

Corso di Laurea in Ingegneria  
Informatica  
*Fondamenti di informatica II*  
Modulo "*Basi di dati*"  
a.a. 2017-2018

Docente: Gigliola Vaglini  
Docente laboratorio: Francesco  
Pistolesi

1

## Lezione 4

Progettazione di basi di dati:  
metodologie e modelli

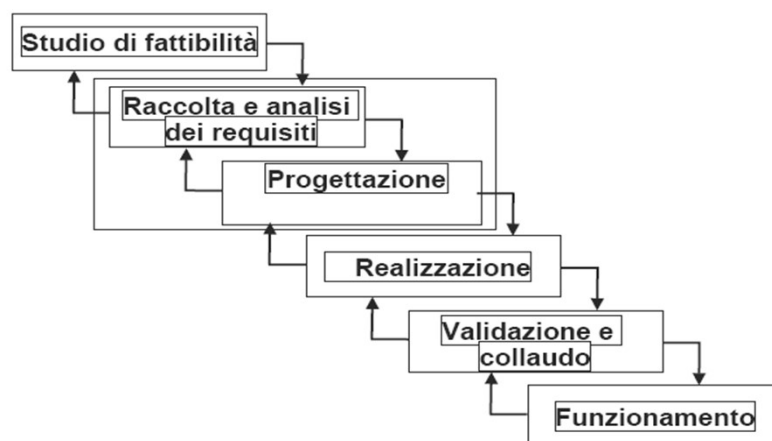
2

## Processo di sviluppo di sistemi software

- Lo sviluppo di sistemi software in generale, e di sistemi informativi in particolare, è un'attività complessa che comprende diverse fasi.
- Ciclo di vita: sequenza di attività, anche ripetute ciclicamente, svolte da analisti, progettisti, utenti, nello sviluppo e nell'uso dei sistemi sw

3

## Modello del ciclo di vita a cascata (Waterfall model)



4

## Fasi del ciclo di vita

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

5

## La progettazione è una fase del ciclo di vita

- Adesso ci focalizziamo sul passo 3 del ciclo di vita più alcuni aspetti di raccolta e analisi dei requisiti.
- La progettazione di un sistema software consta fondamentalmente di due aspetti
  - progettazione dei dati
  - progettazione delle applicazioni
- Nel caso di sistemi informativi il progetto dei dati ha un ruolo centrale

6

## Un buon progetto

- Per garantire prodotti di buona qualità è opportuno seguire una
  - metodologia di progetto, basata su
    - modelli per rappresentare i dati che siano facili da usare
    - decomposizione delle attività in fasi successive
    - strategie e criteri di scelta nei vari passi

7

## Progettare per livelli di astrazione

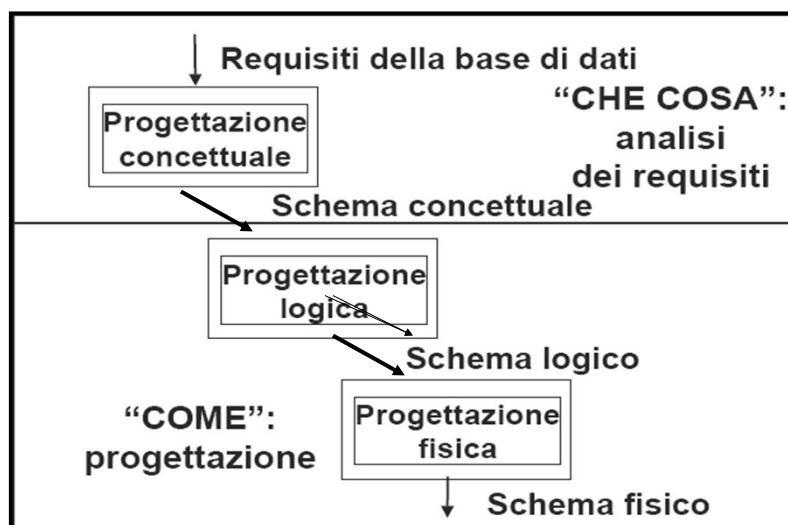
- Basandosi sull'architettura ANSI/SPARC, anche la progettazione procede per livelli:
  - **Livello concettuale.** Esprime i requisiti di un sistema in una descrizione adatta all'analisi dal punto di vista esterno
  - **Livello logico.** Evidenzia l'organizzazione dei dati dal punto di vista del loro contenuto informativo, descrivendo la struttura di ciascun record e i collegamenti tra record diversi.
  - **Livello fisico.** A questo livello la base di dati è vista come un insieme di blocchi fisici su disco. Qui viene decisa l'allocazione dei dati e le modalità di memorizzazione dei dati sul disco.

8

# Schemi di basi di dati

- I prodotti della varie fasi della progettazione sono schemi della base di dati, ognuno basato su un insieme di costrutti per organizzare i dati e descriverne la dinamica
  - Schema concettuale
  - Schema logico
  - Schema fisico

9



10

## Raccolta e analisi dei requisiti

- acquisizione dei requisiti: il reperimento dei requisiti è un'attività difficile e non standardizzabile
- analisi dei requisiti: l'attività di analisi inizia con i primi requisiti raccolti e spesso indirizza verso altre acquisizioni

11

## Come si acquisiscono i requisiti

- direttamente dagli utenti
  - interviste
  - documentazione apposita
- da documentazione esistente:
  - normative (leggi, regolamenti di settore)
  - regolamenti interni, procedure aziendali
  - realizzazioni preesistenti

12

## Interazione con gli utenti

### ■ Problemi

- utenti diversi possono fornire informazioni diverse
- utenti a livello più alto hanno spesso una visione più ampia ma meno dettagliata
- spesso l'acquisizione dei requisiti avviene "per raffinamenti successivi"

13

## Interazione con gli utenti

### ■ Spunti:

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

14

## Interazione con gli utenti tramite documentazione

- Regole generali:
  - standardizzare la struttura delle frasi
  - separare le frasi sui dati da quelle sulle funzioni
  - organizzare termini e concetti
    - costruire un glossario dei termini (individuare i sinonimi)
    - unificare i termini
    - rendere esplicito il riferimento fra termini
    - riorganizzare le frasi per concetti

15

## Un esempio

### **Società di formazione (1)**

**Si vuole realizzare una base di dati per una società che eroga corsi: di ogni corso vogliamo rappresentare i dati dei partecipanti 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 già frequentato (le materie sono in tutto circa 200) e il giudizio finale.**

16



### **Società di formazione (2)**

**Rappresentiamo anche i corsi attualmente attivi 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.**

17

### **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.**

18

## Glossario dei termini

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

19

## Strutturazione dei requisiti in gruppi di frasi omogenee

20

<b>Fraasi di carattere generale</b>
-------------------------------------

<b>Si vuole realizzare una base di dati per una società che eroga corsi: di ogni corso vogliamo rappresentare i dati dei partecipanti e dei docenti.</b>
--

