

Programma del Corso

l'informazione e la sua codifica nei modelli informatici

- **SISTEMI INFORMATIVI E INFORMATICI**
- **MODELLO RELAZIONALE DEI DATI**
- **ALGEBRA RELAZIONALE: operazioni**
- **LINGUAGGIO SQL**
- **SISTEMI PER LA GESTIONE DEI DB**

- **SISTEMI PER LA GESTIONE DEI TESTI**
- **INDICIZZAZIONE**
- **RECUPERO**

Nella prima parte del corso lo scopo è quello di descrivere le principali problematiche relative alla progettazione di database, con particolare riferimento alla modellizzazione e ai linguaggi di programmazione

Programma del Corso

I'XML come linguaggio per documenti strutturati

- COS'È IL MARKUP
- STORIA E INTRODUZIONE ALL'XML
- DOCUMENTI, DATI E RAPPRESENTAZIONI
- DEFINIZIONI E APPLICAZIONI DI GRAMMATICHE
- MODELLI E PRESENTAZIONI DEI DOCUMENTI XML
- STRUMENTI E INDIRIZZI PER XML

Esso risulta essere abbastanza generale per poter essere utilizzato nei più disparati contesti: dalla **definizione della struttura di documenti** allo scambio di informazioni tra sistemi diversi, dalla rappresentazione di immagini alla definizione di formati di dati. Questo aspetto rappresentava una rivoluzione.

Sistemi per il recupero delle informazioni

SISTEMI INFORMATIVI E INFORMATICI

“L’informazione migliora la vita”

Il possesso di informazioni giuste, al momento opportuno, consente di affrontare con successo situazioni di incertezza in cui occorre prendere decisioni.

La situazione più frequente non è quella in cui mancano le informazioni, ma quella in cui non si sa come ritrovare le informazioni, sebbene presenti.

Uno degli usi più comuni dei calcolatori è proprio per la **memorizzazione e recupero di informazioni**.

Esempio. Disponendo di un elenco telefonico automatizzato, si vuole trovare l’indirizzo di un abbonato dal suo numero di telefono.

ORGANIZZAZIONE

Si considera il caso del trattamento delle informazioni in un'organizzazione, per migliorarne il funzionamento.

DEFINIZIONE. Chiamiamo **organizzazione** un insieme organizzato di uomini, risorse, strumenti e procedure, finalizzato al conseguimento di alcuni obiettivi o a fornire dei servizi.

Esempi. Uno studio professionale, un'azienda, un ente pubblico o un ente di servizi (banca, assicurazione, ecc.)

Ogni organizzazione, per il suo funzionamento, ha bisogno di disporre di informazioni accurate e di poterle altresì elaborare tempestivamente.

Quando l'attenzione è sulla **risorsa informazione** dell'organizzazione, si dice che di essa interessa il suo **sistema informativo**.

SISTEMI INFORMATIVI

DEFINIZIONE. Un **sistema informativo** è una combinazione di risorse, umane e materiali, e di procedure organizzate per la raccolta, l'archiviazione, l'elaborazione e lo scambio delle informazioni necessarie alle attività operative, alle attività di gestione e alle attività di programmazione, controllo e valutazione.

Il termine *SISTEMA* evidenzia il fatto che esiste un insieme organizzato di elementi, mentre il termine *INFORMATIVO* precisa che tutto ciò è finalizzato alla gestione delle informazioni.

Delle informazioni trattate da un'organizzazione, interessa qui prendere in considerazione solo quelle **strutturate**, con formato predeterminato; si tralasciano invece informazioni non strutturate (es. disegni, immagini).

SISTEMI INFORMATIVI

Ogni sistema informativo comprende raccolte di **informazioni di tipo omogeneo**

Ad es. il sistema dell'università comprende informazioni sugli studenti, le aule e i docenti

Esistono delle correlazioni logiche tra questi insiemi di informazioni (fra insegnamenti ed aule, studenti e esami, ...)

Il sistema informativo è parte del sistema organizzativo .

Quasi sempre, il sistema informativo è di supporto ad altri sottosistemi, e va quindi studiato nel contesto in cui è inserito

Il concetto di “sistema informativo” è indipendente da qualsiasi automatizzazione: esistono organizzazioni la cui ragion d'essere è la gestione di informazioni (p. es. servizi anagrafici e banche) e che operano da secoli

ESEMPI DI ORGANIZZAZIONI

Vediamo alcuni esempi di organizzazioni, di diversa complessità, evidenziando le finalità del trattamento delle informazioni:

- ☐ Uno **studio medico**, come organizzazione, mantiene informazioni su
 - ☐ pazienti, visite fatte, parcelle richieste

- ☐ Un **comune** gestisce informazioni per svolgere le seguenti attività:
 - ☐ gestione dei servizi demografici (anagrafe, stato civile, servizio elettorale e vaccinale)
 - ☐ gestione della rete viaria
 - ☐ gestione dell'attività finanziaria
 - ☐ gestione del personale per il calcolo della retribuzione
 - ☐ gestione dei servizi amministrativi
 - ☐ gestione dei servizi sanitari delle ASL
 - ☐ gestione della cartografia generale e tematica del territorio

ESEMPI DI ORGANIZZAZIONI

- Una biblioteca gestisce informazioni sui materiali raccolti, sui prestiti, sulle persone che prendono in prestito per svolgere attività rivolte
 - alla raccolta dei documenti
 - descrizione dei documenti
 - organizzazione di archivi e cataloghi
 - alla conservazione e consultazione dei documenti
 - alla gestione della biblioteca

ESEMPI DI ORGANIZZAZIONI

Le informazioni possono essere raggruppate in base al tipo di elaborazioni a cui sono sottoposte

- elaborazioni rivolte alla gestione dei rapporti con l'esterno (**processi produttivi o operativi**)
- elaborazioni rivolte alla gestione operativa dell'organizzazione (**processi gestionali**): conoscenza, gestione e controllo delle risorse: personale, contabilità,...
- elaborazioni rivolte all'attività di programmazione per fissare priorità di interventi (**processi decisionali o di governo**).

SISTEMA INFORMATICO

Le informazioni di un'organizzazione, una volta ridotte a dati, possono essere trattate automaticamente con gli elaboratori elettronici.

DEFINIZIONE. Il **sistema informatico** indica l'insieme degli strumenti informatici impiegati per il trattamento automatico delle informazioni al fine di agevolare le funzioni del sistema informativo

Strumenti tipici:

- hardware
- software
- metodologie (algoritmi, modelli matematici)

Dato ed informazione

- L'automazione consente di soddisfare esigenze di **informazione** ben piu' sofisticate che nei sistemi tradizionali
 - quali sono le opere tradotte dal tedesco dopo il 1968?
 - informazioni statistiche tipo la frequenza nei musei d'estate
 - la media dei voti degli studenti per un dato esame
- In un **sistema informatico** le informazioni sono rappresentate per mezzo di **dati** che diventano informazione dopo un processo di **interpretazione**

Dato ed informazione

(definizioni dal Vocabolario della lingua italiana 1987)

informazione: notizia, dato o elemento che consente di avere conoscenza più o meno esatta di fatti, situazioni, modi di essere.

dato: ciò che è immediatamente presente alla conoscenza, prima di ogni elaborazione; (in informatica) elementi di informazione costituiti da simboli che debbono essere elaborati

Esempio:

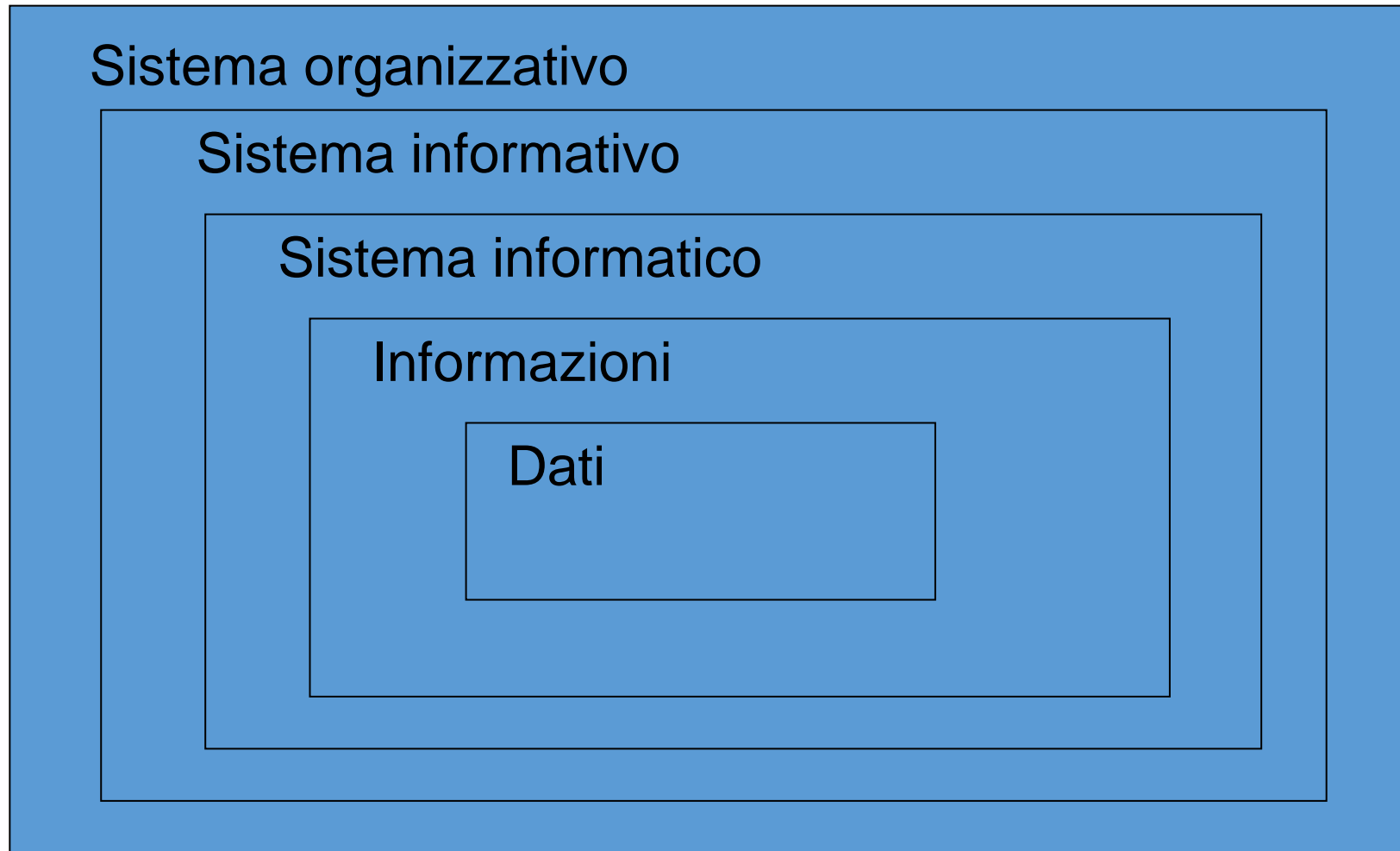
Mario 528275

sono due dati su un foglio di carta

Se il foglio di carta viene fornito in risposta alla domanda: “A chi mi devo rivolgere per questo problema? Qual e’ il suo numero di telefono?”

allora i due dati diventano informazione

SCHEMA RIASSUNTIVO



EVOLUZIONE DEI SISTEMI INFORMATICI

Anni 60

L'esigenza più sentita era una strumentazione per migliorare l'efficienza e la produttività di alcune parti dei processi operativi. Ciò ha determinato una diffusione della tecnologia informatica per applicazioni settoriali con l'obiettivo di automatizzare quelle attività che richiedono l'elaborazione di grandi quantità di dati.

vantaggi:

la correttezza dei risultati, la riduzione dei costi e la maggiore produttività settoriale.

svantaggio:

scarsa integrazione dei dati in comune ai diversi settori

EVOLUZIONE DEI SISTEMI INFORMATICI

Anni 70

Il progresso tecnologico ha permesso un *trattamento globale* dei dati. Ciascuna informazione benché rappresentata una volta sola era utilizzabile per settori diversi.

Si passa da un sistema informatico settoriale a sistemi informatici per l'organizzazione

Anni 80 ad oggi

L'aumento delle velocità di elaborazione dei calcolatori e i progressi nella memorizzazione dei dati hanno consentito l'archiviazione e l'elaborazione dell'informazione nelle varie forme che essa può assumere – dato, testo, suono, documento, disegno, immagine, etc. – allargando lo spettro delle applicazioni di questa tecnologia.

Strumenti informatici per la gestione dei dati

- **Sistemi per la gestione di base dati** (SGBD) o Data Base Management System (DBMS)
 - una base di dati è una collezione di dati atti a rappresentare le informazioni che interessano un sistema informativo
 - un SGBD è un sistema software in grado di gestire (costruire, modificare, utilizzare, ...) una base dati
- **Sistemi di recupero delle informazioni** (SRI) o Information Retrieval System (IRS)

PERCHE` GESTIRE I DATI?

