



Conorzio per la formazione e la ricerca in Ingegneria dell'Informazione

Progettare una base di dati

Sistematizzare è meglio che improvvisare

Docente:
 Cesare Colombo
 CEFRIEL
 colombo@cefriel.it
<http://www.cefriel.it>

Basi di Dati



Il ciclo di vita di un Sistema Informativo



- **Studio di fattibilità:** definire i costi delle possibili soluzioni e definire le priorità di realizzazione
- **Raccolta e analisi dei requisiti:** individuare e studiare le funzionalità che il sistema dovrà implementare
- **Progettazione**
- **Implementazione:** realizzare il sistema informativo
- **Validazione e collaudo:** verificare il corretto funzionamento del sistema
- **Operazione e manutenzione** del sistema

- **Progettazione**
 - ▶ **Dei dati:** individua la struttura e l'organizzazione dei dati
 - ▶ Delle applicazioni: si definiscono le caratteristiche dei programmi
- **NOTE**
 - ▶ Il flusso di progettazione non è quasi mai sequenziale, ma ciclico
 - ▶ Noi ci occuperemo della progettazione dei dati

Progettare una base di dati

- 2 -

Basi di Dati



Progettare una base di dati



- Lo scopo della progettazione è quello di definire lo schema della base di dati e quindi di rappresentare in forma strutturata una serie di informazioni provenienti da sorgenti diverse
 - ▶ Il problema è passare da un insieme di informazioni a una collezione strutturata di dati secondo un certo modello logico
- E' chiaro che questo passo è fondamentale, errori a questo livello possono compromettere l'uso della base di dati
 - ▶ Non rappresento tutti le informazioni interessanti
 - ▶ Non posso compiere alcune elaborazioni
 - ▶ ...



Progetto (1)



- **Progettazione Concettuale**
 - ▶ Ha lo scopo di rappresentare le specifiche informali della realtà 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 (non settoriale)
 - **Indipendente** dalla realizzazione fisica
 - **Leggibile**: è possibile usare questa rappresentazione per comunicare con l'utente finale
 - ▶ Si concentra sul **COSA** rappresentare in una base di dati
 - ▶ Usa un **modello concettuale** per produrre uno **schema concettuale**



Progetto (2)



- **Progettazione Logica**
 - ▶ Descrive i dati secondo una rappresentazione indipendente da dettagli fisici, ma concreta perché disponibile nei DBMS
 - Esempio: il modello relazionale richiederà di formalizzare tutto in tabelle
 - ▶ Traduce lo schema concettuale in uno **schema logico** secondo il **modello logico** prescelto (**COME**)
- **Progettazione fisica**
 - ▶ Lo schema logico viene completato con i parametri fisici di memorizzazione dei dati
 - Organizzazione dei file e degli indici
 - ▶ Strettamente dipendente dal particolare DBMS utilizzato
 - ▶ Il prodotto di questa fase è lo **schema fisico** e fa riferimento ad un **modello fisico** dei dati

Progettare una base di dati
- 5 -
Basi di Dati



Progetto (3)



1. **Proget. Concettuale**
2. **Proget. Logica**
3. **Proget. Fisica**

• Fatture

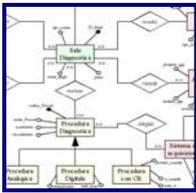
• Ordini dei clienti

• Indirizzi dei clienti

• Archivio prodotti in magazzino

Requisiti

1



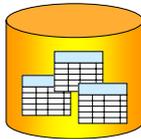
Schema concettuale

2

Fatture
Ordini
Clienti
Magazzino

Schema logico

3



Schema fisico

Progettare una base di dati
- 6 -
Basi di Dati



Astrazione



- L'obiettivo del progettista è quello di rappresentare la conoscenza
- E' necessario operare una **ASTRAZIONE**, ovvero un processo mentale che:
 - ▶ **evidenzia** alcune proprietà **rilevanti** ai fini dell'applicazione
 - ▶ **esclude** le proprietà **irrilevanti** ai fini dell'applicazione
- Esempio: dato un insieme di oggetti capirne le caratteristiche comuni a prescindere dalle loro diversità evidenti
- Esistono tre tipi fondamentali di astrazione: classificazione, associazione e generalizzazione