21

<b>Fraasi relative ai partecipanti</b>
--

<b>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.</b>
---

22

<b>Fraasi relative ai datori di lavoro</b>
--

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

<b>Fraasi relative ai corsi</b>
---------------------------------

<b>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.</b>
---

23

<b>Fraasi relative a tipi specifici di partecipanti</b>
---

<b>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.</b>
--

24

### **Fraasi 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.**

25

## **Modello concettuale**

26

## Modelli concettuali

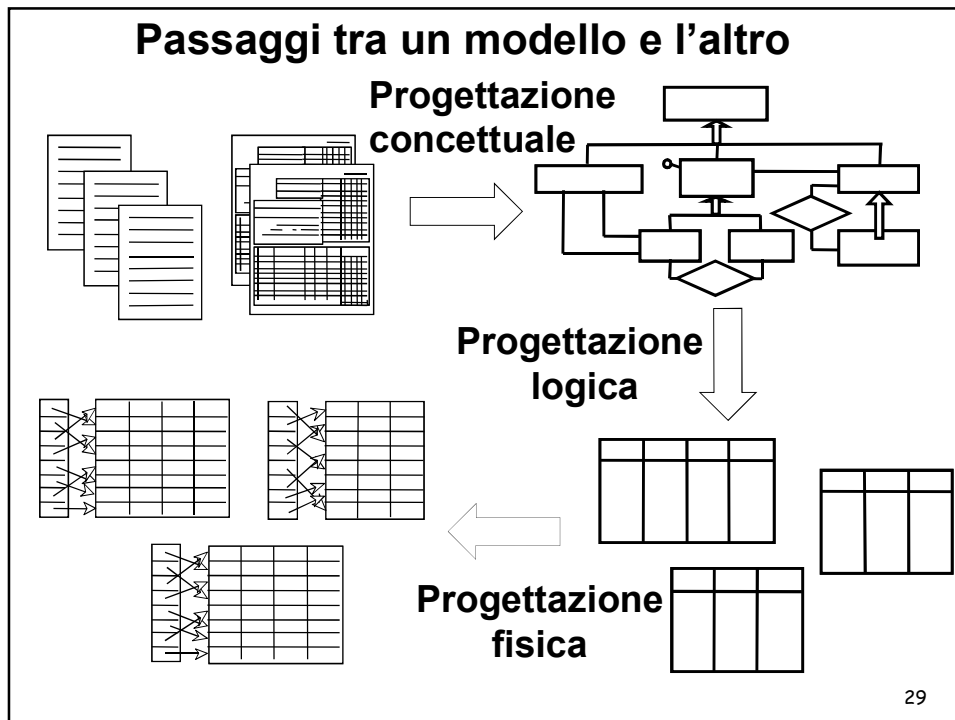
- I dati sono rappresentati in modo più vicino al modo di pensare umano
  - cercano di replicare i concetti del mondo reale
- Prevedono efficaci rappresentazioni grafiche (utili anche per la documentazione e la comunicazione)
  - i modelli successivi sono più rigidi, partendo da loro ci perderemmo nei dettagli

27

## La progettazione concettuale

- Traduce i requisiti di un qualsiasi sistema informatico in una descrizione:
  - **formale**: espressa in modo non ambiguo, ma adeguato a catturare le caratteristiche fondamentali del mondo da descrivere
  - **integrata**: la descrizione si riferisce alla totalità dell'ambiente

28



## Il modello utilizzato

- Il modello E-R (Entity-Relationship, P.P.Chen 1976) si è ormai affermato nelle metodologie di progetto e nei sistemi SW di ausilio alla progettazione (strumenti CASE)

## Costrutti del modello E-R

- Costrutti di base
  - Entità
  - Relationship
  - Attributo
- Altri costrutti
  - Identificatore
  - Generalizzazione
  - ....

31

## Entity

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

32



## Occorrenza di Entità

- **Entità**
  - classe di oggetti, persone, ... "omogenei"
    - es. l'entità "impiegato"
- **Occorrenza (o istanza) di entità**
  - elemento della classe (l'oggetto, la persona, ..., non un valore dei dati legati all'oggetto)
    - es. un "impiegato", non so nulla di lui, ma esiste

33

## Rappresentazione grafica di entità

**Impiegato**

**Dipartimento**

**Città**

**Vendita**

34

## Entità: caratteristiche

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

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

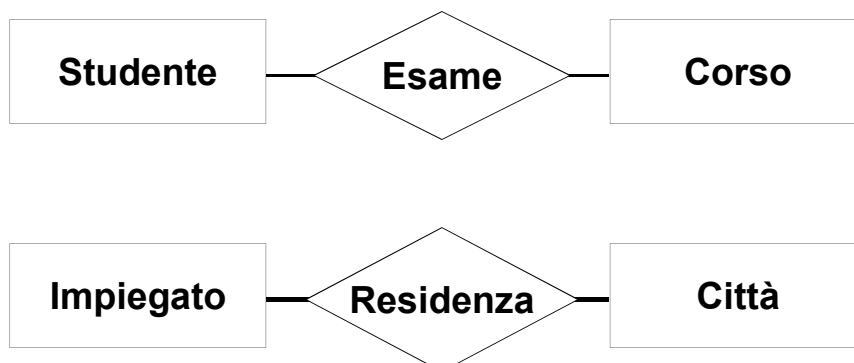
36

## Relationship: occorrenze

- Un'occorrenza di relationship binaria è una coppia di occorrenze di entità, una per ogni entità coinvolta
- Un'occorrenza di relationship n-aria è una n-upla di occorrenze di entità, una per ciascuna entità coinvolta
  - non ci possono essere occorrenze ripetute sulle stesse occorrenze di entità

37

## Rappresentazione grafica di relationship



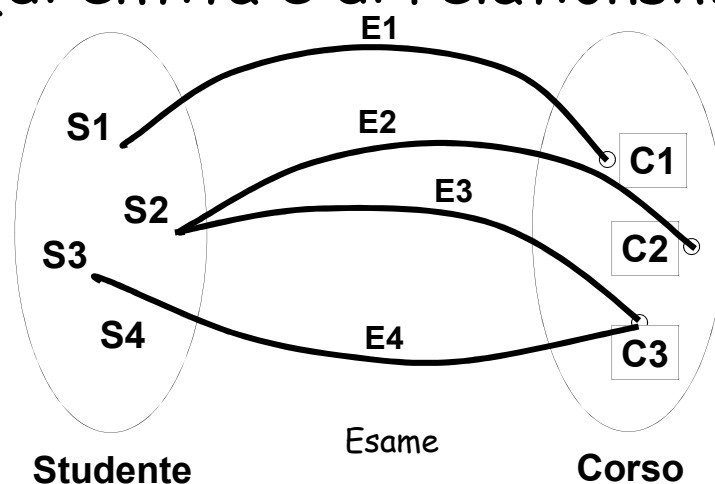
38

## Relationship: caratteristiche

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

39

## Esempi di occorrenze (di entità e di relationship)



40

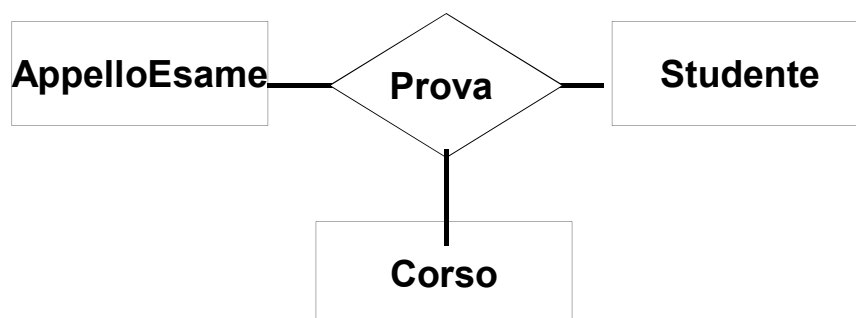
## Relationship



- ✓ Problema.
- ✓ Ma se uno studente sostiene piu' volte lo stesso esame prima di essere promosso, come si vede?
- ✓ Non si può vedere perchè risulterebbero coppie di occorrenze di entità uguali
- ✓ Esame, studente, corso devono diventare tutte entità, messe in relazione dalla relationship Prova, ad esempio