- La rappresentazione precisa di forme più ricche di informazione e conoscenza è difficile
- I dati costituiscono spesso una risorsa strategica, perché più stabili nel tempo di altre componenti (processi, tecnologie, ruoli umani):
 - ad esempio, i dati delle banche o delle anagrafi

Caratteristiche di un DB

- Un DB è costituito da una grande quantità di dati, organizzati in insiemi omogenei in correlazione tra loro
 - i dati sono memorizzati e gestiti in memoria permanente (disco)
 - i dati sono *condivisi fra più utenti e fra più applicazioni* per ridurre la *ridondanza* e la possibilità di *inconsistenze*.
- In un DB esiste:
 - una parte sostanzialmente invariante nel tempo, detto **schema della base di dati**, che descrive le caratteristiche dei dati.
 - una parte variabile nel tempo, detta **istanza** o stato del DB, costituita dai valori effettivi dei dati

Caratteristiche di un DB

- Esempio: per rappresentare i corsi attivati ed i relativi docenti si può usare uno schema del tipo

(Corso, Nome Docente)

- questo e' uno schema invariante nel tempo,

- mentre le coppie

(Informatica Generale, Righi)

(Basi di Dati e Sistemi Informativi, Leoni)

- sono istanze dello schema precedente e possono variare nel tempo

DATABASE

- accezione generica, metodologica
 - insieme organizzato di dati utilizzati per il supporto allo svolgimento delle attività di un ente (azienda, ufficio, persona)
- accezione specifica, metodologica e tecnologica)
 - insieme di dati gestito da un DBMS
- Nella nostra accezione i database sono:
 - un supporto digitale in cui i dati sono immagazzinati
 - un software che permetta il recupero dei dati
 - una rete che permetta un'accessibilità condivisa ai dati

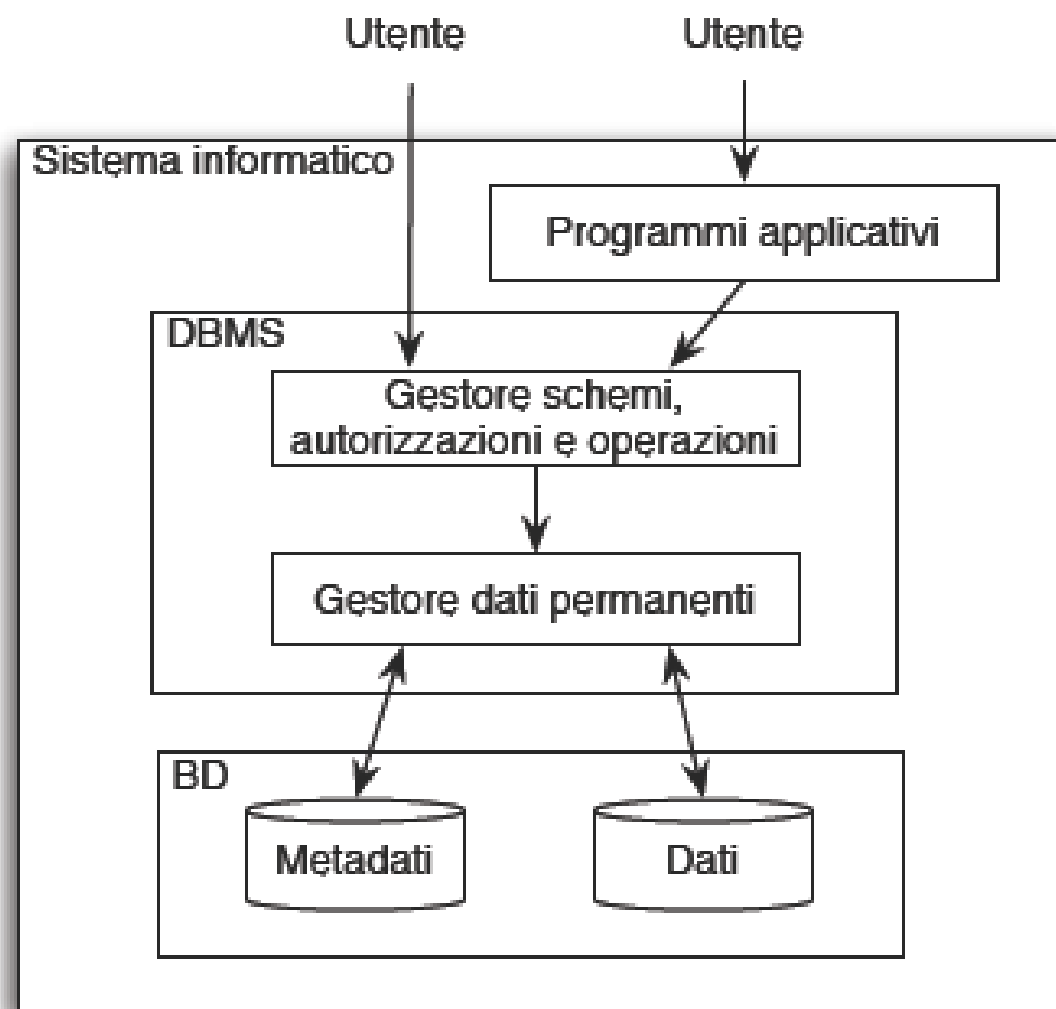
SCHEMA E ISTANZA

Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	9:45
Chimica	Nicola Mori	N1	9:45
Fisica I	Mario Bruni	N1	11:45
Fisica II	Mario Bruni	N3	9:45
Sistemi inform.	Piero Rossi	N3	8:00

colonne: lo schema della relazione

righe: l'istanza della relazione



Sistemi per base di dati - DBMS

- **Definizione:** Un **DBMS** è un sistema centralizzato o distribuito che offre opportuni linguaggi per:
 - definire lo schema di una basi di dati (lo schema va definito prima di creare dati),
 - scegliere le strutture dati per la memorizzazione dei dati,
 - memorizzare i dati rispettando i vincoli definiti nello schema;
 - recuperare e modificare i dati interattivamente (linguaggio di interrogazione o query language) o da programmi.

Sistemi per base di dati - DBMS

- I DBMS offrono dei **linguaggi** per
 - la definizione dello schema del DB
 - l'uso della base di dati (*accesso ai dati, inserimento, modifica, cancellazione*)
- I DBMS garantiscono una gestione **efficiente** del DB
 - le operazioni devono essere eseguite utilizzando un insieme di risorse (*tempo e spazio*) che siano *accettabili per gli utenti*.
- I DBMS garantiscono **l'affidabilità** dell DB
 - i dati sono protetti da errori dovuti a malfunzionamenti e interferenze.
- I DBMS garantiscono la **privatezza** del DB
 - i dati sono protetti dall'accesso e dall'uso da parte di utenti non autorizzati.

Sistemi per base di dati - DBMS

- Diverse categorie di persone possono interagire con una base di dati tramite un DBMS.
 - *Amministratore della base di dati (DBA).*
 - responsabile della progettazione, controllo e amministrazione della base di dati.
 - *I progettisti e programmatori di applicazioni.*
 - realizzano i programmi che accedono e interrogano la base di dati.
 - Gli *utenti non programmatori che utilizzano* la base di dati per le proprie attività.

CARATTERISTICHE DEI DATI GESTITI DAI DBMS

- Sono organizzati in insiemi omogenei, fra i quali sono definite delle relazioni. La struttura dei dati e le relazioni sono descritte nello schema usando i meccanismi di astrazione del modello dei dati del DBMS;
- Sono molti, in assoluto e rispetto ai metadati, e non possono essere gestiti in memoria temporanea;
- Sono accessibili mediante transazioni, unità di lavoro atomiche che non possono avere effetti parziali;
- Sono protetti sia da accesso da parte di utenti non autorizzati, sia da corruzione dovuta a malfunzionamenti hardware e software;
- Sono utilizzabili contemporaneamente da utenti diversi.

DB e DBMS

- I DATABASE SONO **GRANDI**...
 - dimensioni (molto) maggiori della memoria centrale dei sistemi di calcolo utilizzati
 - il limite deve essere solo quello fisico dei dispositivi
 - hanno un tempo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano
- I DATABASE SONO **CONDIVISI**
 - Ogni organizzazione (specie se grande) è divisa in settori o comunque svolge diverse attività
 - Ciascun settore/attività ha un (sotto)sistema informativo (non necessariamente disgiunto)

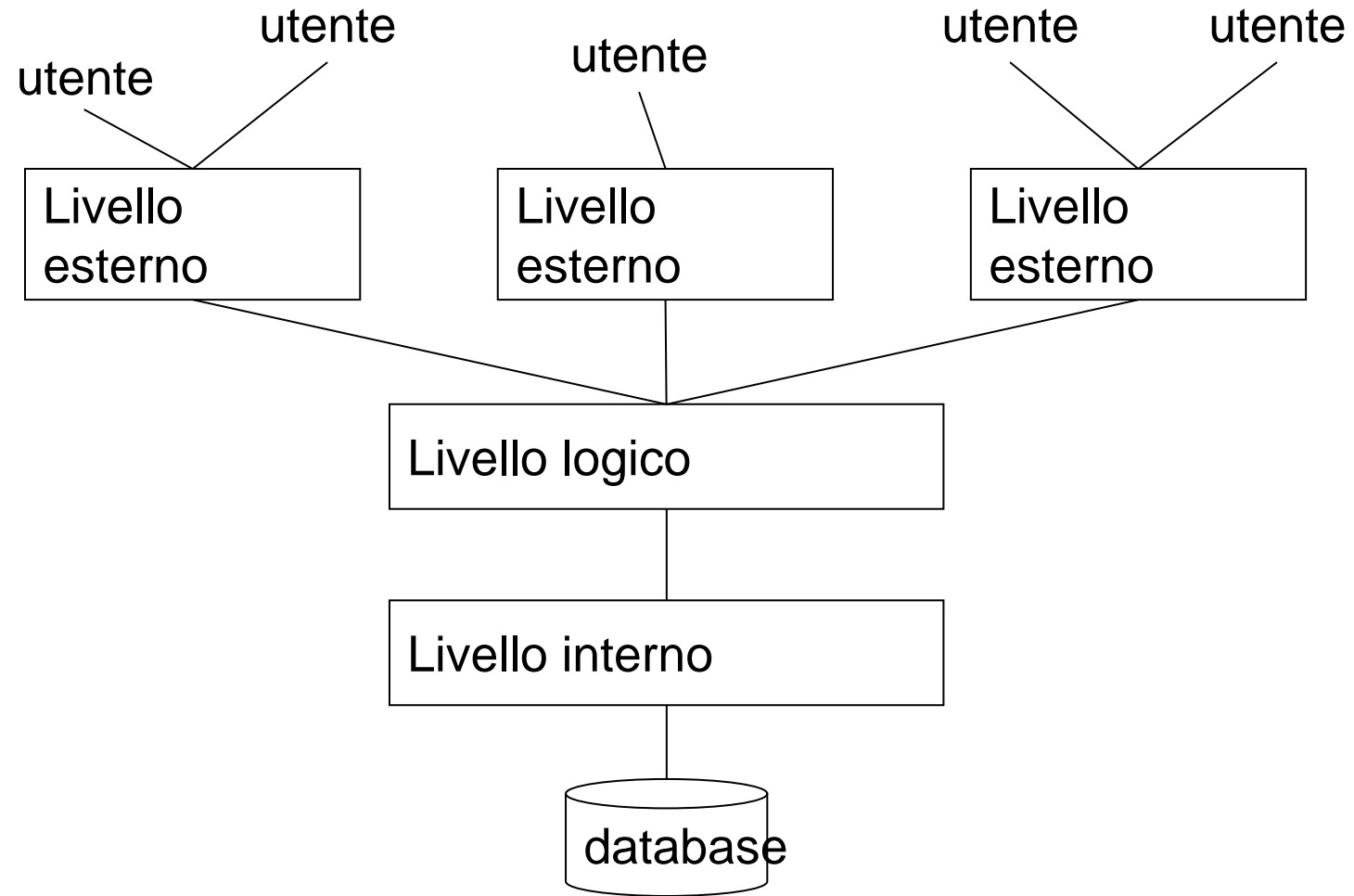
DB e DBMS

- I DBMS GARANTISCONO **PRIVATEZZA**
 - Si possono definire meccanismi di autorizzazione
 - l'utente A è autorizzato a leggere tutti i dati e a modificare quelli sul ricevimento
 - l'utente B è autorizzato a leggere X e a modificare Y
- I DBMS GARANTISCONO **AFFIDABILITA`**
 - Affidabilità (per le basi di dati):
 - resistenza a malfunzionamenti hardware e software
- Una base di dati è una risorsa pregiata e quindi deve essere conservata a lungo termine

DB e DBMS

- I DBMS DEVONO ESSERE EFFICIENTI
 - Cercano di utilizzare al meglio le risorse di spazio di memoria (principale e secondaria) e tempo (di esecuzione e di risposta)
 - I DBMS, con tante funzioni, rischiano l'inefficienza e per questo ci sono grandi investimenti e competizione
 - L'efficienza è anche il risultato della qualità delle applicazioni
- I DBMS DEVONO ESSERE EFFICACI
 - Cercano di rendere produttive le attività dei loro utilizzatori, offrendo funzionalità articolate, potenti e flessibili:
 - il corso è in buona parte dedicato ad illustrare come i DBMS perseguono l'efficacia

ARCHITETTURA STANDARD - I



ARCHITETTURA STANDARD - II

livello logico: descrizione dell'intera base di dati nel modello logico "principale" del DBMS

livello interno (o fisico): rappresentazione dello schema logico per mezzo di strutture fisiche di memorizzazione (es.: files)

livello esterno: descrizione di parte della base di dati in un modello logico ("viste" parziali, derivate, anche in modelli diversi)

LIVELLO ESTERNO: ESEMPIO

Corsi

Corso	Docente	Aula
Basi di dati	Rossi	DS3
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G

Aule

Nome	Edificio	Piano
DS3	OMI	Terra
N3	OMI	Terra
G	Pincherle	Primo

CorsiSedi

Corso	Aula	Edificio	Piano
Sistemi	N3	OMI	Terra
Reti	N3	OMI	Terra
Controlli	G	Pincherle	Primo

INDIPENDENZA DEI DATI

- L'approccio con tre livelli di descrizione dei dati è stato proposto come un modo per garantire le proprietà di indipendenza logica e fisica dei dati, che sono fra gli obiettivi più importanti dei DBMS.
- l'accesso avviene solo tramite il livello esterno (che può coincidere con il livello logico)
- due forme:
 - indipendenza fisica
 - indipendenza logica

INDIPENDENZA FISICA

- il livello logico e quello esterno sono indipendenti da quello fisico
 - una relazione è utilizzata nello stesso modo qualunque sia la sua realizzazione fisica
 - la realizzazione fisica può cambiare senza che debbano essere modificati i programmi
 - i programmi applicativi non devono essere modificati in seguito a modifiche dell'organizzazione fisica dei dati.