Classificazione



- **Classificazione: è membro di...**
 - ▶ E' la capacità di costruire un concetto in termini di classe
 - es. {gennaio, febbraio, marzo...} = "mese"
 - ▶ La classe è tanto meglio definita quanto più è definita la regola di appartenenza ad essa
 - ▶ Un oggetto può essere membro di più classi
 - Es. Giovanni e' membro di **Uomo**, Giovanni e' membro di **Persona**, Giovanni e' membro di **Impiegato**
 - ▶ E' una relazione tra oggetti e classi
 - ▶ Lo scopo del progettista è pertanto quello di scegliere le classi che meglio rappresentano il mondo reale
 - Deve saper **CLASSIFICARE**



Associazione



- ◆ **Associazione: è associato a...**
 - ▶ Una classe viene **associata** ad **altre classi** (quindi la costruzione di concetti complessi a partire da concetti elementari)
 - ▶ Es. **Studenti** e **Corsi** possono essere associati dal concetto di **Esame**
 - ▶ Le stesse classi possono essere associate in più modi esprimendo concetti differenti
 - ◆ **Esame**: associa Studente e Corso e indica che un certo studente ha fatto un certo esame **oppure**
 - ◆ **Iscrizione**: associa Studente e Corso e indica che un certo studente frequenta un certo corso
 - ▶ E' la scoperta di nuovi concetti ottenuti mettendo in relazione concetti dati
 - ▶ Lo scopo del progettista è di saper **ASSOCIARE** opportunamente le classi del mondo in esame



Generalizzazione



- ◆ **Generalizzazione: e' sottoinsieme di...**
 - ▶ Definisce una relazione di **sottoinsieme** tra classi: una classe puo' essere generalizzazione di **una o piu' classi**
 - ▶ Es. "studente" è sottoinsieme di "persona". "persona" è una generalizzazione di "studente"
 - ▶ Ereditarietà: le **ASTRAZIONI (PROPRIETA')** definite per la superclasse vengono ereditate dalle sottoclassi
- ◆ **Esempio**
 - ▶ **Persona**: nome, cognome, codice fiscale, indirizzo, data di nascita, n. telefono,...
 - ▶ **Studente**: **si aggiungono** numero di matricola, corso di laurea, esami sostenuti, voti, anno di corso...
 - ▶ **Paziente**: **si aggiungono** gruppo sanguigno, malattie avute, risultati esami medici,...



Progettazione Concettuale



- E' un passo fondamentale perché dà il via alla rappresentazione dell'informazione del mondo reale
- Normalmente sono utilizzati modelli che consentono di descrivere la realtà di interesse in modo facile da comprendere e a prescindere da come i dati saranno rappresentati sul computer (ricorda che l'obiettivo è il **COSA**)
- Il modello più utilizzato è detto Modello E-R (Entity-Relationship)
 - ▶ Il modello E-R usa le tre tipologie di astrazione esaminate
 - ▶ E' stato inventato da Peter Chen nel 1976



Modello E-R



- Il modello E-R fornisce **costrutti** che consentono di rappresentare l'organizzazione e la struttura dei dati
- I costrutti principali del modello sono
 - ▶ Entità
 - ▶ Relazioni
 - ▶ Attributi
- Per ogni costrutto esiste una rappresentazione grafica in modo che il modello possa essere definito tramite un diagramma



Entità



- Rappresentano classi di oggetti (fatti, persone,...) che hanno proprietà comuni ed esistenza “autonoma” ai fini dell’applicazione di interesse
 - ▶ Esempi: Città, Dipartimento, Impiegato

Città

Dipartimento

- Un occorrenza di un’entità è un oggetto della classe che l’entità rappresenta
 - ▶ Esempi: Milano, Elettronica, Mario Rossi
- Il concetto di entità si basa sull’astrazione di **classificazione**

Progettare una base di dati
- 13 -
Basi di Dati



Relazioni (o associazioni) (1)



- Rappresentano legami logici tra due o più entità (significativi per l’applicazione)
 - ▶ Esempio: **Esame** è una relazione fra **Studente** e **Corso**
- Una occorrenza di relazione è un abbinamento fra occorrenze di entità
 - ▶ Esempio: **Analisi I** e **Mario Rossi** è una occorrenza di **Esame**
- Si rappresenta con un rombo, con il nome all’interno, che unisce le entità interessate
 - ▶ Normalmente per il nome si usa un sostantivo invece che un verbo (per evitare di considerare un “verso”)
- La relazione si basa sull’astrazione di **associazione**