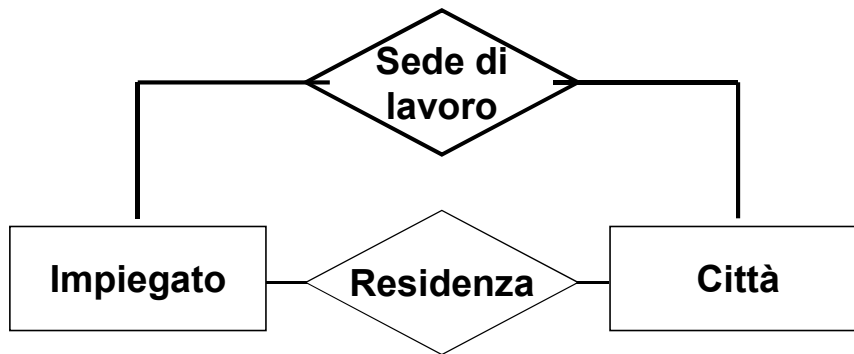
41

## Soluzione: Relationship n-aria



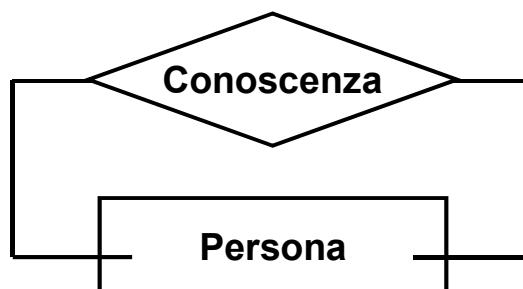
42

## Relationship diverse sulle stesse entità



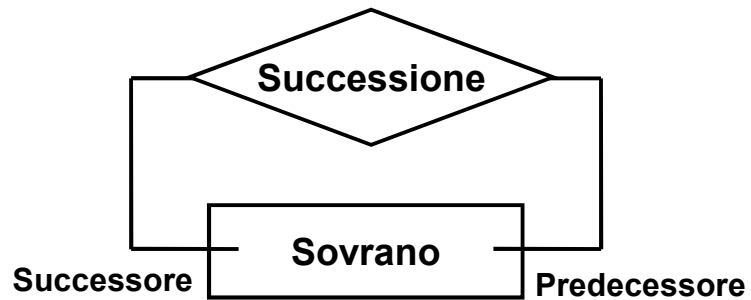
43

## Relationship ricorsiva



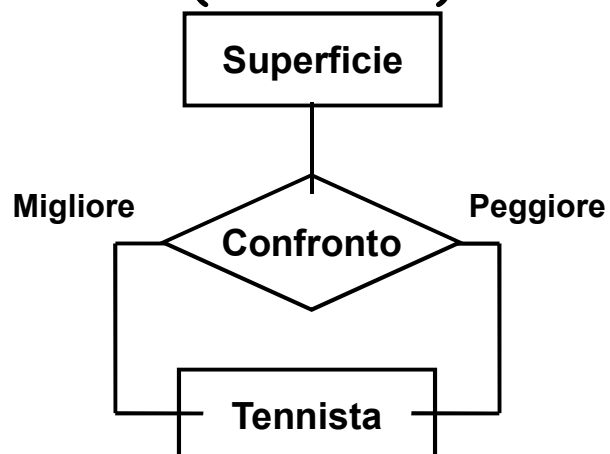
44

## Relationship ricorsiva con "ruoli"



45

## Relationship mista (ternaria)



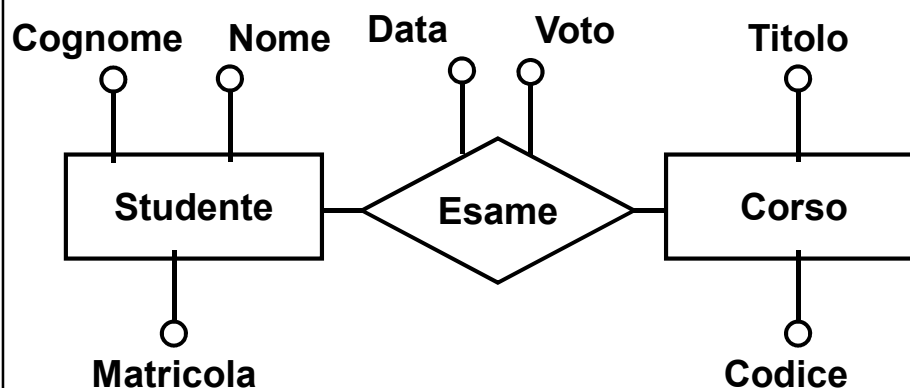
46

## Attributo

- Proprietà elementare di un'entità o di una relationship
- Associa ad ogni occorrenza di entità o relationship un valore appartenente a un insieme detto dominio dell'attributo