INDIPENDENZA LOGICA

- il livello esterno è indipendente da quello logico
- aggiunte o modifiche alle viste non richiedono modifiche al livello logico
- modifiche allo schema logico che lascino inalterato lo schema esterno sono trasparenti
- i programmi applicativi non devono essere modificati in seguito a modifiche dello schema logico.

LINGUAGGI PER DATABASE

- Un altro contributo all'efficacia: disponibilità di vari linguaggi e interfacce
 - ⇒ linguaggi testuali interattivi (SQL)
 - ⇒ comandi (SQL) immersi in un linguaggio ospite (Pascal, Java, C ...)
 - ⇒ comandi (SQL) immersi in un linguaggio ad hoc, con anche altre funzionalità (p.es. per grafici o stampe strutturate)
 - ⇒ con interfacce amichevoli (senza linguaggio testuale)

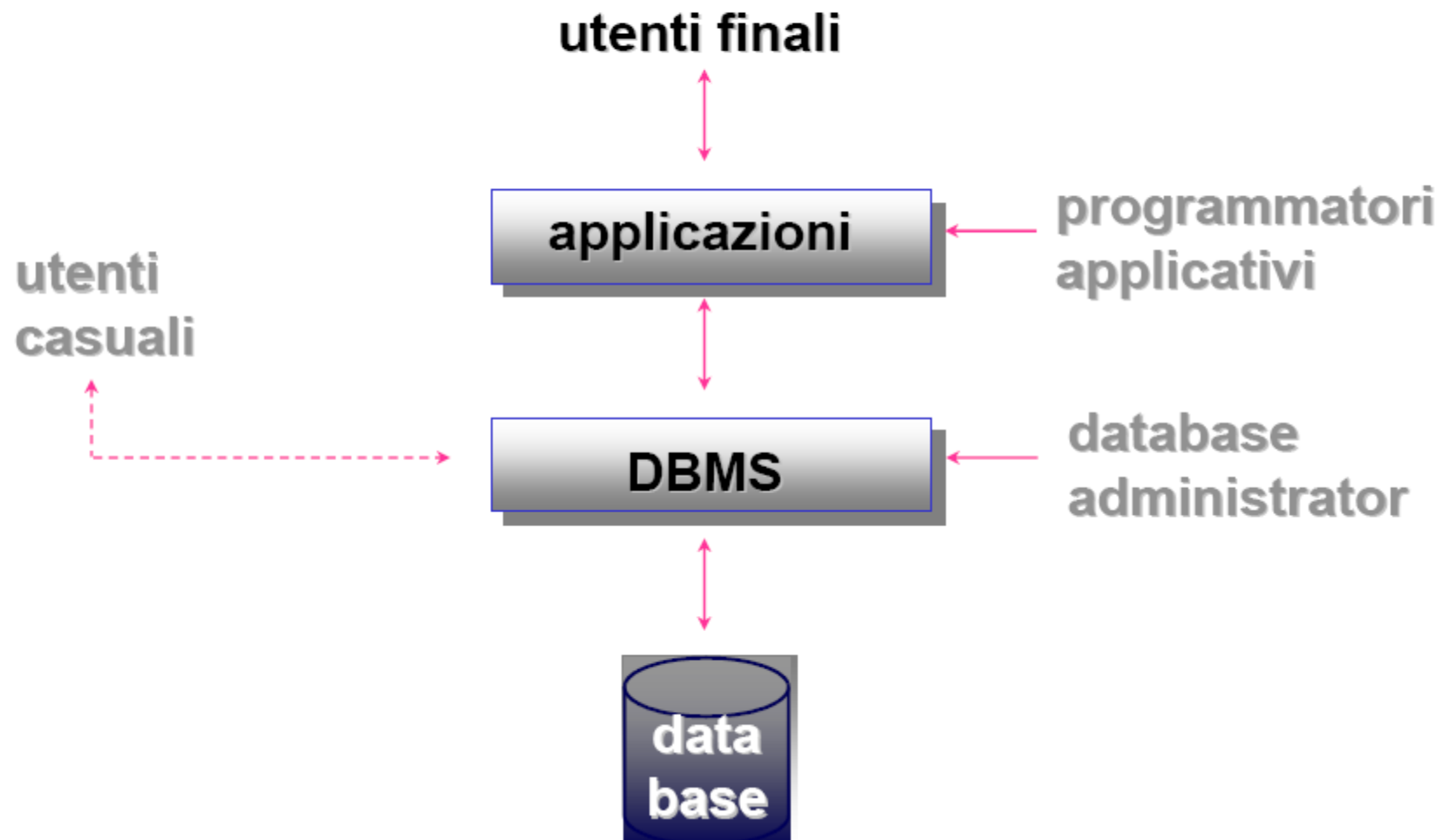
PERSONAGGI E INTERPRETI

- progettisti e realizzatori di DBMS
- progettisti del database e amministratori del database
- progettisti e programmatori di applicazioni
- utenti
 - utenti finali (terminalisti): eseguono applicazioni predefinite (transazioni)
 - utenti casuali: eseguono operazioni non previste a priori, usando linguaggi interattivi

DATABASE ADMINISTRATOR (DBA)

- Persona o gruppo di persone responsabile del controllo centralizzato e della gestione del sistema, delle prestazioni, dell'affidabilità, delle autorizzazioni
- Le funzioni del DBA includono quelle di progettazione, anche se in progetti complessi ci possono essere distinzioni

Utenti del DBMS



TRANSAZIONI (PER L'UTENTE)

- Programmi che realizzano attività frequenti e predefinite, con poche eccezioni, previste a priori.
- Esempi:
 - versamento presso uno sportello bancario
 - emissione di certificato anagrafico
 - dichiarazione presso l'ufficio di stato civile
 - prenotazione aerea
- Le transazioni sono di solito realizzate in linguaggio ospite (tradizionale o ad hoc)
- Per l'utente:
 - programma a disposizione, da eseguire per realizzare una funzione di interesse
- Per il sistema:
 - sequenza indivisibile di operazioni

VANTAGGI E SVANTAGGI DEI DBMS

Pro

- dati come risorsa comune, database come modello della realtà
- gestione centralizzata con possibilità di standardizzazione ed “economia di scala”
- disponibilità di servizi integrati
- riduzione di ridondanze e inconsistenze
- indipendenza dei dati (favorisce lo sviluppo e la manutenzione delle applicazioni)

Contro

- costo dei prodotti e della transizione verso di essi
- non scorporabilità delle funzionalità (con riduzione di efficienza)

Sistemi per il recupero delle informazioni

MODELLI

MODELLI

L'informatica offre metodologie e strumenti per la costruzione di modelli di situazioni reali che ricorrono in ogni campo che richiede un'attività di progettazione. Essi permettono di riprodurre le caratteristiche essenziali di fenomeni reali, omettendo dettagli che costituirebbero inutili complicazioni.

Per poter gestire un sistema informativo con strumenti informatici dobbiamo realizzare una descrizione della realtà d'interesse per mezzo di simboli

Un **modello di dati e' un insieme di concetti utilizzati per organizzare i dati e descriverne la struttura in modo che essa risulti comprensibile a un elaboratore**

Ad es per poter gestire una biblioteca è necessario dare al calcolatore una sua descrizione simbolica

Una tale descrizione costituisce un modello della realtà

MODELLI ASTRATTI

astrazione: processo mentale mediante il quale sono selezionate alcune proprietà caratteristiche di un insieme di oggetti e ne sono ignorate altre

- **modelli in scala** Essi sono la riproduzione in dimensioni ridotte di strutture che hanno grandi dimensioni e che vanno studiate, prima di passare alla loro realizzazione, in condizioni il più possibile simili a quelle reali; gli oggetti sono riprodotti in dimensioni diverse da quelle reali
- **modelli matematici:** le conoscenze relative ad un fenomeno vengono rappresentate con espressioni matematiche
- **modelli simbolici:** le conoscenze relative ad un fenomeno vengono rappresentate con simboli

MODELLI ASTRATTI

Sono definiti come

“la rappresentazione formale di idee e conoscenze relative a un fenomeno”.

Aspetti fondamentali di un modello astratto:

- a) è la rappresentazione di certi fatti;
- b) la rappresentazione è data con un linguaggio formale;
- c) il modello è il risultato di un processo di interpretazione, guidato dalle idee e conoscenze possedute dal soggetto che interpreta.

MODELLI ASTRATTI

L'informatica consente di costruire modelli astratti diversi:

- ☐ **modelli per l'analisi del problema**
- ☐ **modelli per la progettazione della soluzione**
- ☐ **modelli per la realizzazione del progetto**

La differenza fra le varie categorie di modelli sta nel diverso livello di astrazione a cui si opera

MODELLI INFORMATICI

Sono **modelli simbolici**: rappresentazione formale delle idee e conoscenze relative ad un fenomeno

Si hanno diverse categorie di **modelli informatici** che differiscono per i tipi di fatti che si considerano, ossia per i diversi livelli di astrazione a cui si opera

- **modelli concettuali**: sono considerati solo i fatti relativi alla realtà senza alcun riferimento agli strumenti informatici che dovranno usare il modello
- **modelli logici**: sono considerati anche fatti relativi agli strumenti informatici; il modello è adeguato alle caratteristiche del DBMS, in modo che possa essere da questo interpretato
- **modelli fisici**: sono considerate le strutture fisiche usate dal calcolatore per rappresentare i dati

MODELLI ASTRATTI - ESEMPIO

Un'analogia con l'atlante stradale aiuta a comprendere il ruolo dei diversi livelli di astrazione usati nella costruzione di un modello, in funzione del problema da risolvere: quando si pianifica un viaggio

- **uno sguardo delle singole tavole con la cartografia stradale consente di stabilire l'itinerario e le distanze da percorrere;**
- **un esame delle singole tavole consente di programmare deviazioni per strade panoramiche e le vie di accesso alle città da visitare;**
- **l'esame delle piante delle città consente di fissare dettagliati itinerari di visita.**

MODELLI CONCETTUALI

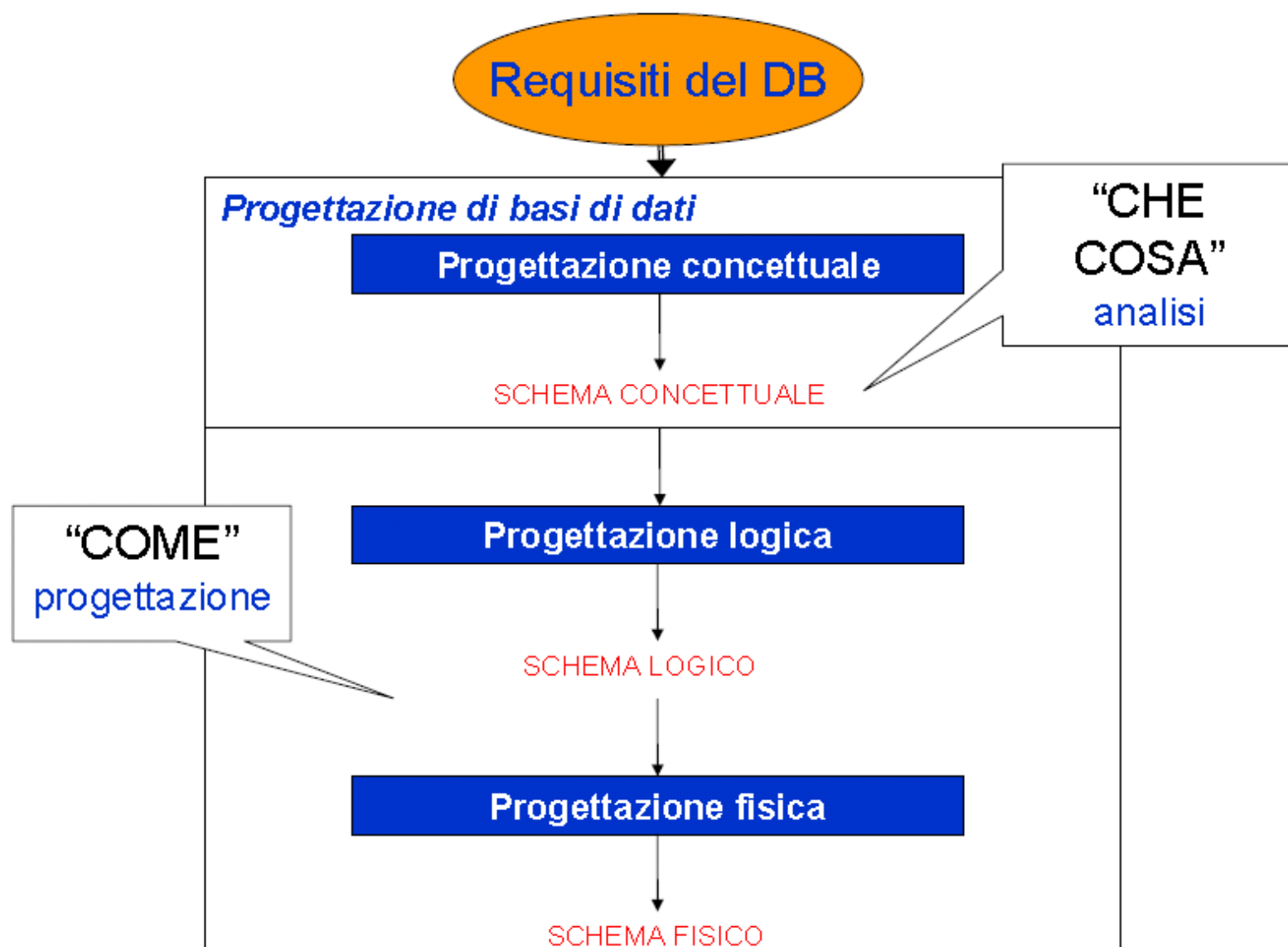
- Permettono di rappresentare i dati in modo indipendente da ogni sistema
 - cercano di descrivere i concetti del mondo reale
 - sono utilizzati nelle fasi preliminari di progettazione
- Il più diffuso è il **modello Entity-Relationship**

MODELLI LOGICI

- Adottati nei DBMS esistenti per l'organizzazione dei dati
 - utilizzati dai programmi
 - indipendenti dalle strutture fisiche
- esempi: relazionale, reticolare, gerarchico, a oggetti
- Modello relazionale:
 - il costruttore e' la relazione
 - la relazione e' una tabella
 - in un database ci sono piu'

Corsi(corso,docente,aula)

Corso	Docente	Aula
Basi di dati	Rossi	DS3
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G



Sistemi per il recupero delle informazioni

COSA SI MODELLA

COSA SI MODELLA

Nella costruzione di un modello si supporrà di modellare la **conoscenza concreta** e la **conoscenza astratta**.

