di un database, vi sono sei livelli di qualità **maggiormente noti**:



Basi di Dati corso A

311

## normalizzazione

- ▶ Processo ripetuto di applicazione di proiezioni per
- ▶ Eliminare ridondanze
- ▶ Eliminare anomalie

Basi di Dati corso A

312

## 1NF

- ▶ Nel modello relazionale, una relazione è in 1NF se e solo se ogni tupla contiene un solo valore per ogni attributo

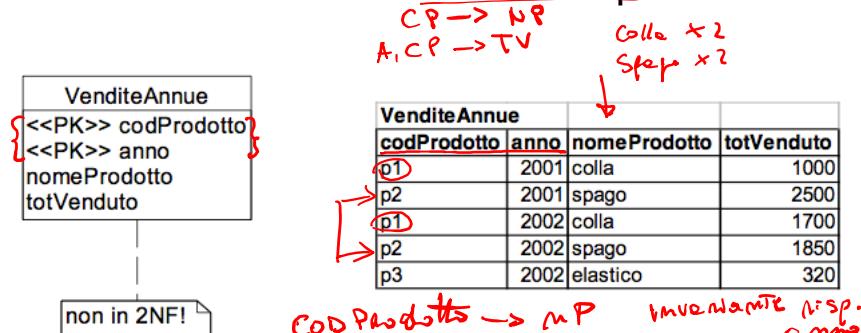
	Attr1	Attr2	Attr3
Tupla 1			
Tupla 2		x	
Tupla 3			
Tupla 4			

Basi di Dati corso A

313

## 2NF

- ▶ Una relazione è in 2NF se è in 1NF e se ogni suo attributo non chiave dipende funzionalmente dalla chiave completa

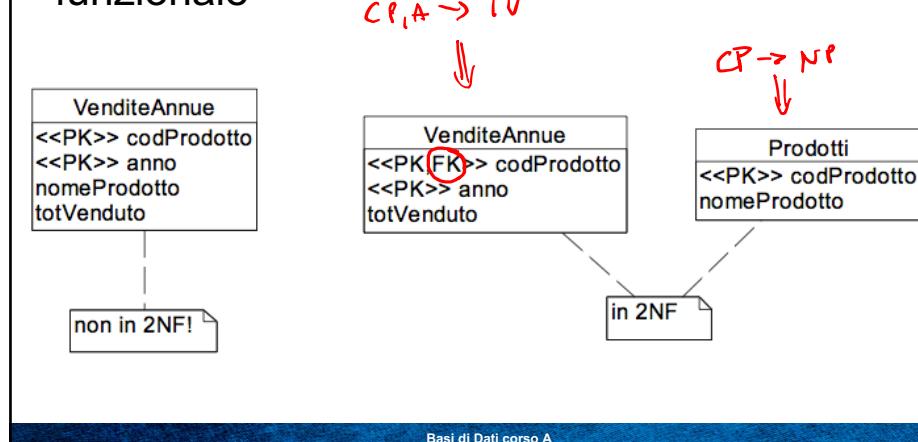


Basi di Dati corso A

314

## 2NF

- ▶ Mediante proiezione si crea una nuova relazione che evidenzia la dipendenza funzionale



Basi di Dati corso A

315

## 3NF

- ▶ Una relazione è in 3NF se è in 2NF e se tra i suoi attributi non chiave non esistono FD transitive

Ordini				
<<PK>> numOrdine				
data				
codCliente				
ragSociale				
importo				

non in 3NF!

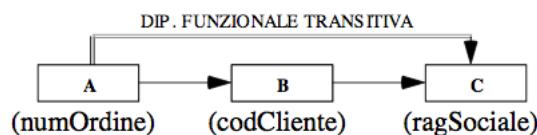
Ordini				
numOrdine	data	codCliente	ragSociale	importo
15	05/04/1999	1	abc	50
234	21/10/2001	2	def	75
567	11/01/2002	1	abc	34
678	31/07/2002	3	ghi	200

Basi di Dati corso A

316

## dip. funzionale transitiva (TFD)

- ▶ In una relazione  $R(A, B, C)$
- ▶ Se  $B \rightarrow C$  e  $A \rightarrow B$
- ▶ Se  $B$  non è chiave candidata di  $R$
- ▶ allora  $C$  ha TFD da  $A$

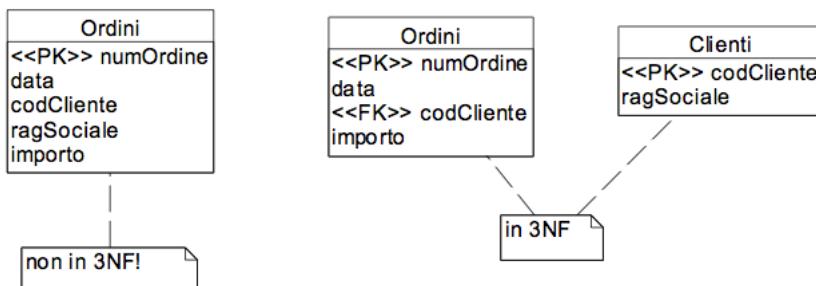


Basi di Dati corso A

317

## 3NF

- ▶ Mediante proiezione si crea una nuova relazione che rimuove la TFD



Basi di Dati corso A

318

## BCNF

- ▶ Una relazione è in BCNF se tutti i suoi determinanti sono candidati chiave
- ▶ Ogni relazione in 3NF è in BCNF a meno che:
  - ▶ abbia 2 o + chiavi candidate
  - ▶ le chiavi candidate siano composte