47

## Attributi: rappresentazione grafica



48

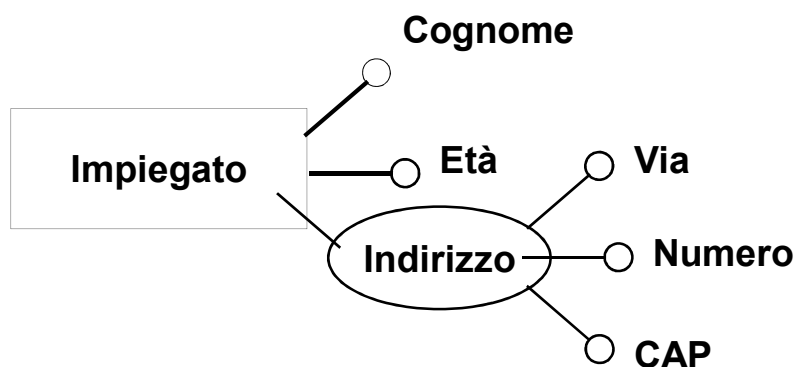


## Attributi composti

- 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

49

## Rappresentazione grafica

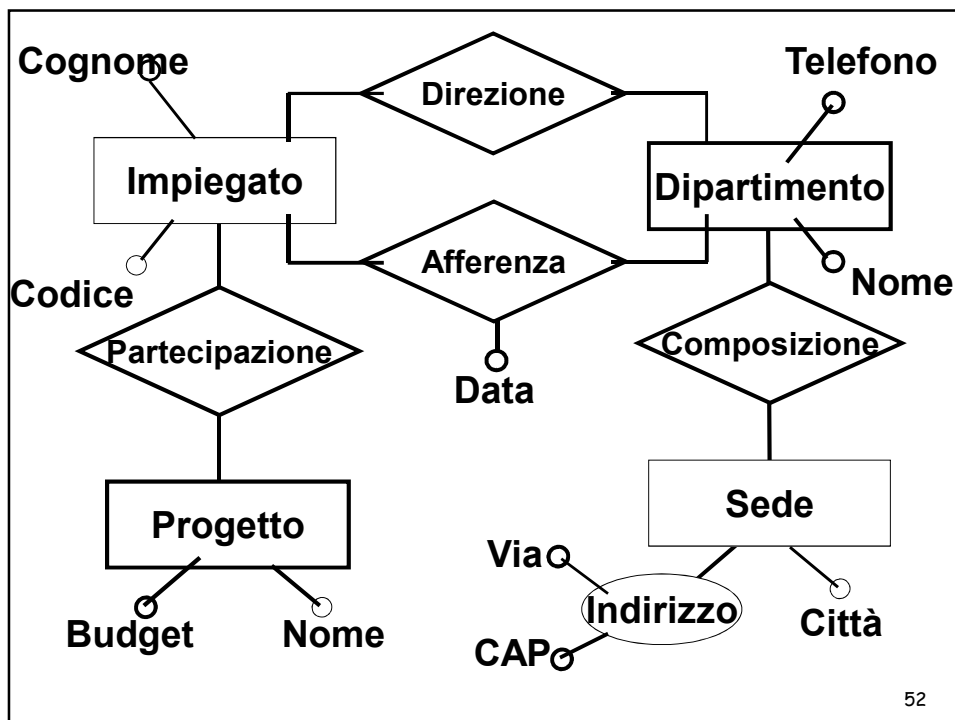


50

## Schema E-R con solo i costrutti base

- Si vuole descrivere l'organizzazione di un'azienda
  - Con sedi diverse
  - Ogni sede è composta di vari dipartimenti
  - Gli impiegati dell'azienda afferiscono ai vari dipartimenti e un'impiegato li dirige
  - Gli impiegati lavorano su progetti
  - Ogni entità o relationship può avere vari attributi.

51



52

## Altri costrutti del modello E-R

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

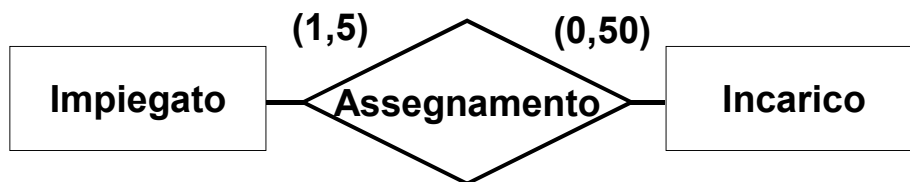
53

## Cardinalità di relationship

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

54

## Esempio di cardinalità

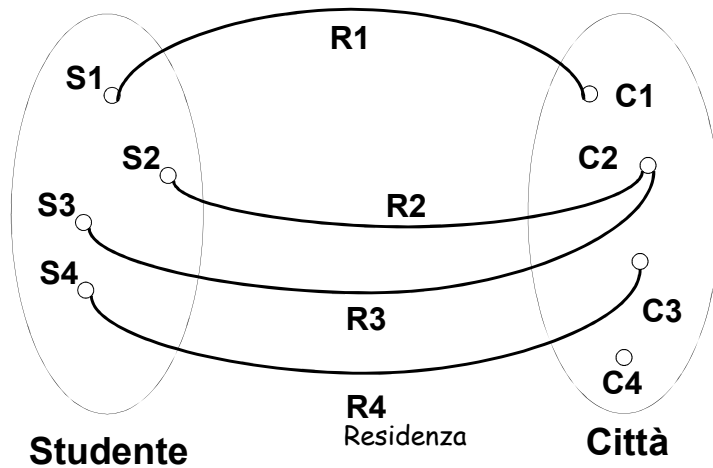


55

- per semplicità usiamo solo tre simboli:
- 0 e 1 per la cardinalità minima:
  - 0 = "partecipazione opzionale"
  - 1 = "partecipazione obbligatoria"
- 1 e "N" per la massima:
  - "N" non pone alcun limite