Altri tipi di conoscenze: **conoscenza procedurale**

- ad esempio, quale procedura va seguita per la catalogazione di una nuova pubblicazione in una biblioteca oppure quale procedura va seguita quando si acquista un nuovo libro

LA CONOSCENZA CONCRETA

La conoscenza concreta riguarda i **fatti specifici** che si vogliono rappresentare.

Si suppone che la realtà consista di **entità** che hanno alcune **proprietà**.

Si suppone inoltre che le entità omogenee siano raggruppabili in collezioni e che siano connesse fra di loro da **associazioni** che evolvano nel tempo.

CONOSCENZA CONCRETA: ENTITA' e PROPRIETA'

Le **entità** sono ciò che esiste e di cui interessa rappresentare alcuni fatti o proprietà.

Ad esempio sono entità i libri la “Divina Commedia” o il “Decamerone”, gli utenti “Anna” e “Fabio” di una biblioteca in esame.

Le **proprietà** costituiscono i fatti che interessano soltanto perché descrivono caratteristiche di determinate entità.

Ad esempio Il cognome e il recapito di un utente sono proprietà.

Le proprietà sono fatti che non interessano di per sé, ma solo come caratterizzazione di altri fatti interpretati come entità.

Per meglio comprendere una realtà è opportuno semplificarla e organizzarla con un processo di astrazione che individui tipi di entità.

CONOSCENZA CONCRETA: TIPI DI ENTITA'

Un **tipo entità** è una descrizione astratta di ciò che accomuna un insieme di entità omogenee (della stessa natura), esistenti o possibili.

Un tipo non è una specifica collezione di entità, ma descrive la struttura di tutte le entità “possibili”: va pensato come una collezione infinita di entità possibili.

Il tipo persona descrive non solo tutte le persone che esistono, ma anche quelle che esisteranno o che potrebbero esistere.

CONOSCENZA CONCRETA: TIPI DI ENTITA'

Ad un tipo sono associate le proprietà delle entità che appartengono a tale tipo, nonché le caratteristiche di tali proprietà.

Il tipo utente ha le proprietà cognome e recapito, intendendo con questo che ogni utente ha un cognome e un recapito, ma con un valore in generale diverso da quello di tutti gli altri.

Nell'esame di una realtà, tra tutte le possibili proprietà di entità omogenee, con il processo di astrazione che porta a definire il loro tipo si isolano solo quelle che sono interessanti per il fine che ci si prefigge.

Ad esempio, per gli utenti della biblioteca si ritengono interessanti il cognome e il recapito, ma non il colore degli occhi o dei capelli.

Esempi

- **Tipo entità**

- studente
- libri
- esame

- **Attributi**

- nome, anno nascita, matricola
- titolo, autore, ...
- materia, candidato, docente

Certi fatti possono essere interpretati come entità in certi contesti e come proprietà in altri

- **Descrizione bibliografica con attributi Autori, Titolo, Editore, LuogoEdizione, Anno**

oppure

- **Autore con attributi Nome, Nazionalità, AnnoNascita ...**
- **Editore con attributi Nome, Indirizzo, e-mail, Anno ...**
- **Descrizione bibliografica con attributi Titolo ...**

CONOSCENZA CONCRETA:

CARATTERISTICHE DELLE PROPRIETA'

Ogni proprietà ha un nome, detto anche **attributo**, e un **dominio**, cioè l'insieme dei possibili valori che può assumere.

Ogni proprietà è una coppia (attributo, valore di un certo tipo)

Ogni proprietà può essere classificata come segue:

- **atomica/strutturata**
- **unione / semplice**
- **univoca / multivalore**
- **totale (obbligatoria)/parziale**
- **costante/variabile**
- **calcolata/non calcolata**

CONOSCENZA CONCRETA:

CARATTERISTICHE DELLE PROPRIETA'

Una proprietà è **atomica** se il suo valore non è scomponibile
(es. il cognome di una persona); altrimenti e' detta **strutturata** (es. la residenza
è scomponibile in indirizzo, CAP,..);

Una proprietà è di tipo **unione**, se il suo valore può essere di tipi diversi
(es. la proprietà titolare di un corso può essere un professore associato o un
professore ordinario); altrimenti è detta **semplice**.

Una proprietà è **univoca**, se il suo valore è unico
(es. cognome di un utente ha un unico valore); altrimenti è detta **multivalore**.
(es. la proprietà “recapiti telefonici” di una persona è multivalore se si ammette
che alcune persone possano essere raggiunte attraverso più numeri telefonici);

CONOSCENZA CONCRETA:

CARATTERISTICHE DELLE PROPRIETA'

Una proprietà è **totale (obbligatoria)**, se ogni entità dell'universo del discorso ha per essa un valore specificato, altrimenti è detta **parziale (opzionale)**

(es. si può considerare il cognome di un utente una proprietà totale ed il suo recapito telefonico una proprietà parziale);

Una proprietà è **costante**, se il suo valore non cambia nel tempo, altrimenti è detta **variabile**

(es. data di nascita è una costante, l'indirizzo è variabile);

Una proprietà è **calcolata**, se il suo valore può essere determinato a partire dal valore di altre proprietà, altrimenti è detta **non calcolata**.

(es. l'età può essere calcolata dalla data di nascita).

CONOSCENZA CONCRETA: COLLEZIONI DI ENTITA'

Le entità dello stesso tipo possono essere raggruppate in collezioni che chiameremo **classi**.

DEFINIZIONE. Una **classe** è un insieme variabile nel tempo di entità omogenee (dello stesso tipo).

Esempio. La classe dei libri è l'insieme dei libri che la biblioteca possiede ad un certo istante.

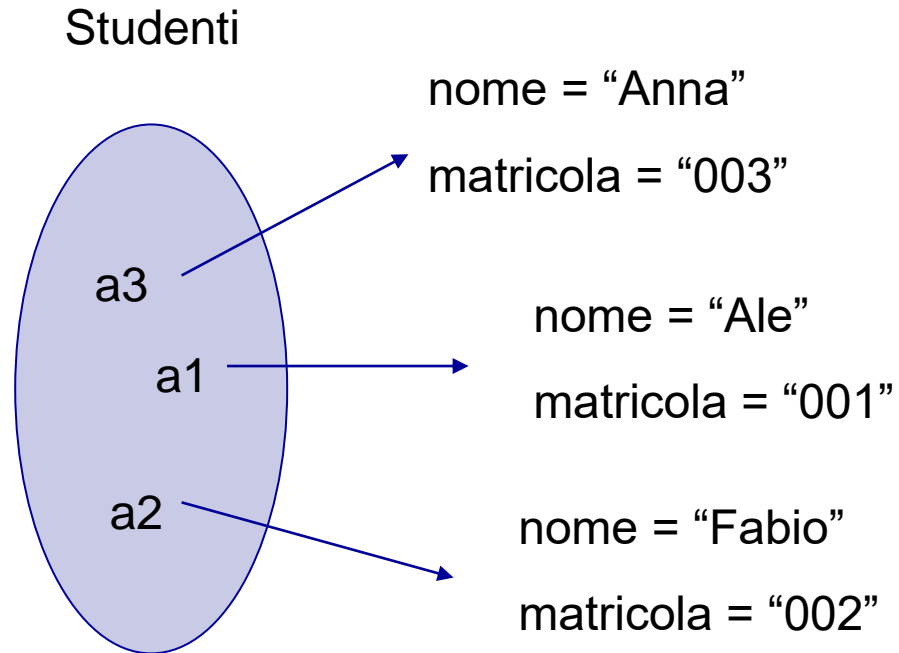
Una classe ha due aspetti:

- uno **intensionale**, invariante nel tempo
- uno **estensionale**, variabile nel tempo

L'aspetto intensionale riguarda il tipo degli elementi, mentre l'aspetto estensionale riguarda l'insieme dei suoi elementi.

Come accade in generale per gli **insiemi**, gli elementi di una classe possono essere dati in due modi: elencandoli in modo esplicito (ad esempio, gli elementi della classe delle persone sono Mario, Giorgio ecc.), oppure caratterizzandoli mediante una condizione sui valori delle loro proprietà (ad esempio, i minorenni sono tutte le persone con età inferiore a 18 anni).

Esempio di classe



CONOSCENZA CONCRETA:

GERARCHIE DI COLLEZIONI

Le classi di entità sono organizzate in **gerarchie di specializzazione o di generalizzazione**

DEFINIZIONE. Una **gerarchia** modella, all'interno delle classi, entità ad un diverso livello di dettaglio. Una classe della gerarchia minore di altre viene detta **sottoclasse**, mentre una classe della gerarchia maggiore di altre viene detta **superclasse**.

Gli elementi di una sottoclasse ereditano le caratteristiche degli elementi della superclasse.

Esempio. Classificazione degli organismi animali e vegetali: dicendo che “i marsupiali sono mammiferi” intendiamo che le femmine sono dotate di ghiandole mammarie per l'allattamento dei piccoli, proprietà di ogni mammifero, ma hanno come proprietà specifica il marsupio.

Come le classi **anche le sottoclassi hanno un aspetto intensionale**, riguardante il tipo dei loro elementi, **e uno estensionale**, riguardante gli elementi che le compongono.

CONOSCENZA CONCRETA:

GERARCHIE DI COLLEZIONI

Per quanto riguarda l'aspetto intensionale, se C è una sottoclasse di D, allora il tipo degli elementi di C è un sottotipo del tipo degli elementi di D, ovvero gli elementi di C ereditano tutte le proprietà degli elementi di D, ma possono avere anche altre proprietà specifiche (*vincolo intensionale*).

Ad esempio, se il tipo studente è definito come sottotipo del tipo utente, uno studente automaticamente eredita le proprietà delle persone, ma può anche avere altre proprietà come matricola, corso di laurea, etc.

Per quanto riguarda l'aspetto estensionale, se C è una sottoclasse di D, allora ogni elemento di C è anche un elemento di D, ovvero gli elementi di C sono sempre un sottoinsieme degli elementi di D (*vincolo estensionale*).

CONOSCENZA CONCRETA:

GERARCHIE DI COLLEZIONI

E' utile distinguere almeno tre modalità di definizione di sottoclassi: per **sottoinsieme**, per **sottoinsiemi disgiunti**, e per **partizione**.

Le sottoclassi sottoinsieme, specializzazioni della stessa superclasse, non sono in generale fra loro disgiunte. Un elemento della classe utenti può essere contemporaneamente sia un elemento della sottoclasse studenti, che della sottoclasse impiegati.

Le sottoclassi sottoinsiemi disgiunti sono un gruppo di sottoclassi, specializzazioni della stessa superclasse, con elementi disgiunti, ma la cui unione è un sottoinsieme degli elementi della superclasse. Ad esempio, le sottoclassi matricole e laureandi, sono sottoinsiemi disgiunti della superclasse studenti.

Le sottoclassi partizione di una stessa classe sono fra loro disgiunte, ma l'unione dei loro elementi coincide con gli elementi della superclasse. Ad esempio, le sottoclassi Maschi e Femmine, partizione della superclasse Persone.

CONOSCENZA CONCRETA: ASSOCIAZIONI

Introduciamo le associazioni per definire i collegamenti tra collezioni di entità.

Istanze di associazioni. Un fatto che correla due o più entità, stabilendo un legame logico fra di loro.

Grado dell'associazione. Il numero delle classi coinvolte nell'istanza di associazione.

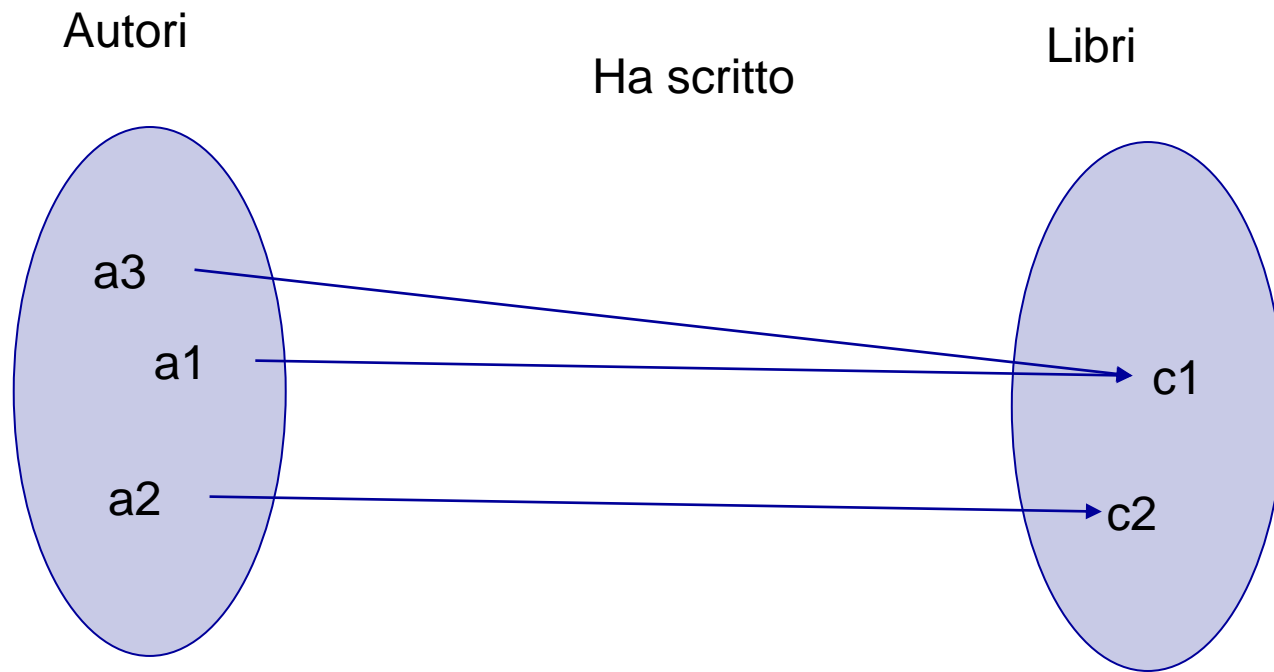
Associazioni. Un insieme di istanze di associazioni fra due classi.