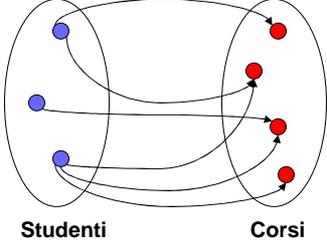
Progettare una base di dati
- 14 -
Basi di Dati



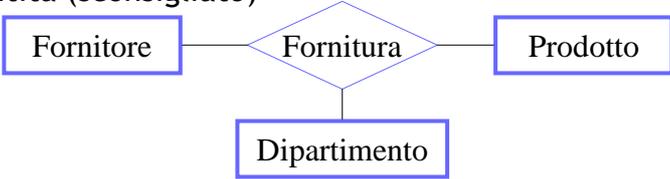
Relazioni (o associazioni) (2)



- Esempio di occorrenze della relazione esame
 - ▶ Come rappresento il fatto che una persona può sostenere più volte lo stesso esame?
 - Anche l'esame diventa una entità...



- Le relazioni possono coinvolgere anche più di due entità (sconsigliato)



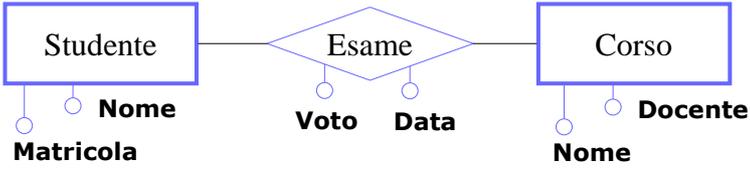
Progettare una base di dati
- 15 -
Basi di Dati



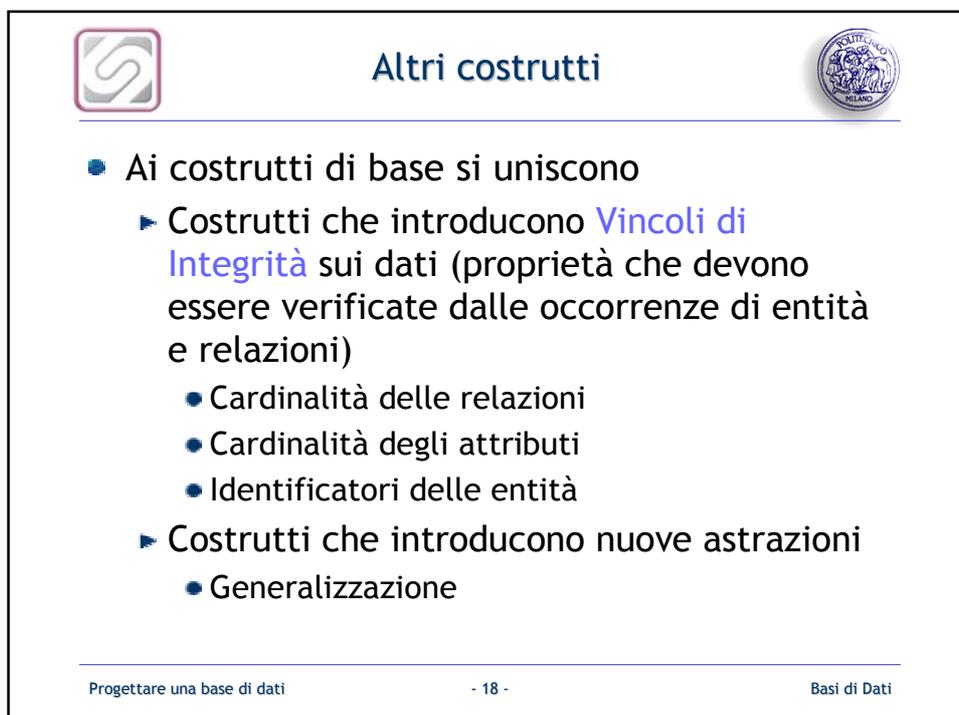
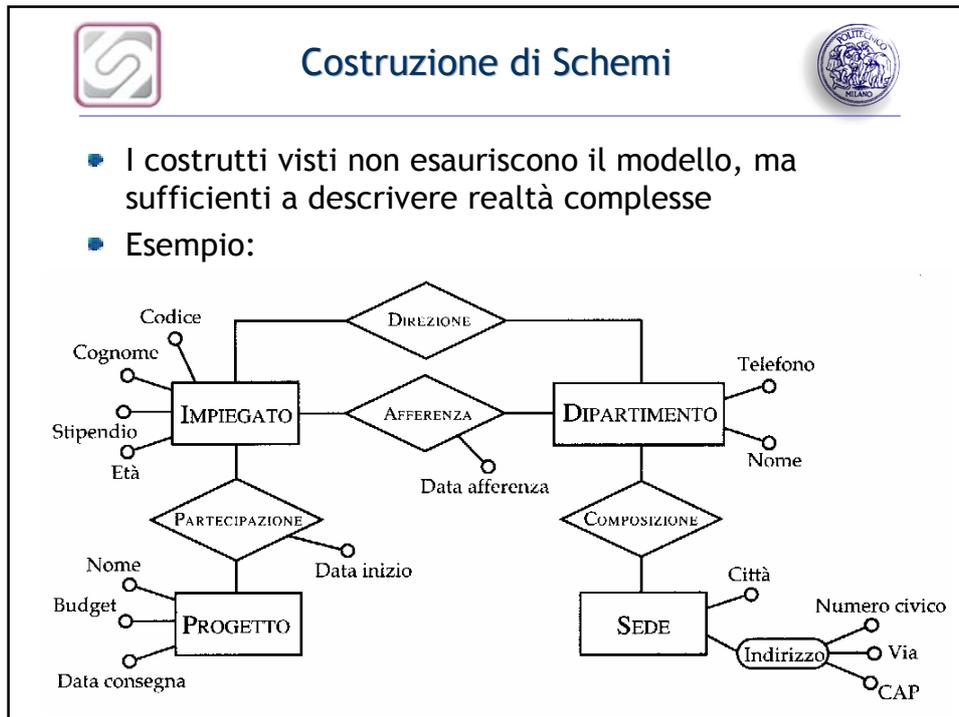
Attributi