56

## Occorrenze di Residenza



57

## Cardinalità di Residenza



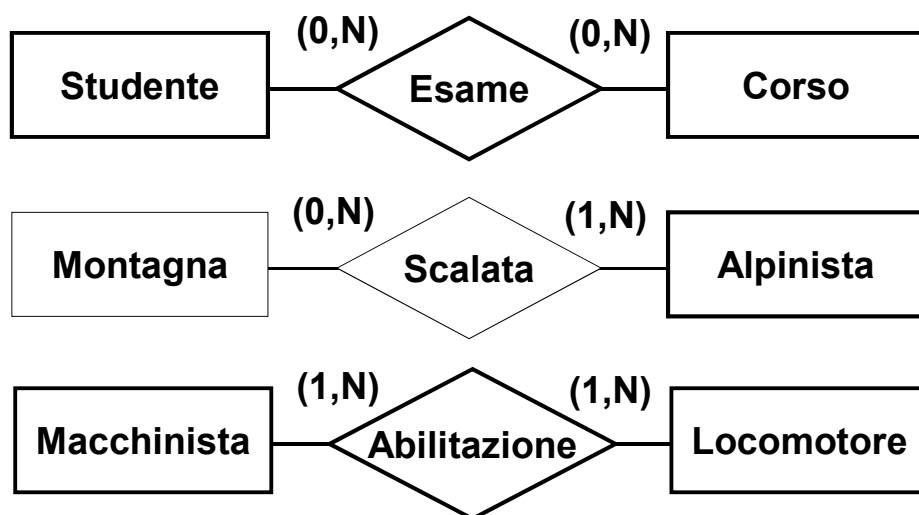
58

## Tipi di relationship

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

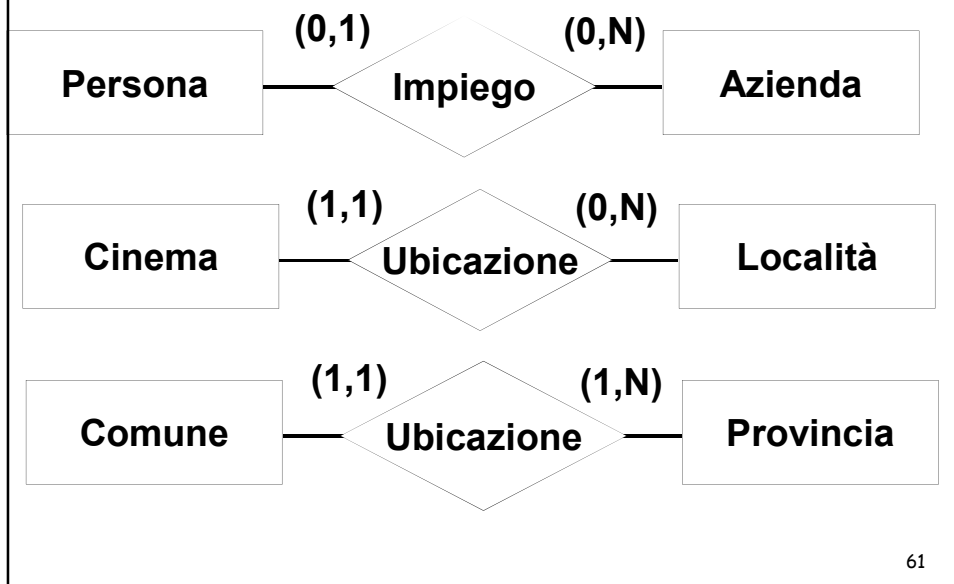
59

## Relationship "molti a molti"

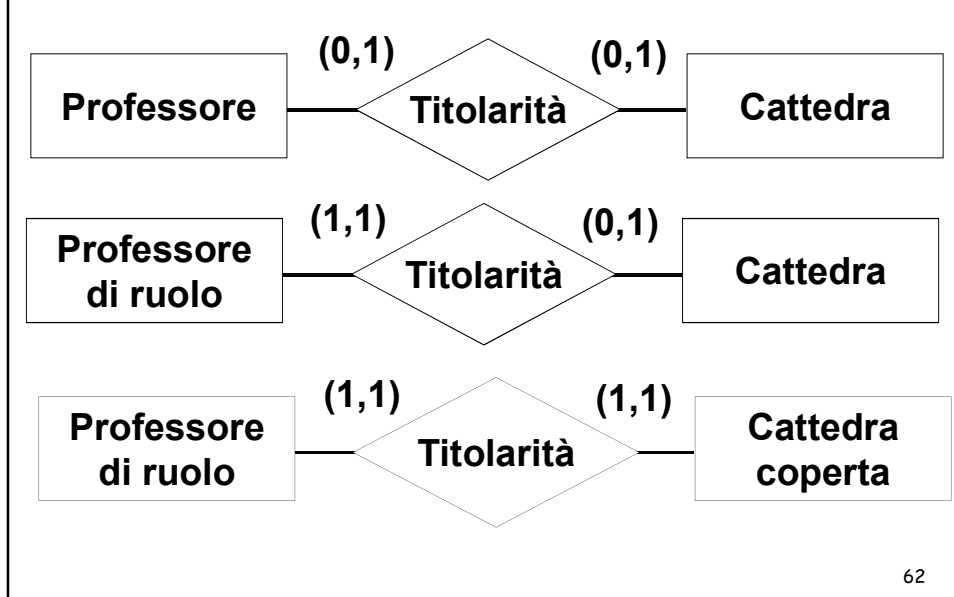


60

## Relationship "uno a molti"



## Relationship "uno a uno"

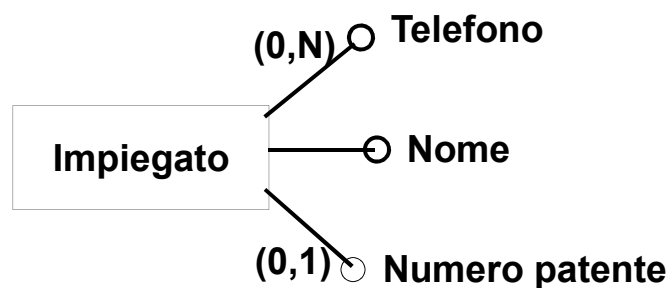


## Cardinalità di attributi

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

63

## Rappresentazione grafica



64

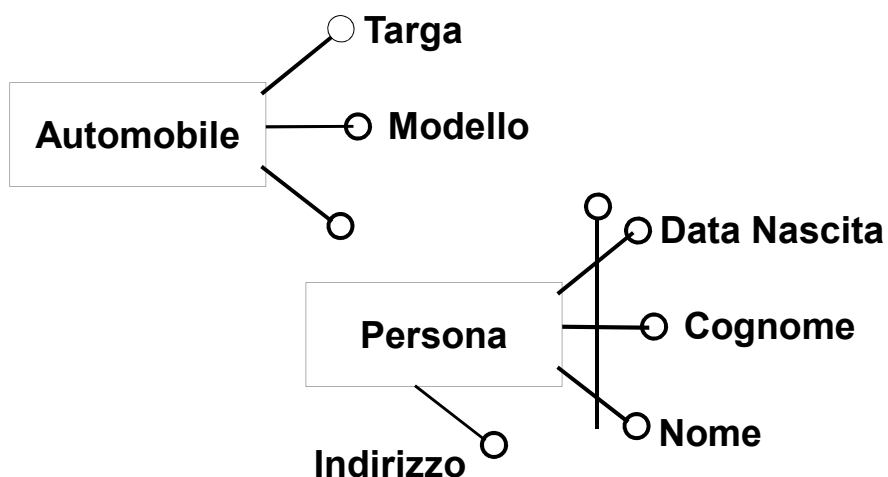


## Identificatore di una entità

- "strumento" per l'identificazione univoca delle occorrenze di un'entità
- costituito da:
  - attributi dell'entità
    - identificatore interno (o chiave)
  - (attributi +) entità esterne attraverso relationship
    - identificatore esterno

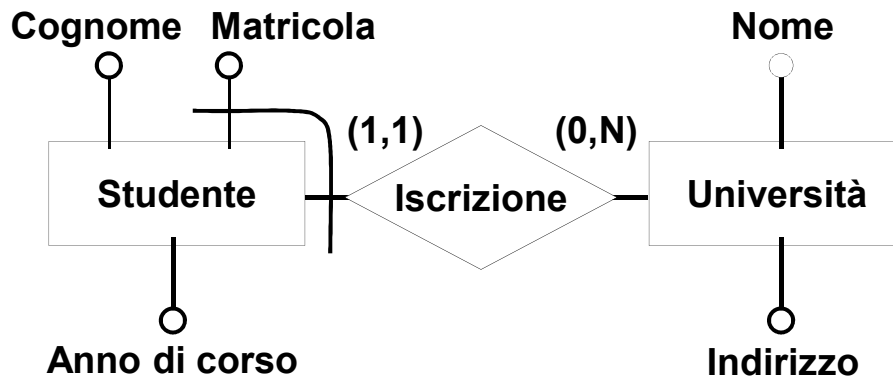
65

## Identificatori interni



66

## Identificatore esterno



67

## Identificatori: caratteristiche

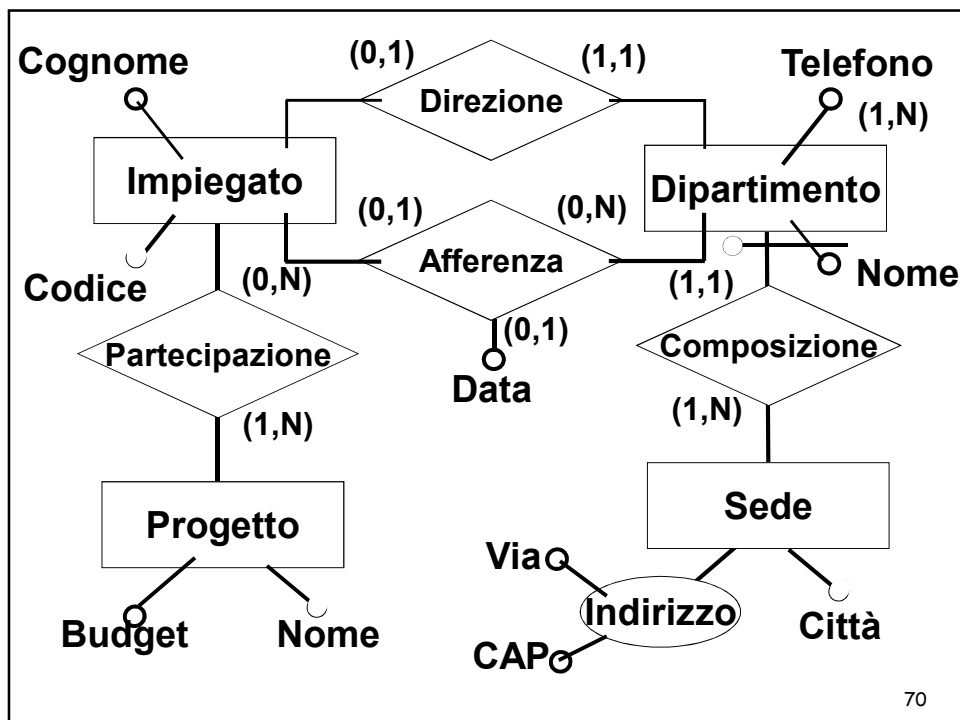
- ogni entità deve possedere almeno un identificatore, ma può averne in generale più di uno
- una identificazione esterna è possibile solo attraverso una relationship a cui l'entità da identificare partecipa con cardinalità (1,1)

68

## Schema E-R con costrutti di base, identificatori e cardinalità

- Riprendiamo l'esempio dell'azienda ed introduciamo la cardinalità per attributi e relazioni, ed individuiamo gli identificatori delle entità coinvolte.