Dominio dell'associazione. Un sottoinsieme del prodotto cartesiano tra due classi:

$$C_1 \times C_2 := \{ (c_1, c_2) \mid c_1 \in C_1, c_2 \in C_2 \}$$

Esempio. Se “Dante” e “Boccaccio” sono gli elementi della classe autori e la “Divina Commedia” e il “Decamerone” sono elementi della classe dei libri, l'associazione “ha scritto” fra le classi autori e libri ha come istanze le coppie (Dante, “Divina Commedia”) e (Boccaccio, “Decamerone”).

Esempio



CONOSCENZA CONCRETA: ASSOCIAZIONI

PROPRIETA' STRUTTURALI DELLE ASSOCIAZIONI

Un'associazione è caratterizzata, oltre che dal suo dominio e dalle caratteristiche delle eventuali proprietà, anche dalle seguenti proprietà strutturali:

- *la molteplicità*
- *la cardinalità*
- *la totalità*

CONOSCENZA CONCRETA: ASSOCIAZIONI

molteplicità

Molteplicità di un'associazione fra X e Y riguarda il numero massimo di elementi di Y che possono trovarsi in relazione con un elemento di X e viceversa. Si dice che l'associazione è

- ❖ **univoca** da X ad Y se ogni elemento di X può essere in relazione con al più un elemento di Y.
- ❖ **multivalore** se non esiste il vincolo di unicità da X a Y

CONOSCENZA CONCRETA: ASSOCIAZIONI

cardinalità

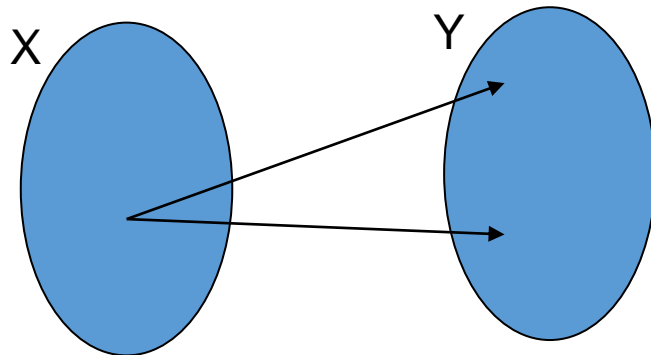
La **cardinalità** di un'associazione fra X e Y descrive contemporaneamente la molteplicità dell'associazione e della sua inversa.

Si dice che la cardinalità è:

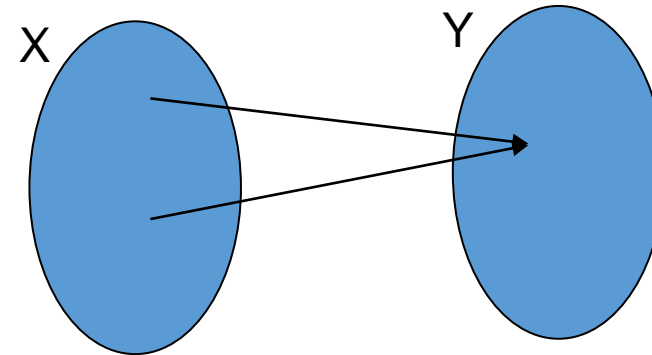
- ❖ **uno a molti** (1:N) se l'associazione è multivalore da X ad Y ed univoca da Y ad X
- ❖ **molti ad uno** (N:1) se l'associazione è univoca da X ad Y e multivalore da Y ad X
- ❖ **molti a molti** (N:M) se l'associazione è multivalore in entrambe le direzioni
- ❖ **uno ad uno** (1:1) se l'associazione è univoca in entrambe le direzioni

PROPRIETA' STRUTTURALI DELLE ASSOCIAZIONI

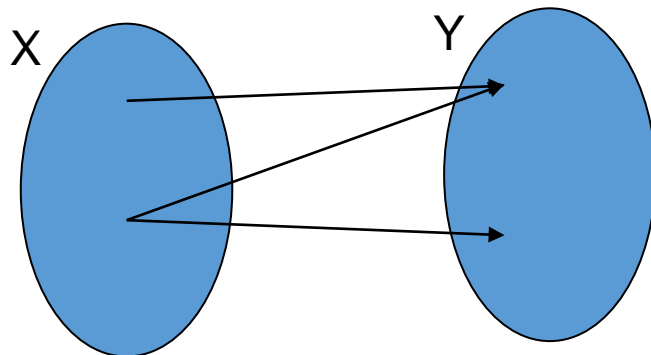
schematizzazione



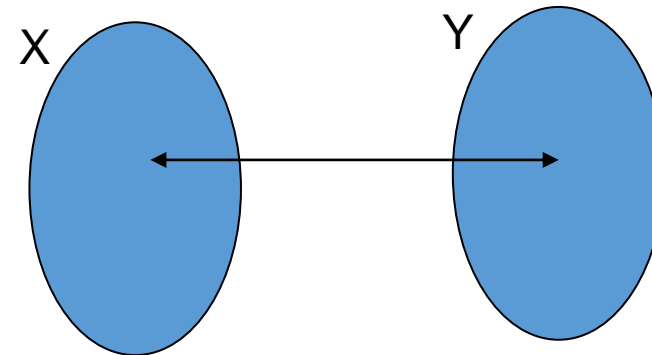
cardinalità è 1 a molti (1:N)



cardinalità è molti a 1 (N:1)



cardinalità è molti a molti (N:M)



cardinalità è 1 a 1 (1:1)

PROPRIETA' STRUTTURALI DELLE ASSOCIAZIONI

esempi

In un universo del discorso popolato da studenti, dipartimenti, corsi del piano di studi e professori, se indichiamo con la notazione $A(C_1, C_2)$ un'associazione A con dominio il prodotto cartesiano di C_1 e C_2 , si ha che

associazione	cardinalità
Frequenta(Studenti, Corsi)	(M:N)
Insegna(Professori, Corsi)	(1:N)
Dirige(Professori, Dipartimenti)	(1:1)

CONOSCENZA CONCRETA: ASSOCIAZIONI

totalità

L'altra proprietà strutturale di un'associazione fra due collezioni X e Y, detta **totalità**, riguarda il numero minimo di elementi di Y che sono associati ad ogni elemento di X.

Se almeno un elemento di Y è associato ad ogni elemento di X, si dice che l'associazione è **totale su X**, e viceversa sostituendo X con Y ed Y con X. Quando non sussiste tale vincolo, si dice che l'associazione è parziale.

Pensando alla notazione insiemistica, nel caso di un'associazione parziale fra le classi X ed Y accade che esistono punti di X che non sono collegati con punti di Y.

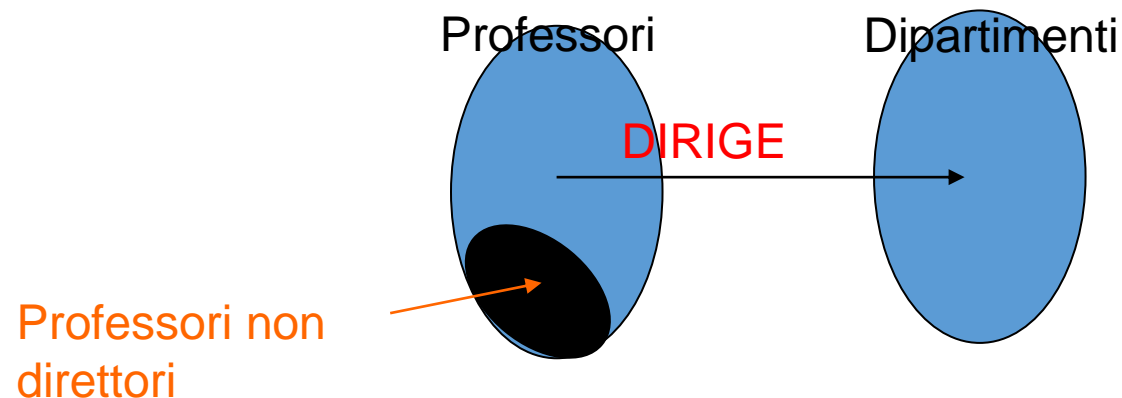
PROPRIETA' STRUTTURALI DELLE ASSOCIAZIONI

esempio

L'associazione

Dirige(Professori, Dipartimenti)

- è totale su Dipartimenti, in quanto ogni dipartimento ha un direttore
- non è totale su Professori, in quanto non tutti i professori sono direttori di dipartimenti.



CONOSCENZA ASTRATTA

La conoscenza astratta riguarda i **fatti generali** che descrivono

1. struttura della conoscenza concreta (collezioni, tipi di entità, associazioni)
2. restrizioni sui valori possibili della conoscenza concreta e sui modi in cui essi possono evolvere nel tempo (**vincoli d'integrità**)
3. regole per derivare nuovi fatti da altri noti.

CONOSCENZA ASTRATTA

Vincoli di integrità

Un **vincolo d'integrità** può essere:

statico: se riguarda ogni singolo stato dell'universo del discorso,

dinamico: se riguarda le transizioni da uno stato ad un altro.

CONOSCENZA ASTRATTA

Vincoli di integrità: statici

I **vincoli d'integrità statici** definiscono delle condizioni sui valori della conoscenza concreta che devono essere soddisfatte indipendentemente da come evolve l'universo del discorso.

Le condizioni possono riguardare:

- **i valori di una proprietà.**
Ad esempio, un utente ha le proprietà codice fiscale, cognome, residenza, con valori di tipo stringa di caratteri alfanumerici e anno di nascita, con valori di tipo intero;
- **i valori di proprietà diverse di una stessa entità.**
Ad esempio, per ogni impiegato le trattenute sulla paga devono essere inferiori ad un quinto dello stipendio;
- **i valori di proprietà di entità diverse di uno stesso insieme.**
Ad esempio, le matricole degli studenti sono tutte diverse;

CONOSCENZA ASTRATTA

Vincoli di integrità: statici

Un attributo, o un insieme di attributi, è detto **chiave** rispetto ad una classe di elementi, se i suoi valori identificano univocamente un elemento della collezione, e se ogni attributo della chiave è necessario a questo fine.

Un esempio di chiave in persone è il codice fiscale.

In generale possono esistere più chiavi per una classe di entità, come accade per la classe degli studenti che hanno come chiave non solo il codice fiscale ma anche la matricola. In questi casi se ne sceglie una come principale e viene chiamata chiave primaria

CONOSCENZA ASTRATTA

Vincoli di integrità: dinamici

I vincoli d'integrità dinamici definiscono delle condizioni sul modo in cui la conoscenza concreta può evolvere nel tempo.

Ad esempio, una persona coniugata non potrà cambiare stato civile ritornando ad essere scapolo o nubile; una data di nascita non può essere modificata.

In conclusione, mentre un vincolo statico riguarda ogni singolo stato dell'universo del discorso, un vincolo dinamico riguarda le transizioni da uno stato ad un altro.

Sistemi per il recupero delle informazioni

PARTE - III

COME SI MODELLA

INTRODUZIONE

Nella costruzione di un modello informatico,

- ❑ prima si “**definisce**” il modello, descrivendo
 - conoscenza concreta
 - conoscenza astratta,
- ❑ poi si “**costruisce**” la rappresentazione della conoscenza concreta.

ESEMPIO

Per costruire un modello informatico per la gestione di informazioni sui libri, prima si devono definire quelle che interessano ai fini dell'applicazione (titolo, autore, editore ecc.).

Una volta definite le proprietà interessanti comuni a tutti i possibili libri, si passa a costruire per ogni entità “libro” della realtà oggetto di studio una rappresentazione nel modello informatico.

IL MODELLO

Per la definizione del modello si possono usare diversi tipi di formalismi, che si differenziano per il “*modello dei dati*” che supportano, cioè per meccanismi di astrazione offerti per rappresentare la realtà.

Nel seguito si presentano due tipi di modelli dei dati: il

- **modello a oggetti**, usato come esempio di formalismo per la **progettazione** di una basi di dati
- **modello relazionale**, usato come esempio di formalismo per la **realizzazione** di una base di dati.

IL MODELLO DEI DATI A OGGETTO

Per rappresentare in maniera naturale e diretta l'idea che il progettista si fa del mondo osservato, il **modello dei dati a oggetti** prevede i seguenti meccanismi d'astrazione:

- **oggetto**
- **tipo di oggetto**
- **classe**
- **gerarchie fra classi**

Verranno dati esempi di utilizzo di questi meccanismi tramite un formalismo grafico che serva a definire lo schema di una base di dati, ovvero la struttura della conoscenza concreta, che chiameremo **schema concettuale**.

Il formalismo grafico viene chiamato diagramma **entità-relazione**

IL MODELLO DEI DATI A OGGETTO

questioni terminologiche

dominio del discorso
entità
tipo entità
collezione
associazione

modello informatico (ER)
oggetto (entity instance)
tipo oggetto (entity type)
classe (entity)
associazione o relazione

OGGETTO E TIPO DI OGGETTO

Un **oggetto** è un'entità **software** con **stato** ed **identità**, che modella un'entità dell'universo del discorso.

- . Lo **stato** è costituito da un insieme di campi, che
 - sono valori costanti o variabili associati ad un **nome**, detto **attributo**.
 - possono assumere valori di qualsiasi complessità
 - modellano le proprietà dell'entità

Come accade per le proprietà delle entità, un attributo di un oggetto può avere valori di tipo atomico o strutturato, semplice o unione, univoco o multivalore, totale o parziale, costante o modificabile.