- Descrivono proprietà elementari di entità o relazioni (significativi per l'applicazione)
 - ▶ Esempio: **Data** e **Voto Esame** sono attributi per la relazione **Esame**
- Un attributo associa a ciascuna occorrenza di entità un valore scelto in un opportuno dominio
 - ▶ Esempio: un **numero intero** fra 0 e 30 per **Voto Esame**



Progettare una base di dati
- 16 -
Basi di Dati





Cardinalità delle relazioni (1)



- Viene specificata per ogni entità che partecipa ad una relazione e dice quante volte una occorrenza di una entità può essere legata a occorrenze di altre entità in una relazione
 - ▶ Viene indicato il numero minimo (0 o 1) e il numero massimo (1 o N)
 - ▶ Esempio: Città - Residenza - Persona
 - 1 città può essere residenza di un numero di persone da 0 a N
 - 1 persona può risiedere in 1 sola città

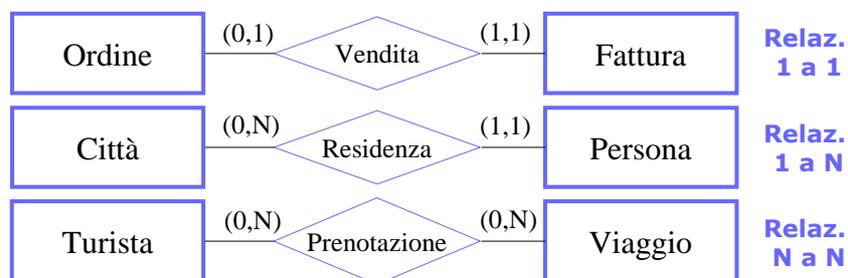


Cardinalità delle relazioni (2)



Le relazioni binarie possono essere classificate in base alla cardinalità massima

- ▶ 1:1 → Relazione uno a uno
- ▶ 1:N o N:1 → Relazione uno a molti
- ▶ N:N → Relazione molti a molti

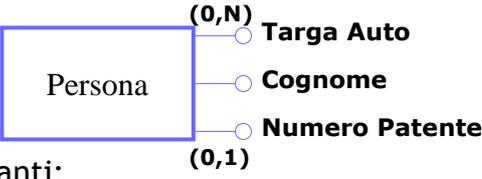




Cardinalità degli attributi



- Descrive il numero minimo e massimo di valori dell'attributo associati a ogni occorrenza di entità o relazione
 - ▶ Nella maggior parte dei casi è (1,1) e viene omessa
 - ▶ Esempio