69



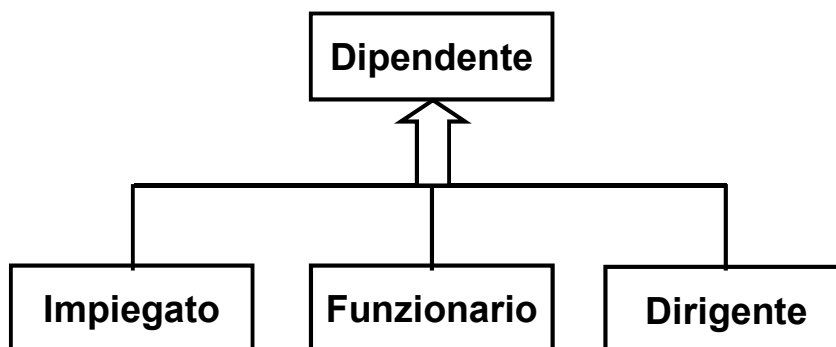
70

## Generalizzazione

- Mette in relazione una o più entità  $E_1, E_2, \dots, E_n$  con una entità  $E$ , che le comprende come casi particolari
  - $E$  è generalizzazione di  $E_1, E_2, \dots, E_n$
  - $E_1, E_2, \dots, E_n$  sono specializzazioni (o sottotipi) di  $E$

71

## Generalizzazione: rappresentazione grafica



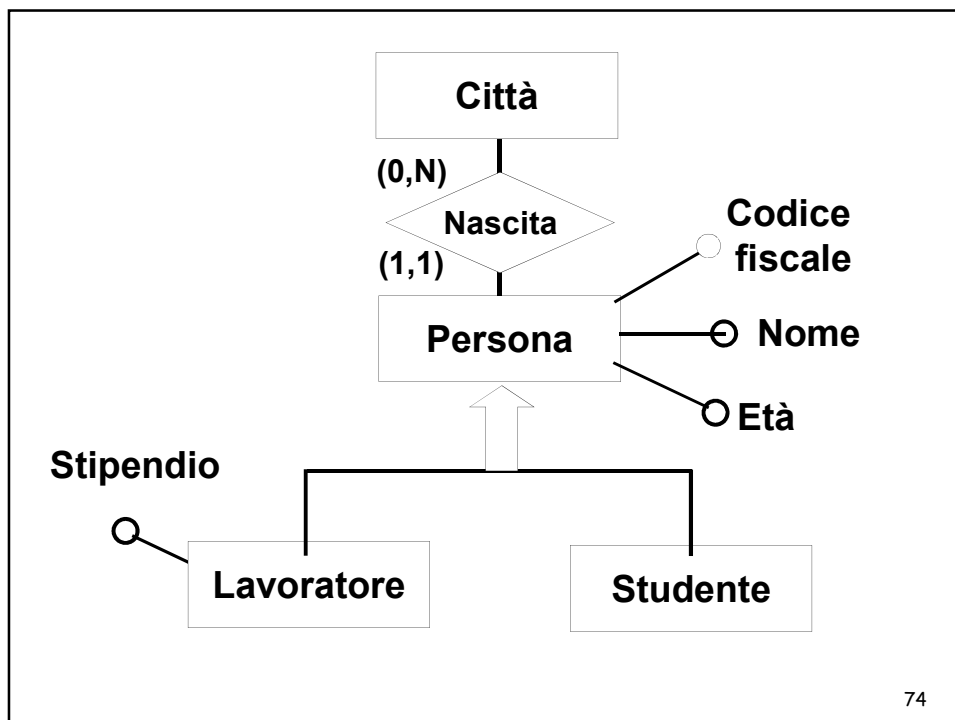
72

## Proprietà delle generalizzazioni

Se E (genitore) è generalizzazione di E1, E2, ..., En (figlie):

- ogni proprietà di E è significativa per E1, E2, ..., En
- ogni occorrenza di E1, E2, ..., En è occorrenza anche di E

73



74

## Ereditarietà

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

75

## Tipi di generalizzazioni

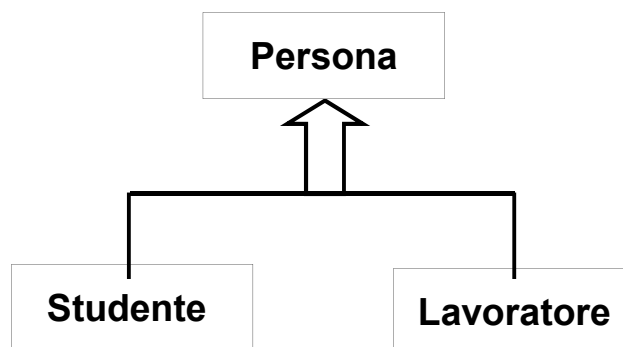
- totale se ogni occorrenza dell'entità genitore è occorrenza di almeno una delle entità figlie, altrimenti è parziale
- esclusiva se ogni occorrenza dell'entità genitore è occorrenza di al più una delle entità figlie, altrimenti è sovrapposta