OGGETTO E TIPO DI OGGETTO - Esempio

Nel caso della biblioteca, alcuni possibili attributi degli elementi della classe Utenti sono CodiceFiscale, Cognome, Residenza, AnnoDiNascita. I primi due sono associati a valori di tipo stringa (sequenza di caratteri alfanumerici), il terzo strutturato, il quarto associato a valori di tipo intero.

Ogni oggetto è un valore di un tipo che specifica la struttura di un insieme di possibili oggetti, ovvero quali sono gli attributi e il tipo dei valori che possono assumere.

CLASSI E ASSOCIAZIONI

La **classe** modella un insieme di oggetti dello stesso tipo del mondo reale

- è costruita col processo di astrazione detto classificazione: entità diverse vengono unificate per costituire un nuovo concetto. Si astrae dalle differenze fra le entità per evidenziare ciò che le accomuna.
- tutti gli oggetti di una classe sono caratterizzati dagli stessi attributi.

L' **associazione** modella una correlazione fra entità del mondo reale

- è una corrispondenza tra classi, e stabilisce una correlazione logica fra oggetti appartenenti a classi diverse (o anche alla stessa classe).

CONCETTI IMPORTANTI

dominio di un attributo: è l'insieme dei valori che può assumere l'attributo stesso

- esempio: il dominio dell'attributo mese è costituito dall'insieme {gennaio, febbraio, marzo, ...,dicembre} oppure dall'insieme {01, 02, 03, ...,12}

chiave di una classe: è un attributo (o un insieme di attributi) che identifica univocamente gli oggetti di una classe

- K è chiave della classe C se non esistono due oggetti in C che hanno lo stesso valore di K
- esempio: l'attributo matricola è chiave della classe STUDENTI

DIAGRAMMI E-R

Nel **formalismo grafico** che si adotta, una **classe** si rappresenta con un rettangolo etichettato con il nome della classe.

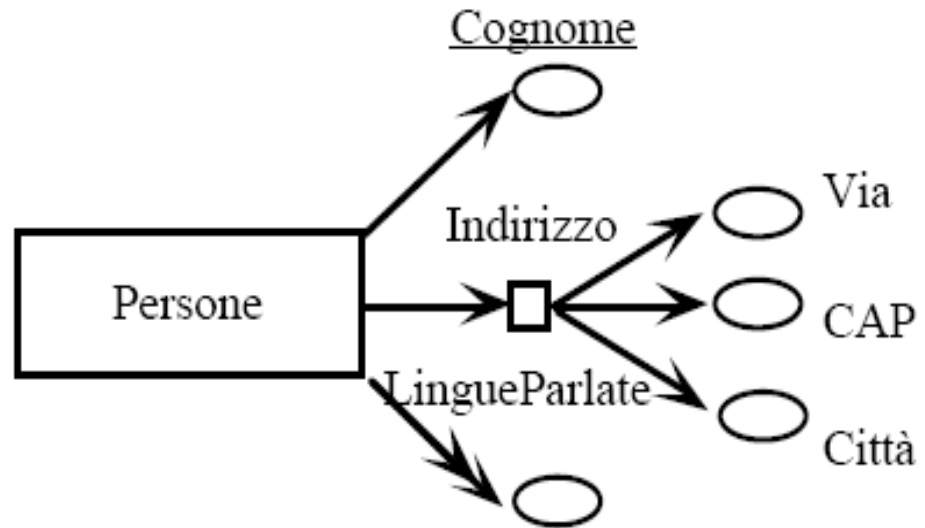
Per gli **attributi** si usano le seguenti convenzioni:

1. attributi con valori atomici sono rappresentati con ovali collegati alla classe da un arco che termina con una freccia singola, se l'attributo è univoco, oppure con una freccia doppia, se l'attributo è multivalore;
2. se l'attributo ha valori strutturati, l'ovale si sostituisce con un quadratino dal quale escono archi verso ovali in numero pari ai campi del valore strutturato.
3. se l'attributo ha valori unione, si procede come nel caso precedente sostituendo il quadratino con un cerchietto;
4. proprietà che possono avere valori non specificati si rappresentano con un taglio sull'arco;
5. gli attributi della chiave primaria si sottolineano.

Esempi

Una classe con gli attributi.

Cognome e' la chiave primaria



Una classe



DESCRITTORE DI CLASSI

Poiché di solito i tipi degli oggetti hanno molti attributi, per non appesantire la rappresentazione grafica, si preferisce descriverli separatamente con un cosiddetto **descrittore di classe**, in cui si specificano il nome della classe, il nome del tipo degli oggetti e poi si elencano gli attributi e i tipi dei loro valori.

Per descrivere i tipi dei valori useremo la seguente notazione:

1. tipi primitivi: integer, real, bool e string,
2. tipi record: insieme di coppie “Attributo :Tipo del valore”, quanti sono i campi del record, separate da un punto e virgola e racchiuse fra quadre.
3. tipi unione: insiemi di coppie “Attributo :Tipo del valore”, quante sono le alternative, separate da un punto e virgola e racchiuse fra parentesi tonde.
4. tipo sequenza di valori di un tipo T, indicato come “seq T”
5. valori opzionali di tipo T si indicano con “optional T”.

DESCRITTORE DI CLASSI: esempio

Classe Persone

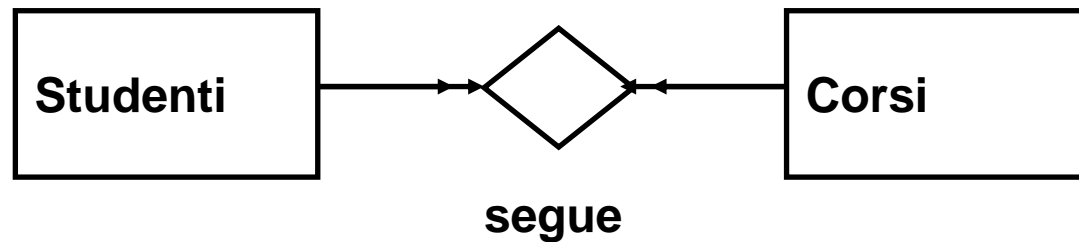
Tipo oggetto Persona

Attributi Cognome :string;
 Indirizzo :[Via :string; Numero :integer; Città :string];
 LingueParlate :seq string;
 Sesso :(M; F)

ASSOCIAZIONI

Un'associazione binaria tra classi si rappresenta con un rombo collegato con degli archi alle classi associate.

Il rombo è etichettato con il nome dell'associazione

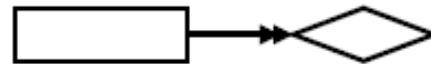


ASSOCIAZIONI: caratteristiche

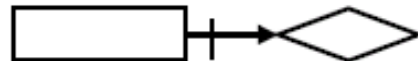
- **univocità**: ad un oggetto di una classe può essere associato un solo oggetto dell'altra classe
- **totale**: ad un oggetto di una classe è sempre associato qualche oggetto dell'altra classe



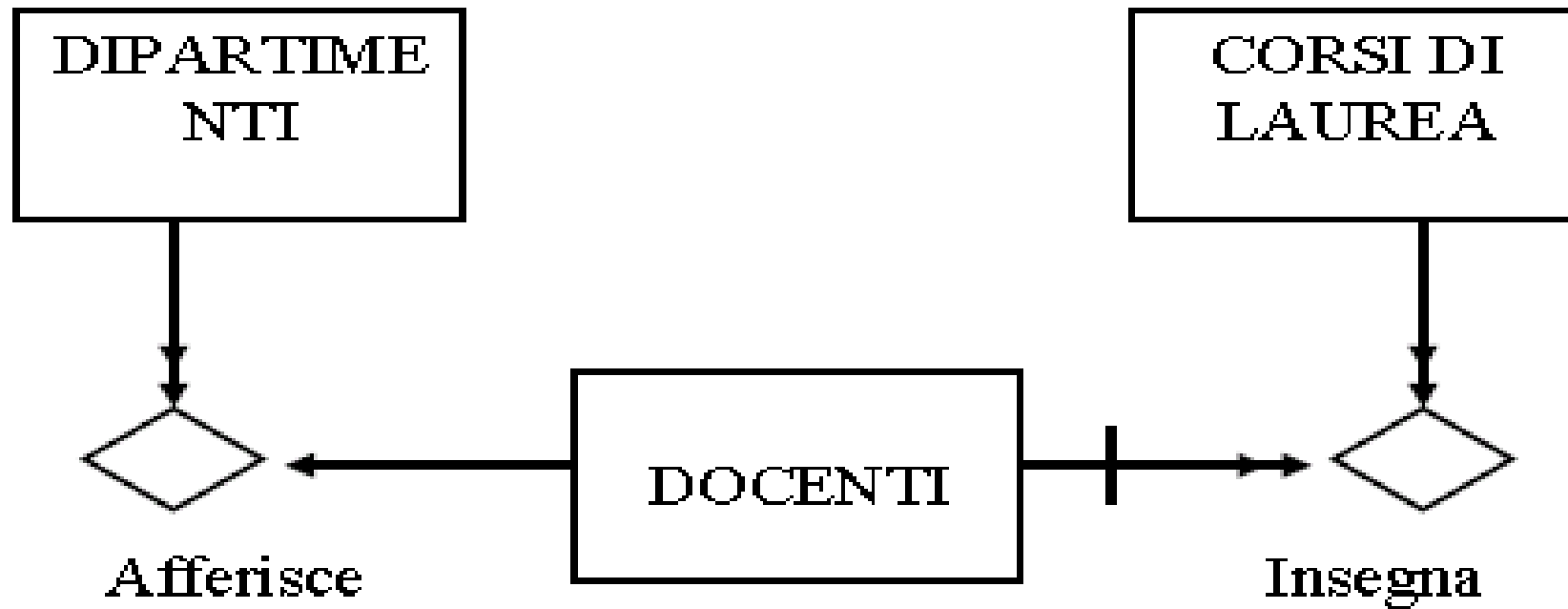
- **multivalore**: ad un oggetto di una classe possono essere associati più oggetti dell'altra classe



- **parziale**: ad un oggetto di una classe non è sempre associato qualche oggetto dell'altra classe



ASSOCIAZIONI: esempio

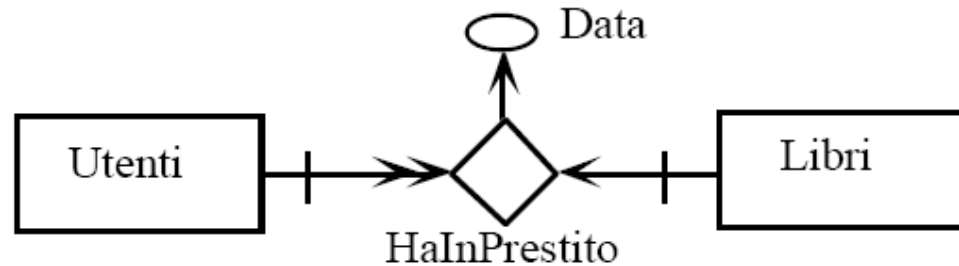


ASSOCIAZIONI: esempio

- Afferisce è **univoca** da Docenti a Dipartimenti: ad un oggetto della classe Docenti è associato un solo oggetto della classe Dipartimenti; si modella il fatto che un docente possa afferire ad un solo dipartimento
- Afferisce è **multivalore** da Dipartimenti a Docenti: ad un oggetto della classe Dipartimenti possono essere associati più oggetti della classe Docenti; si modella il fatto che ad un dipartimento possano afferire molti docenti
- Insegna è **totale** da Corsi di Laurea a Docenti: ad un oggetto della classe Corsi di Laurea è sempre associato qualche oggetto della classe Docenti; si modella il fatto che in un corso di laurea debbano necessariamente insegnare dei docenti
- Insegna è **parziale** da Docenti a Corsi di Laurea: ad un oggetto della classe Docenti non è sempre associato qualche oggetto della classe Corsi di Laurea ; si modella il fatto che un docente potrebbe non insegnare

ASSOCIAZIONI: proprietà

Se l'associazione binaria ha delle proprietà, si aggiungono degli archi uscenti dal rombo etichettati con il nome della proprietà.

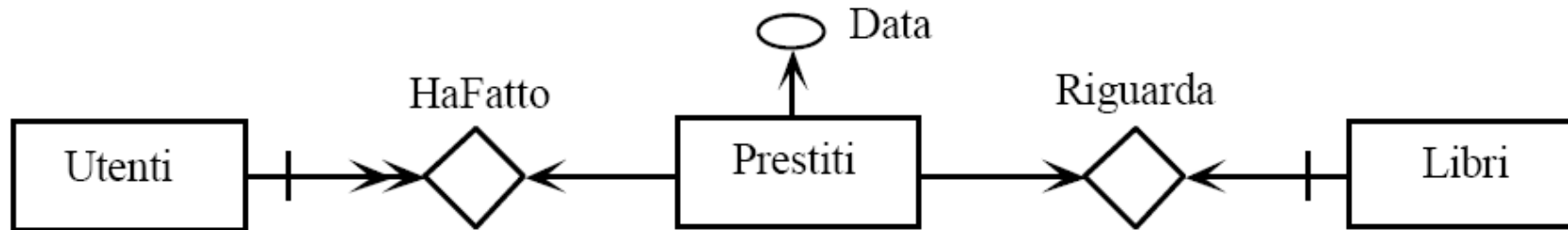


In Figura è mostrato un esempio di un'associazione tra i libri di una biblioteca e gli utenti, che modella i prestiti, e che ha una proprietà "Data".