- Casi interessanti:
 - ▶ Cardinalità minima 0: attributo opzionale
 - ▶ Cardinalità minima 1: attributo obbligatorio
 - ▶ Cardinalità massima N: attributo multivalore
 - Attenzione: spesso possono essere visti come entità

Progettare una base di dati
- 21 -
Basi di Dati



Identificatori delle Entità (1)



- Sono specificati per ciascuna entità di uno schema e descrivono i concetti (attributi e/o entità) che permettono di identificare in maniera univoca le occorrenze delle entità
 - ▶ Se l'identificatore è costituito da soli attributi è detto identificatore *interno* o CHIAVE
 - ▶ Se è costituito anche da entità è detto identificatore *esterno*



Progettare una base di dati
- 22 -
Basi di Dati



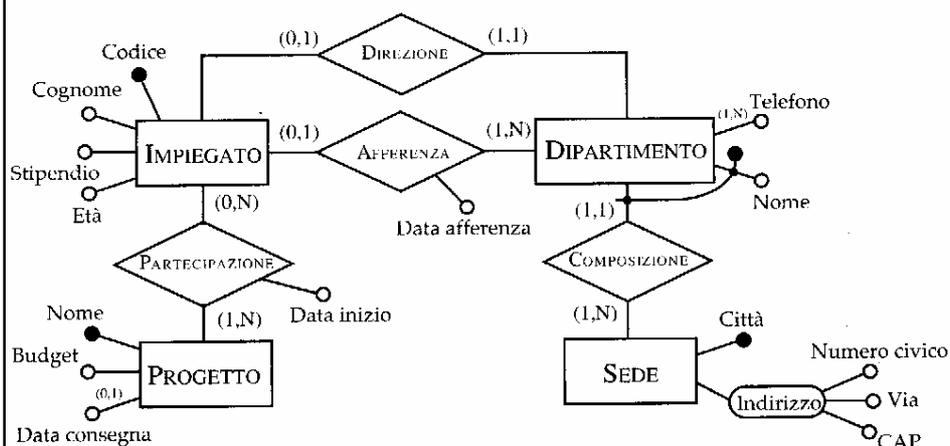
Identificatori delle Entità (2)



- ▶ Regole generali
 - ▶ Un identificatore può coinvolgere uno o più attributi ognuno con cardinalità (1,1)
 - ▶ Un'identificatore esterno può coinvolgere una o più entità ognuna membro di una relazione alla quale l'entità in esame partecipa con cardinalità (1,1)
 - ▶ Una identificazione esterna può coinvolgere entità che a sua volta sono identificate esternamente purché non vengano generati cicli
 - ▶ Ogni entità deve avere almeno un identificatore



Esempio di Schema





Generalizzazione (1)



- Spesso nella analisi di una particolare realtà può risultare che più entità risultino simili o casi particolari l'una dell'altra, derivanti da "viste" diverse da parte dell'utenza
 - ▶ Es. *Ingegnere*, *Medico* e *Avvocato* possono essere differenti "viste" dell'entità *Professionista*
 - Ricorda: è una relazione "è sottoinsieme di"
- Emerge quindi la necessità di evidenziare sottoclassi di alcune classi
- Si definisce pertanto **generalizzazione** il legame logico che esiste tra classi e sottoclassi
- Ovviamente si tratta di una implementazione dell'astrazione di **generalizzazione**



Generalizzazione (2)



- La generalizzazione è il legame logico tra un'entità padre E ed alcune entità figlie E1, E2, ... En dove:
 - ▶ E è la generalizzazione di E1, E2, ... En
 - ▶ E1 E2 ... En sono specializzazioni di E
- Esempi
 - ▶ Persona è una generalizzazione di Uomo e Donna
 - ▶ Professionista è una generalizzazione di Ingegnere, medico e avvocato
- Valgono le seguenti proprietà generali
 - ▶ Una occorrenza di una entità figlia è anche occorrenza della entità padre
 - Una occorrenza di Medico è anche occorrenza di Professionista
 - ▶ Ogni proprietà dell'entità padre è anche proprietà delle entità figlie (EREDITARIETA')
 - Se Persona ha attributo Nome, lo hanno anche Uomo e Donna

Tipi di Generalizzazione - 1

- Due tipi di generalizzazioni
 - Una generalizzazione è **TOTALE** se ogni occorrenza della classe padre è occorrenza di almeno un figlio (altrimenti è **PARZIALE**)
 - Una generalizzazione è **ESCLUSIVA** se ogni occorrenza della classe padre è al più occorrenza di un figlio (altrimenti è **SOVRAPPONSTA**)
- Esempio
Un'azienda si avvale dell'opera di professionisti esterni, quindi il suo personale si suddivide in esterni e dipendenti (**TOTALE** e **ESCLUSIVA**)