76

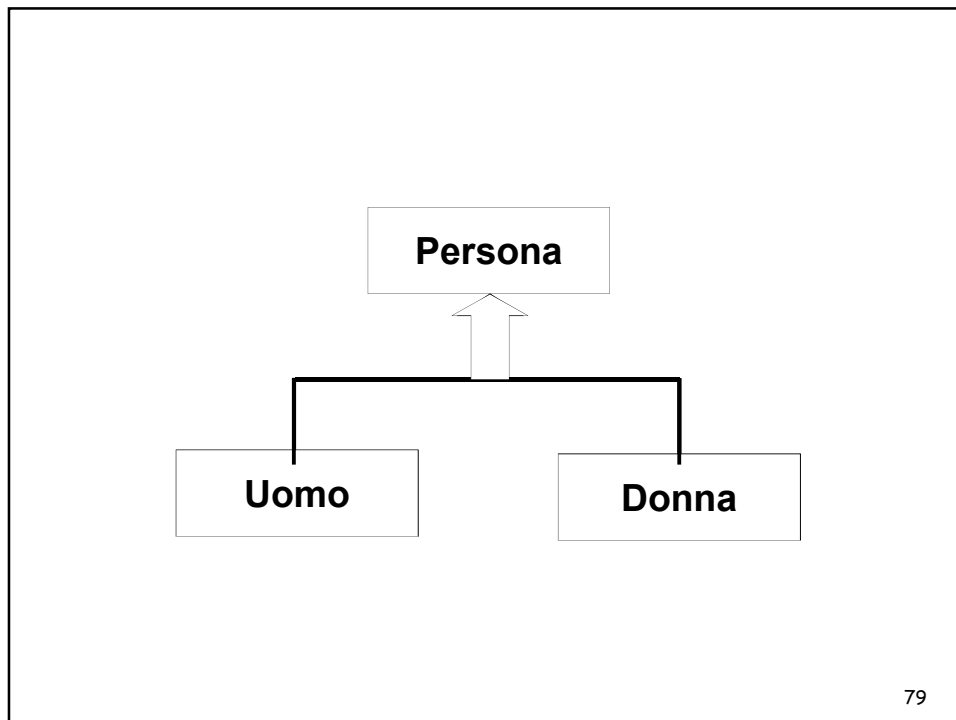
## Tipi di generalizzazioni

- Consideriamo negli schemi solo generalizzazioni esclusive (si può sempre trasformare una generalizzazione sovrapposta in una esclusiva) e distinguiamo fra generalizzazioni parziali e totali

77



78



## 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, di che tipo è questa generalizzazione?
- il genitore di una generalizzazione parziale deve avere un identificatore tra i suoi attributi, quello di una generalizzazione totale no, però? ...

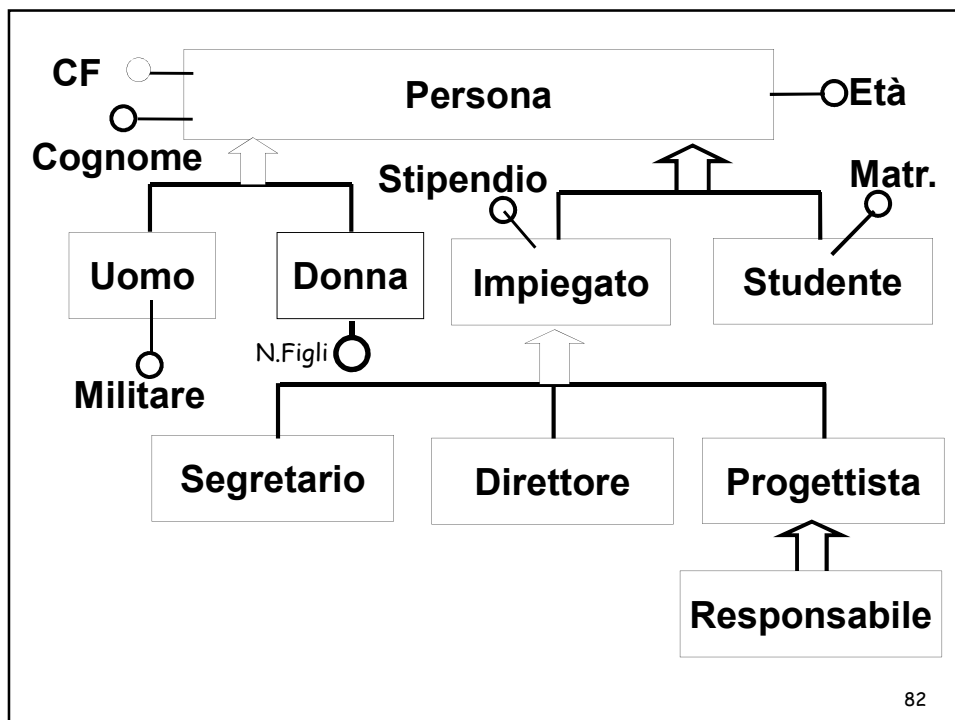
80



## Esercizio

- Le persone hanno CF, cognome ed età, gli uomini la posizione militare, le donne il numero dei figli;
- 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) hanno un numero di matricola;
- esistono persone che non sono né impiegati né studenti (ma i dettagli non ci interessano)

81



82

## Documentazione associata agli schemi concettuali

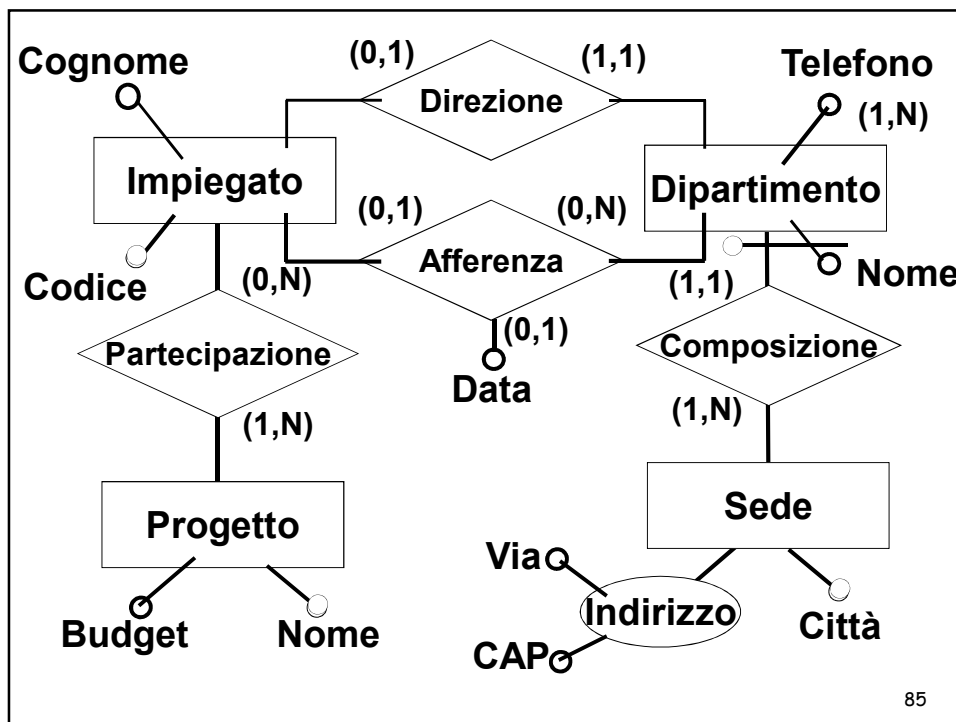
- Uno schema E-R non è quasi mai sufficiente da solo a rappresentare tutti i dettagli di un'applicazione
- Ci sono vincoli non esprimibili
- È necessario associare una documentazione di supporto

83

## Documentazione

- Regole aziendali
  - Vincoli di integrità
  - Possibili derivazioni
- Dizionario dei dati
  - entità
  - relationship

84



85

## Dizionario dei dati (entità)

Entità	Descrizione	Attributi	Identificatore
Impiegato	Dipendente dell'azienda	Codice, Cognome,	Codice
Progetto	Progetti aziendali	Nome, Budget	Nome
Dipartimento	Struttura aziendale	Nome, Telefono	Nome, Sede
Sede	Sede dell'azienda	Città, Indirizzo	Città

86

## 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	

87

## Regole di vincolo

- |  |
|--|
| (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         |

88

## Regole di derivazione

- |  |
|--|
| (1) Il numero di impiegati di un dipartimento si ottiene contando gli impiegati che afferiscono a tale dipartimento    |
| (2) Il budget di un progetto si ottiene moltiplicando per 3 la somma degli stipendi degli impiegati che vi partecipano |