Basi di Dati corso A

319

## Dipendenza multivalente

Prodotto	Taglia	Colore
T-Shirt	S	Bianco
T-Shirt	S	Giallo
T-Shirt	S	Rosso
T-Shirt	M	Bianco
T-Shirt	M	Giallo
T-Shirt	M	Rosso
T-Shirt	L	Bianco
T-Shirt	L	Giallo
T-Shirt	L	Rosso
T-Shirt	XL	Bianco
T-Shirt	XL	Giallo
T-Shirt	XL	Rosso
Camicia	S	Blu
Camicia	S	Verde
Camicia	M	Blu
Camicia	M	Verde
Camicia	L	Blu
Camicia	L	Verde

- ▶ Se un attributo determina in modo indipendente più attributi
- ▶ prodotto →→ taglia
- ▶ prodotto →→ colore
- ▶ Prodotto ‘multidetermina’ taglia e colore
- ▶ Esiste:
  - ▶ Dipendenza multivalente
  - ▶ ridondanza

Basi di Dati corso A

320

## dipendenza multivalente (MVD)

- ▶ Dati U e V sottoinsiemi di attributi di R, se W è l’insieme di attributi di R non compresi in U e in V, la MVD  $U \rightarrow\rightarrow V$  esiste in R se e solo se
  - ▶  $R = R(UV) \text{ JOIN } R(W)$
- ▶ MVD è generalizzazione di FD
  - ▶ Tutte le FD sono MVD
  - ▶ Non tutte le MVD sono FD
- ▶  $(A \rightarrow B)$  può essere visto come caso speciale di  $A \rightarrow\rightarrow B$

Basi di Dati corso A

321

## MVD

- ▶ Se la MVD è FD non causa ridondanze
  - ▶ In prodotti(cod, qtà, prezzo) esistono le seguenti FD: cod→qtà e cod→prezzo
  - ▶ Si può effettuare una proiezione di Prodotti in: (cod,qtà) e (cod,prezzo) e ricostruire con equijoin
  - ▶ In prodotti sussistono cod→→qtà e cod→→prezzo, che sono anche FD

Basi di Dati corso A

322

## 4NF

- ▶ R è in 4NF se e solo se, se compare una dipendenza multivalente  $A \rightarrow\!\!\! \rightarrow B$ , tutti gli attributi di R dipendono anche funzionalmente da A

Prodotto	Taglia	Colore
T-Shirt	S	Bianco
T-Shirt	S	Giallo
T-Shirt	S	Rosso
T-Shirt	M	Bianco
T-Shirt	M	Giallo
T-Shirt	M	Rosso
T-Shirt	L	Bianco
T-Shirt	L	Giallo
T-Shirt	L	Rosso
T-Shirt	XL	Bianco
T-Shirt	XL	Giallo
T-Shirt	XL	Rosso
Camicia	S	Blu
Camicia	S	Verde
Camicia	M	Blu
Camicia	M	Verde
Camicia	L	Blu
Camicia	L	Verde

non in  
4NF!

Prodotto	Taglia
T-Shirt	S
T-Shirt	M
T-Shirt	L
T-Shirt	XL
Camicia	S
Camicia	M
Camicia	L

Prodotto	Colore
T-Shirt	Bianco
T-Shirt	Giallo
T-Shirt	Rosso
Camicia	Blu
Camicia	Verde

in  
4NF

Basi di Dati corso A

323

## dipendenza di Join (JD)

- ▶ Esistono relazioni in 4NF con anomalie
  - ▶ Un concessionario che vende un articolo e rappresenta un produttore. Se il produttore produce quell'articolo, allora il concessionario vende l'articolo del produttore

Concessionario	Articolo	Produttore
Neri	PC	IBM
Neri	PC	HP
Neri	Scanner	HP
Verdi	PC	HP

Basi di Dati corso A

324

## JD

- ▶ Non si tratta di FD, né di MVD ma di dipendenza di join (JD)
- ▶  $R(X,Y,Z)$  soddisfa la JD se e solo se
  - ▶  $R = \pi(X) \text{ JOIN } \pi(Y) \text{ JOIN } \pi(Z)$
- ▶ JD è la dipendenza più generale
  - ▶ FD caso particolare di MVD
  - ▶ MVD caso particolare di JD

Basi di Dati corso A

325

## 5NF

- R è in 5NF se e solo se ogni dipendenza di join presente in essa è conseguenza delle sue chiavi candidate

ConcessionarioArticoloProduttore		
Concessionario	Articolo	Produttore
Neri	PC	IBM
Neri	PC	HP
Neri	Scanner	HP
Verdi	PC	HP

ConcessionarioArticolo	
Concessionario	Articolo
Neri	PC
Neri	Scanner
Verdi	PC

ConcessionarioProduttore	
Concessionario	Produttore
Neri	IBM
Neri	HP
Verdi	HP

Basi di Dati corso A

326

non in  
5NF!

ArticoloProduttore	
Articolo	Produttore
PC	IBM
PC	HP
Scanner	HP

in  
5NF

## JD e 5NF

- Se la dipendenza di join è conseguenza delle chiavi candidate, la relazione è già in 5NF
  - In dipendenti(matricola,CF, nome) matricola e CF sono chiavi candidate
  - Si può proiettare: (matricola,CF), (CF, nome), (matricola, nome), e ricostruire con join su matricola e CF
  - Dipendenti contiene JD, ma non presenta anomalie => è già in 5NF

Basi di Dati corso A

327