Progettare una base di dati
- 27 -
Basi di Dati

Tipi di Generalizzazione - 2

Progettare una base di dati
- 28 -
Basi di Dati



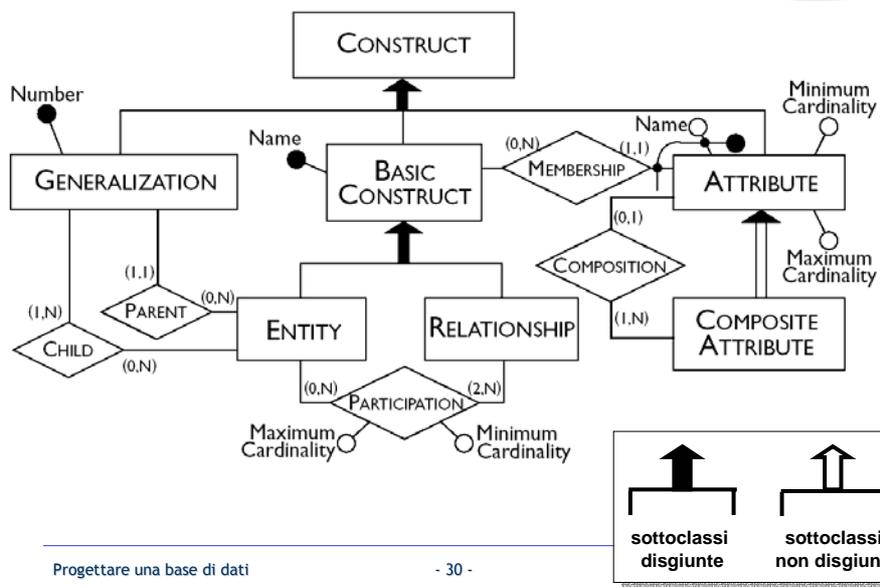
Note sulle generalizzazioni



- Tutte le generalizzazioni possono facilmente essere trasformate in esclusive aggiungendo entità figlie che rappresentano le sovrapposizioni
 - ▶ Esempio: La generalizzazione di *Persona* con *Studente* e *Lavoratore* può diventare esclusiva aggiungendo l'entità figlia *StudenteLavoratore*
- E' possibile effettuare generalizzazioni su più livelli: si ottiene una **gerarchia** di generalizzazioni



Il modello E-R con il modello E-R

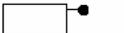




Riepilogo E-R e note



- Gli schemi E-R non sono utili solo in progettazione, ma anche
 - ▶ Documentazione
 - ▶ Descrizione di sistemi già esistenti
- Gli schemi E-R non sono però sufficienti da soli, ma necessitano di documentazione a corredo

Construct	Graphical representation
Entity	
Relationship	
Simple attribute	
Composite attribute	
Cardinality of a	
Cardinality of an attribute	
Internal identifier	
External identifier	
Generalization	
Subset	

Progettare una base di dati

- 31



Documentazione di schemi E-R (1)



- Dizionario dei dati: una tabella per le **entità** e una per le **relazioni**

Entità	Descrizione	Attributi	Identificatore

Relazione	Descrizione	Entità coinvolte	Attributi

- L'uso del dizionario è raccomandato nel caso in cui lo schema sia particolarmente complesso e risulti pesante specificare direttamente su di esso tutti gli attributi di entità e relazioni

Progettare una base di dati

- 32 -

Basi di Dati



Documentazione di schemi E-R (2)



- **Regole di Vincolo e di Derivazione**

- ▶ **Vincolo:** rappresenta un vincolo di integrità non direttamente esprimibile con i costrutti del modello E-R
 - Si possono esprimere con frasi del tipo *<concetto> deve/non deve <espressione su concetti>*
 - Es. un impiegato NON DEVE avere uno stipendio maggiore del direttore del dipartimento a cui afferisce
- ▶ **Derivazioni:** rappresenta un concetto che può essere ottenuto da altri concetti dello schema
 - Si possono esprimere con frasi del tipo: *<concetto> si ottiene <operazioni su concetti>*
 - Es. il numero degli impiegati di un dipartimento si ottiene contando gli impiegati che vi afferiscono