ASSOCIAZIONI: proprietà

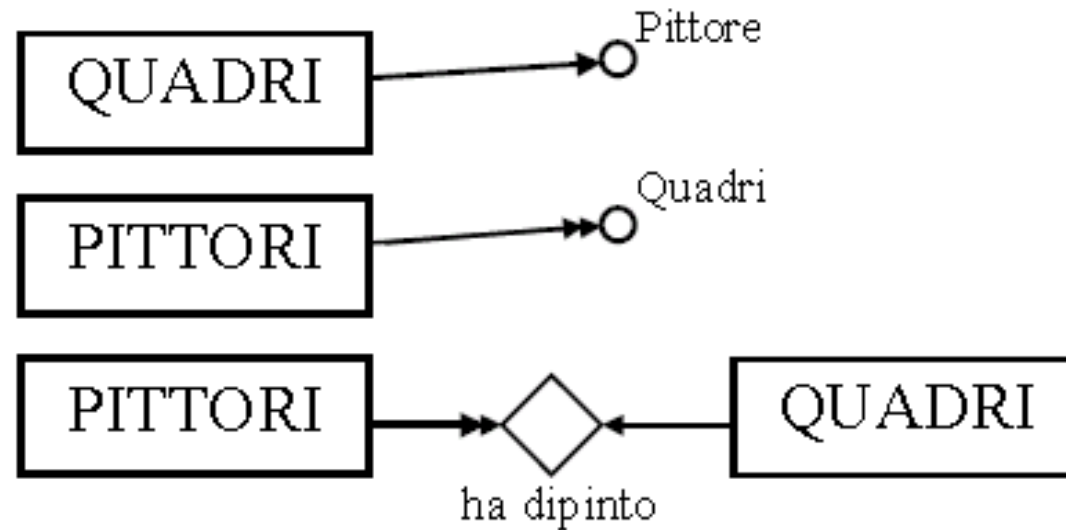
Tuttavia a volte è utile introdurre una nuova classe liberando l'associazione dalla proprietà, e inglobandola nella nuova classe.

Un'associazione con proprietà, come quella tra i libri di una biblioteca e gli utenti della figura precedente, può essere modellata interpretando un'istanza di associazione come un'entità e definendo così una classe Prestiti, associata in modo (1:1) ai Libri e in modo (N:1) agli Utenti, e aggiungendo un attributo "Data" alla classe Prestiti stessa.



Considerazioni - I

- Decidere che cosa modellare con una classe e cosa con una proprietà e quando usare associazioni piuttosto che proprietà è una scelta di progetto e dipende sia dall'interpretazione del progettista che dalla finalità dell'applicazione informatica.



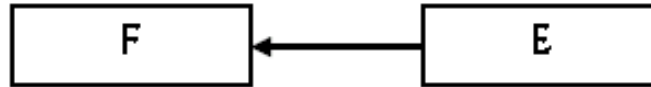
Considerazioni - II

Tutti e tre gli schemi precedenti rappresentano il fatto che un pittore ha dipinto certi quadri

1. L'enfasi è posta sui quadri; di essi si possono rappresentare altre proprietà (dimensioni, tipo, epoca, prezzo,...) mentre dei pittori non si possono rappresentare proprietà
2. L'enfasi è posta sui pittori; di essi si possono rappresentare altre proprietà (nazionalità, nato il, morto il, ...) mentre dei quadri non si possono rappresentare proprietà
3. In questo caso, il più espressivo, si possono rappresentare sia proprietà dei quadri che dei pittori.

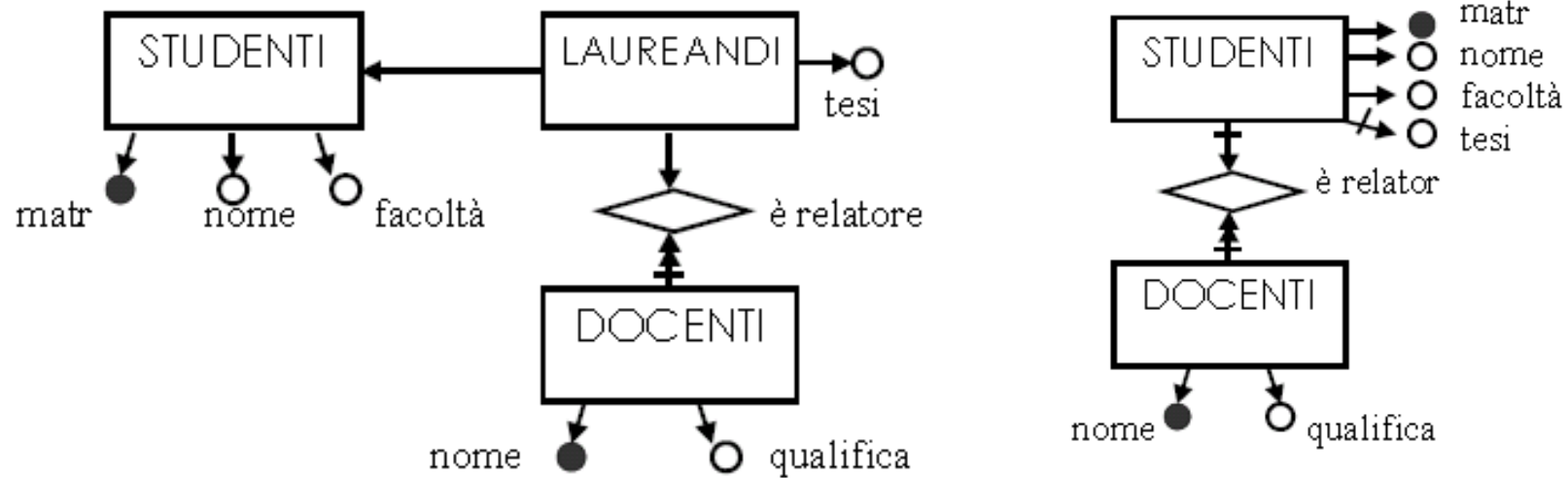
GERARCHIA FRA CLASSI

- Tra due classi E ed F può essere stabilita una **gerarchia**



- E è detta **sottoclasse** o specializzazione di F
- F è detta **superclasse** o generalizzazione di E
- Proprietà delle gerarchie
 - **Vincolo di struttura**: se *E* è sottoclasse di *F*, *E* ha tutti gli attributi di *F*, e partecipa a tutte le associazioni cui partecipa *F* (ereditarietà). *E* può avere altri attributi, e partecipare ad altre associazioni
 - **Vincolo di insieme**: se *E* è specializzazione di *F*, ogni oggetto di *E* è anche un oggetto di *F* (cioè *E* è un sottoinsieme di *F*)

GERARCHIA FRA CLASSI: esempio



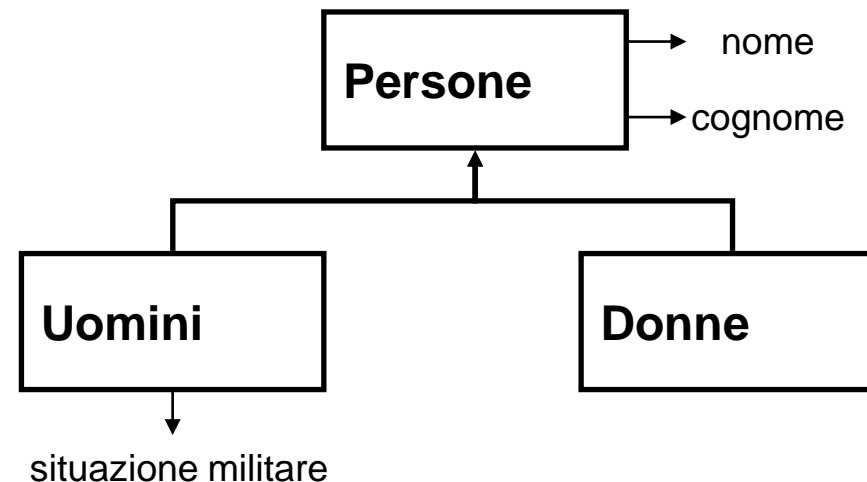
Lo schema a sinistra e' piu' espressivo di quello di destra

GERARCHIA FRA CLASSI: ereditarietà

Gli attributi nelle gerarchie vengono ereditati dalle entità superiori a quelle inferiori.

Normalmente l'eredità tra tipi si usa solo per definire sottotipi (ereditarietà stretta); in questo caso:

- gli attributi possono essere solo aggiunti
- gli attributi possono essere ridefiniti solo specializzandone il tipo



GERARCHIA FRA CLASSI: tipi di sottoclassi

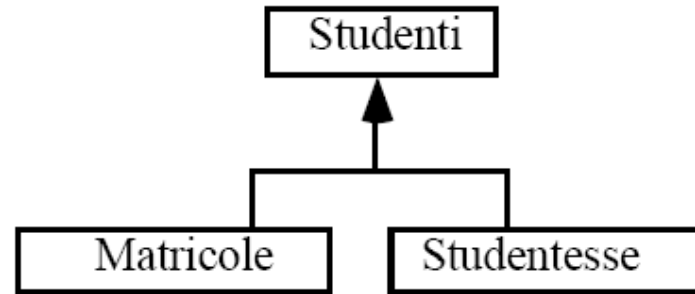
Quando si definiscono più sottoclassi di una stessa classe, su questo insieme di sottoclassi possono essere definiti i seguenti vincoli:

1. un insieme di sottoclassi soddisfa il vincolo di disgiunzione se ogni coppia di sottoclassi in questo insieme è disgiunta, ovvero è priva di elementi comuni (**sottoclassi disgiunte**);
2. un insieme di sottoclassi soddisfa il vincolo di copertura se l'unione degli elementi delle sottoclassi coincide con l'insieme degli elementi della superclasse (**sottoclassi copertura**).

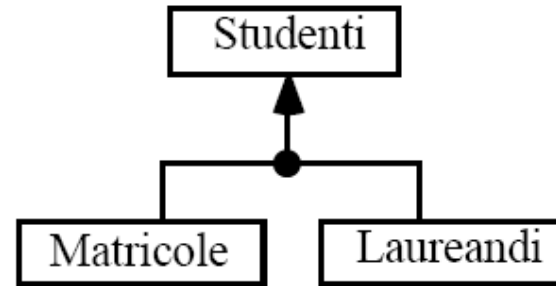
Quando i due vincoli sono entrambi soddisfatti, l'insieme di sottoclassi costituisce una **partizione** della superclasse.

Il vincolo di disgiunzione viene rappresentato con il pallino nero, mentre il vincolo di copertura viene rappresentato con una freccia doppia verso la superclasse.

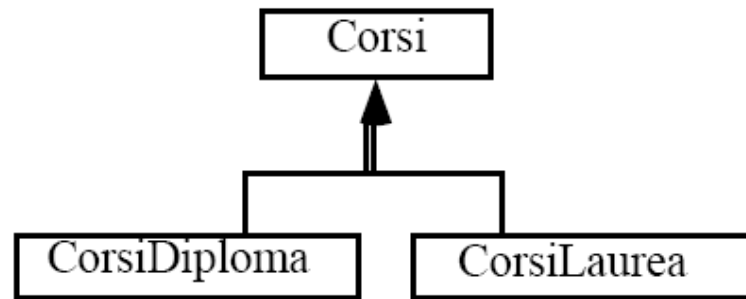
GERARCHIA FRA CLASSI: tipi di sottoclassi



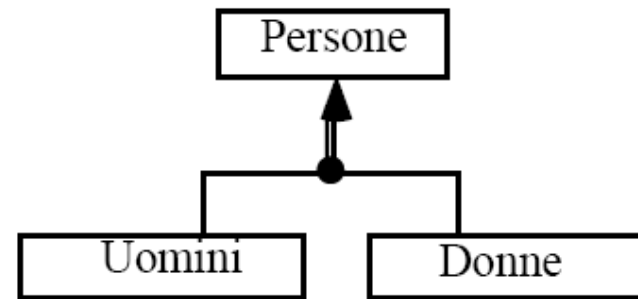
Sottoclassi scorrelate



Sottoclassi disgiunte



Sottoclassi copertura



Sottoclassi partizione

DIAGRAMMI E-R: esempio

A titolo di esempio, mostriamo, ad un primo livello di dettaglio, la rappresentazione con il formalismo grafico di alcuni fatti riguardanti una **biblioteca universitaria**: descrizioni bibliografiche, libri, autori, utenti e prestiti.

Delle entità interessano le seguenti proprietà:

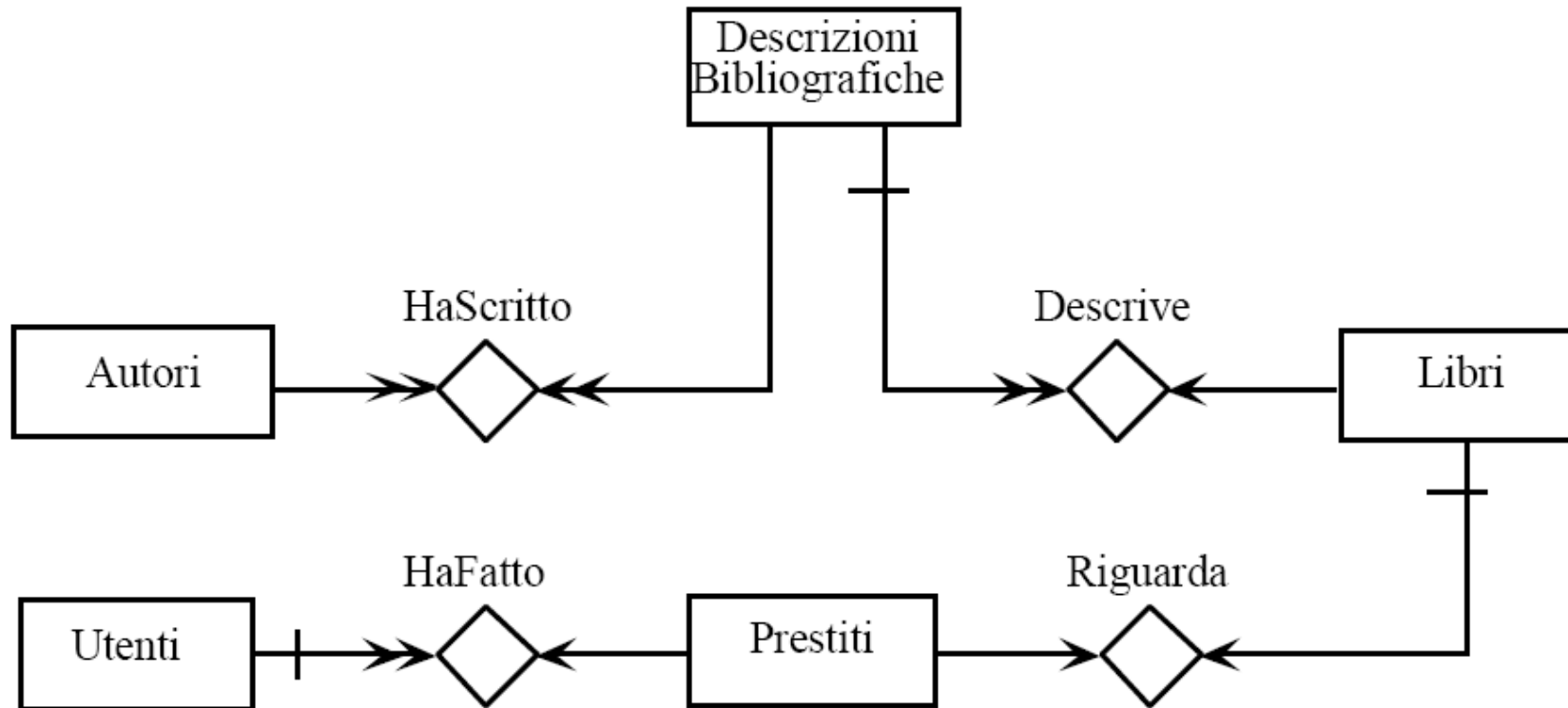
1. Di una **descrizione bibliografica** interessano il codice, il titolo dell'opera, l'editore, l'anno di pubblicazione e un insieme di termini usati per la classificazione del contenuto dell'opera.
2. Di un **libro** interessano la collocazione e il numero della copia.
3. Di un **autore** interessano il nome e cognome, la nazionalità, la data di nascita
4. Di un **utente** interessano il nome, il cognome, l'indirizzo e i recapiti telefonici.
5. Di un **prestito** interessano la data del prestito e la data di restituzione.