89

## Affrontare il progetto

90

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

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

91

- 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
  - associazione
- se è caso particolare di un altro
  - generalizzazione

92

## Strategie di progetto

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

93

## Strategia top-down

- Si parte da uno schema iniziale che viene successivamente raffinato e integrato per mezzo di primitive che lo trasformano in una serie di schemi intermedi per arrivare allo schema E-R finale

94

## Primitive di raffinamento top-down

- Da entità a associazione tra entità
- Da entità a generalizzazione
- Da associazione a insiemi di associazioni
- Da associazione a entità con associazioni
- Introduzione di attributi su entità e associazioni

95

## Strategia bottom-up

- Si parte dalle specifiche iniziali e si suddividono fino a dare specifica ad una componente minima di cui si dà lo schema E-R
- questi schemi vengono fusi e integrati fino ad ottenere lo schema finale

96



## Primitive di trasformazione Bottom-up

- Generazione di entita'
- Generazione di associazione
- Generazione di generalizzazione

97

## 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

98

## Definizione dello schema scheletro

- 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

99

## Qualità di uno schema concettuale

- correttezza
- completezza
- leggibilità
- minimalità

100

## Archivio fotografico

Si vuole rappresentare la base di dati di un archivio fotografico distribuito in varie sedi. Le fotografie sono catalogate in base ad un catalogo di soggetti possibili, ciascun soggetto ha una propria chiave. Le foto hanno una dimensione ed uno stato di conservazione; per le foto a colori, è noto il tipo di stampa (chiaro o opaco). Le foto sono reperibili in archivi, di cui è noto il responsabile, l'indirizzo, il numero telefonico e l'orario di apertura.

Le foto possono descrivere personaggi, luoghi o oggetti. I personaggi hanno un nome ed un sesso; alcuni sono deceduti. Per i personaggi politici, si indica il partito di appartenenza e l'eventuale carica governativa ricoperta. Per gli artisti, si indica la loro attività prevalente (pittura, scultura, ...). Quando le foto descrivono opere artistiche, è noto il nome dell'opera d'arte, l'artista che l'ha realizzata, il luogo dove l'opera risiede e l'anno di realizzazione. Quando le foto descrivono luoghi o oggetti, è noto nome e descrizione.

101

## Ristrutturazione delle frasi

- Si vuole rappresentare la base di dati di un archivio fotografico distribuito in varie sedi.
- Le foto sono reperibili in archivi, di cui è noto il responsabile, l'indirizzo, il numero telefonico e l'orario di apertura.

102

- Le fotografie sono catalogate in base ad un catalogo di soggetti possibili.
- Le foto hanno una dimensione ed uno stato di conservazione; per le foto a colori, è noto il tipo di stampa (chiaro o opaco).

103

- Le foto possono descrivere personaggi, luoghi o oggetti, ciascun soggetto ha una propria chiave.
- I personaggi hanno un nome ed un sesso; alcuni sono deceduti. Per i personaggi politici, si indica il partito di appartenenza e l'eventuale carica governativa ricoperta. Per gli artisti, si indica la loro attività prevalente (pittura, scultura, ...). Quando le foto descrivono opere artistiche, è noto il nome dell'opera d'arte, l'artista che l'ha realizzata, il luogo dove l'opera risiede e l'anno di realizzazione. Quando le foto descrivono luoghi o oggetti, è noto nome e descrizione.

104

## Definizione dello scheletro

- Da una prima lettura del testo individuiamo che le entità fondamentali sono
  - Le fotografie
  - Gli archivi
  - I soggetti
- Queste entità sono in relazione fra loro ed è facilmente individuabile il seguente scheletro di base



- Possiamo ora ad esaminare i singoli elementi dello scheletro...

105

## Analisi dello scheletro (1)

### □ Archivio

- Dal testo si ricavano facilmente le caratteristiche dell'entità



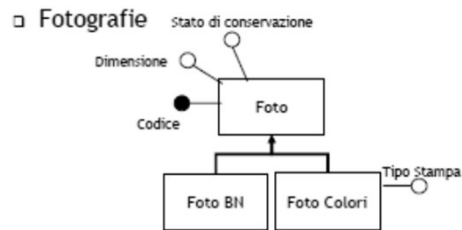
- Per identificatore è stato aggiunto un campo univoco "Codice"

### □ Fotografie

- Dal testo si nota che esistono due tipologie di fotografie: quelle a colori e le altre (in bianco e nero). Si ha cioè una gerarchia totale ed esclusiva.

106

## Analisi dello scheletro (2)



□ Soggetto

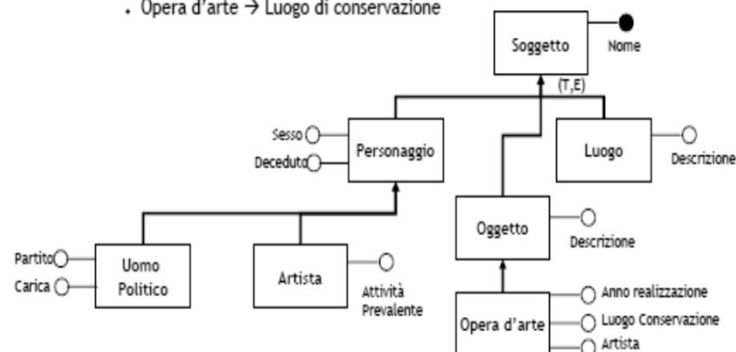
- Dal testo si ricava una gerarchia abbastanza complessa
  - NOTA1: fare bene attenzione a dove si mettono gli attributi
    - Fare attenzione all'attributo nome: ricordare che i figli ereditano dai padri
  - NOTA2: mettere bene in evidenza il tipo di gerarchia (Totale, Esclusiva,...)

107

## Analisi dello scheletro (3)

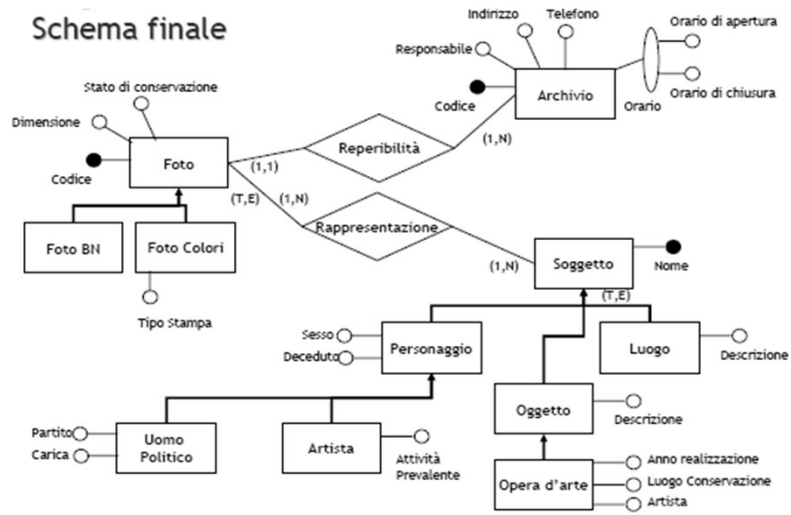
□ Soggetto

- Approfondimento: il testo non lo esplicitava, ma volendo era possibile aggiungere alcune associazioni fra le tipologie di foto
  - Opera d'arte → Artista che l'ha realizzata
  - Opera d'arte → Luogo di conservazione



108

# Schema finale



109