Documentazione di schemi E-R (3)



- I vincoli e le derivazioni essere opportunamente codificate nel modello logico e fisico
- Alcuni possibili approcci
 - ▶ Clausole appropriate del linguaggio SQL
 - ▶ Opportune procedure in un linguaggio di programmazione

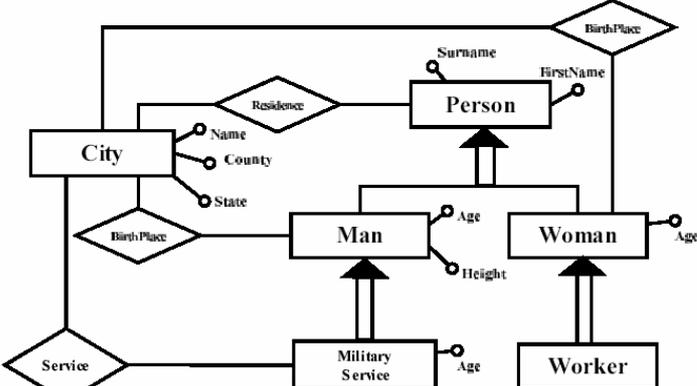
Regole di Vincolo
...
Regole di Derivazione
...



Esercizio



1. Correggere lo schema prendendo in considerazione le proprietà fondamentali delle generalizzazioni
2. Modificare lo schema per considerare tutti i lavoratori
3. Fra gli attributi di "City", possiamo anche vedere "State" come una proprietà di "County". Ristrutturare lo schema in tal senso
4. Indicare le cardinalità e eventuali regole associabili allo schema

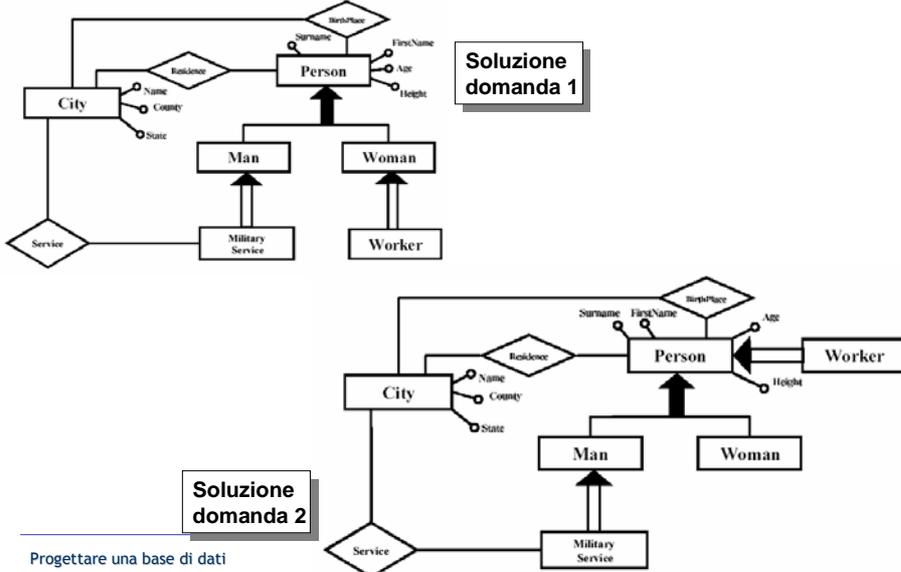


Progettare una base di

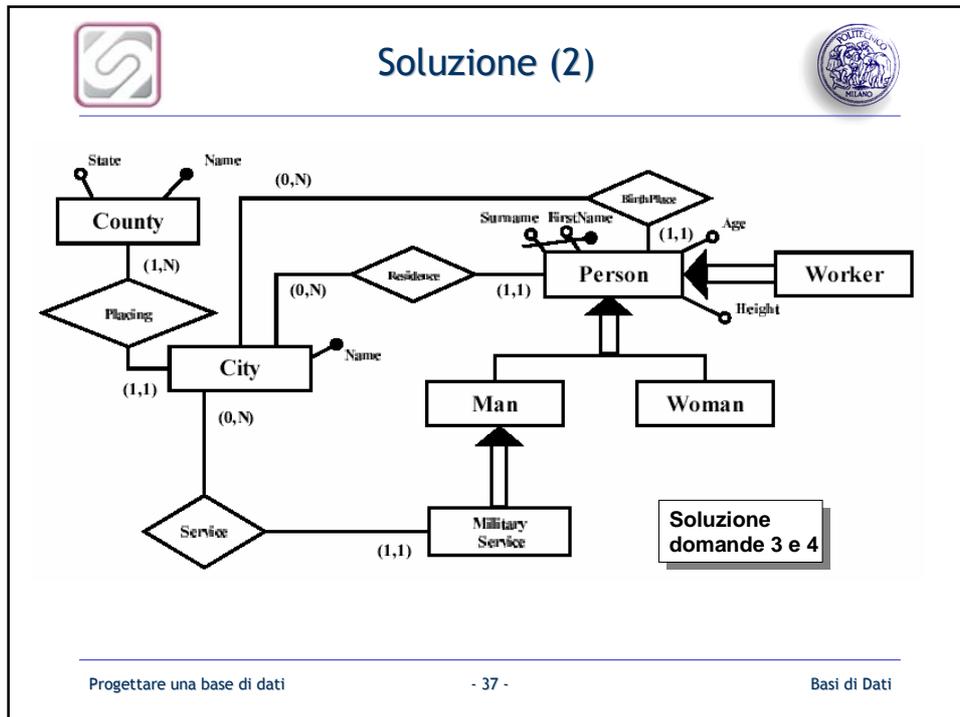


Soluzione (1)



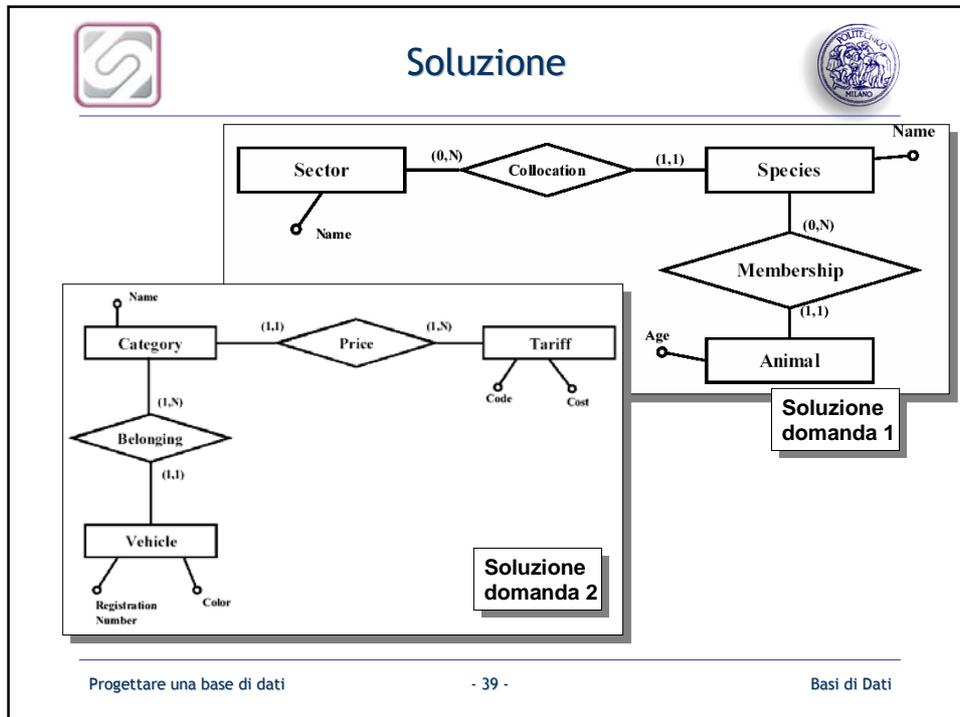


Progettare una base di



The slide, titled "Esercizio", contains two tasks for an Entity-Relationship model:

1. In un giardino zoologico ci sono animali appartenenti a specie aventi una certa età; ogni specie è situata in un settore dello zoo (avente un particolare nome)
2. Un'azienda di noleggio auto ha un insieme di veicoli. Ogni auto ha un numero di registrazione, un colore e appartiene ad una certa categoria. Per ogni categoria c'è una tariffa di noleggio



The slide, titled "Bibliografia", lists the following reference:

- Paolo Atzeni, Stefano Ceri, Stefano Paraboschi, Riccardo Torlone
Basi di Dati - Seconda edizione
Capitolo 5

An illustration of a yellow book with blue rings is shown at the bottom right of the slide.