DIAGRAMMI E-R: esempio

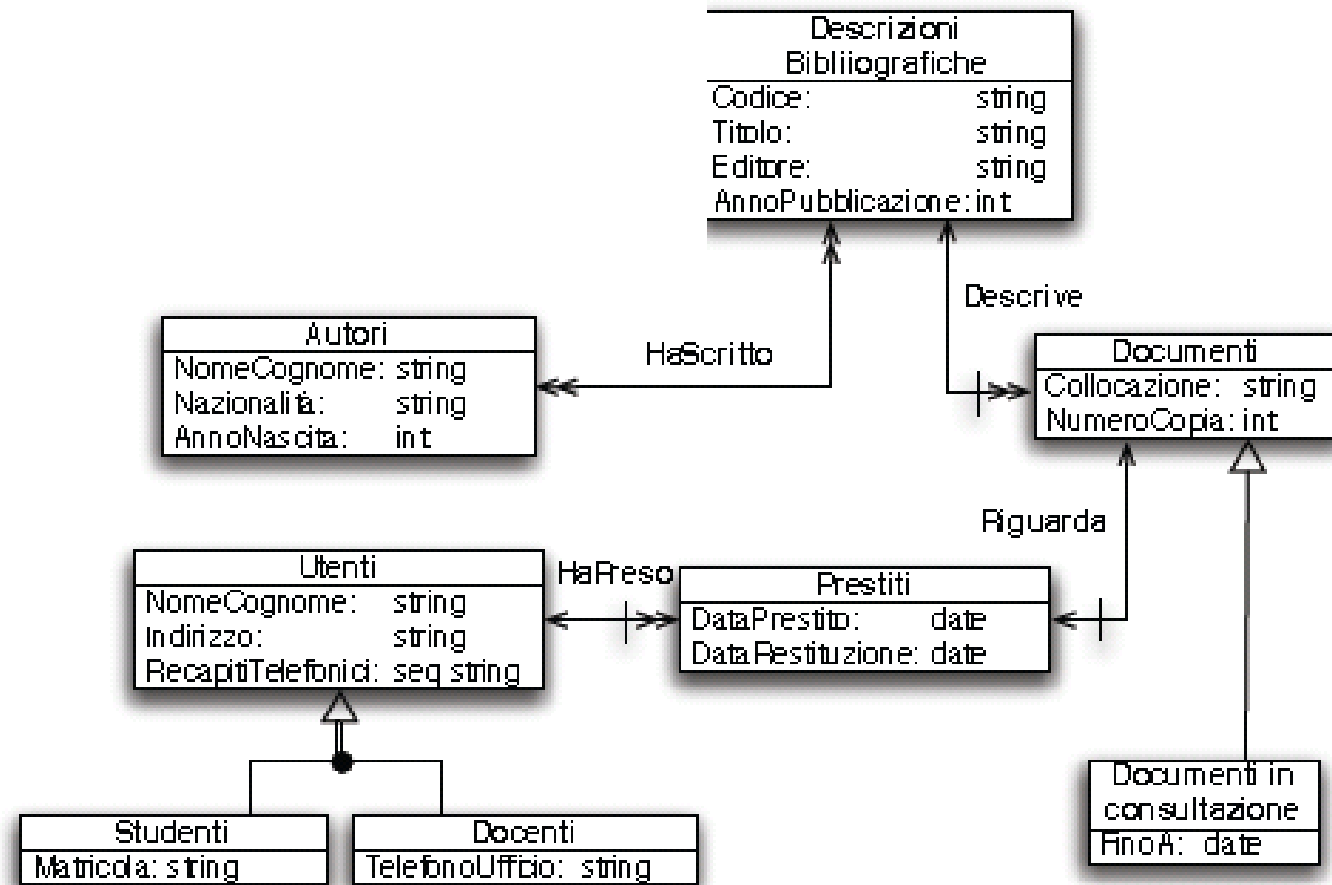
Le associazioni interessanti sono:

1. **HaScritto** (N:M) tra autori e descrizioni bibliografiche, che collega un autore con le descrizioni bibliografiche delle opere che ha scritto. Ogni autore ha scritto almeno un libro e ogni descrizione bibliografica riguarda almeno un autore;
2. **Describe** (N:1) tra descrizioni bibliografiche e libri, che collega una descrizione bibliografica alle copie dei libri presenti in biblioteca. Ogni libro ha una descrizione bibliografica e ogni descrizione bibliografica descrive una o più copie di libri, supporremo inoltre che possa descrivere anche un libro ordinato ma non ancora acquisito dalla biblioteca;
3. **HaFatto** (N:1) tra utenti e prestiti, che collega gli utenti ai prestiti che ha fatto e che non sono ancora scaduti. Ogni utente della biblioteca può avere nessuno, uno o più prestiti, ma un prestito ha sempre associato l'utente che lo ha fatto;
4. **Riguarda** (1:1) tra prestiti e libri, che collega i prestiti alle copie dei libri interessati. Una copia di un libro può essere coinvolta in al più un prestito e un prestito riguarda una copia di un libro.

DIAGRAMMI E-R: esempio



DIAGRAMMI E-R: esempio



Sistemi per il recupero delle informazioni

PARTE - IV

COME SI PROCEDE

INTRODUZIONE

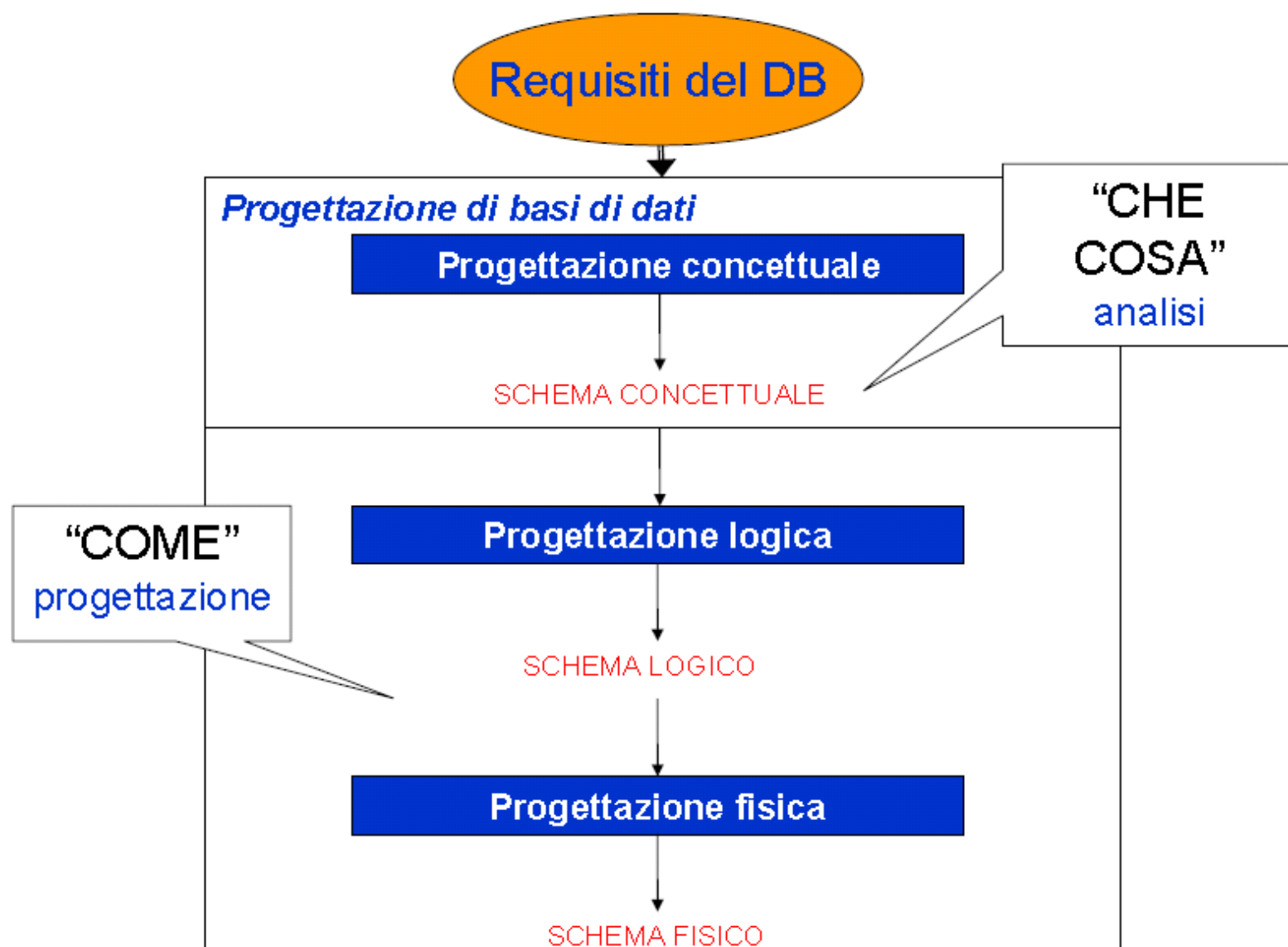
Finora si è visto come definire lo schema concettuale di una base di dati.

Nella pratica non è così semplice ed occorre un lungo procedimento di analisi e studio della situazione da modellare per poter produrre una progettazione concettuale della base di dati e poi una sua realizzazione.

Per dare un'idea di come si procede, si considera una tipica metodologia a più fasi in cui gli aspetti del problema vengono considerati gradualmente per ottenere una realizzazione soddisfacente.

Quattro fasi:

- 1. analisi dei requisiti**
- 2. progettazione concettuale**
- 3. progettazione logica**
- 4. progettazione fisica**



ANALISI DEI REQUISITI

Scopo dell'analisi dei requisiti è la definizione dei bisogni informativi del committente.

Il progettista deve capire di “cosa si parla”. Poi si passa ad un'analisi del problema per raccogliere una descrizione dei bisogni informativi e formulare la cosiddetta specifica dei requisiti in linguaggio naturale.

Quando il problema è di limitata complessità, e basta interagire con una sola persona, il procedimento può essere relativamente veloce, ma quando il problema è complesso e sono coinvolte persone diverse, il procedimento diventa lungo e comporta un lavoro di unificazione dei concetti coinvolti.

Lo scopo dell'analisi dei requisiti è, in altre parole, proprio quello di chiarire la corretta interpretazione dei fatti descritti, riformulando la specifica in modo chiaro.

Esempio

Si vuole progettare una base di dati per gestire informazioni su musei, le opere conservate, gli artisti che hanno creato tali opere.

Di un museo interessano il nome, che lo identifica, la città, l'indirizzo e il nome del direttore.

Un artista è identificato dal nome; di lui interessano la nazionalità, la data di nascita, la eventuale data di morte.

Di un'opera, identificata da un codice, interessano, l'anno di creazione, il titolo, il nome dei personaggi rappresentati.

Un'opera può essere un dipinto od una scultura; se è un dipinto interessano il tipo di pittura e le dimensioni; se è una scultura interessano il materiale, l'altezza ed il peso.

PROGETTAZIONE CONCETTUALE

Scopo della progettazione concettuale è di tradurre la specifica dei requisiti in un progetto della struttura concettuale dei dati descritta utilizzando un formalismo grafico.

Lo schema concettuale si definisce procedendo con i seguenti passi:

1. **identificazione delle classi;**
2. **descrizione delle associazioni fra le classi;**
3. **definizione di sottoclassi;**
4. **definizione delle proprietà degli elementi delle classi.**

PROGETTAZIONE CONCETTUALE

Identificazione delle classi

Si produce una lista preliminare delle classi di oggetti che interessa modellare e si assegna ad ognuna di esse un nome appropriato. Questo elenco iniziale ha un grado di completezza e di significatività che dipende dal grado di comprensione del problema e, in generale, sarà soggetto a modifiche mano a mano che si procede.

PROGETTAZIONE CONCETTUALE

Descrizione delle associazioni fra le classi

Si individuano le possibili associazioni fra le classi finora definite e le loro proprietà strutturali. L'analisi delle associazioni può portare ad eliminare una classe che può essere rappresentata da un'associazione, o ad aggiungere una nuova classe per rappresentare un'associazione.

PROGETTAZIONE CONCETTUALE

Definizione di sottoclassi

Per definire le sottoclassi si esaminano tutte le classi già definite per capire

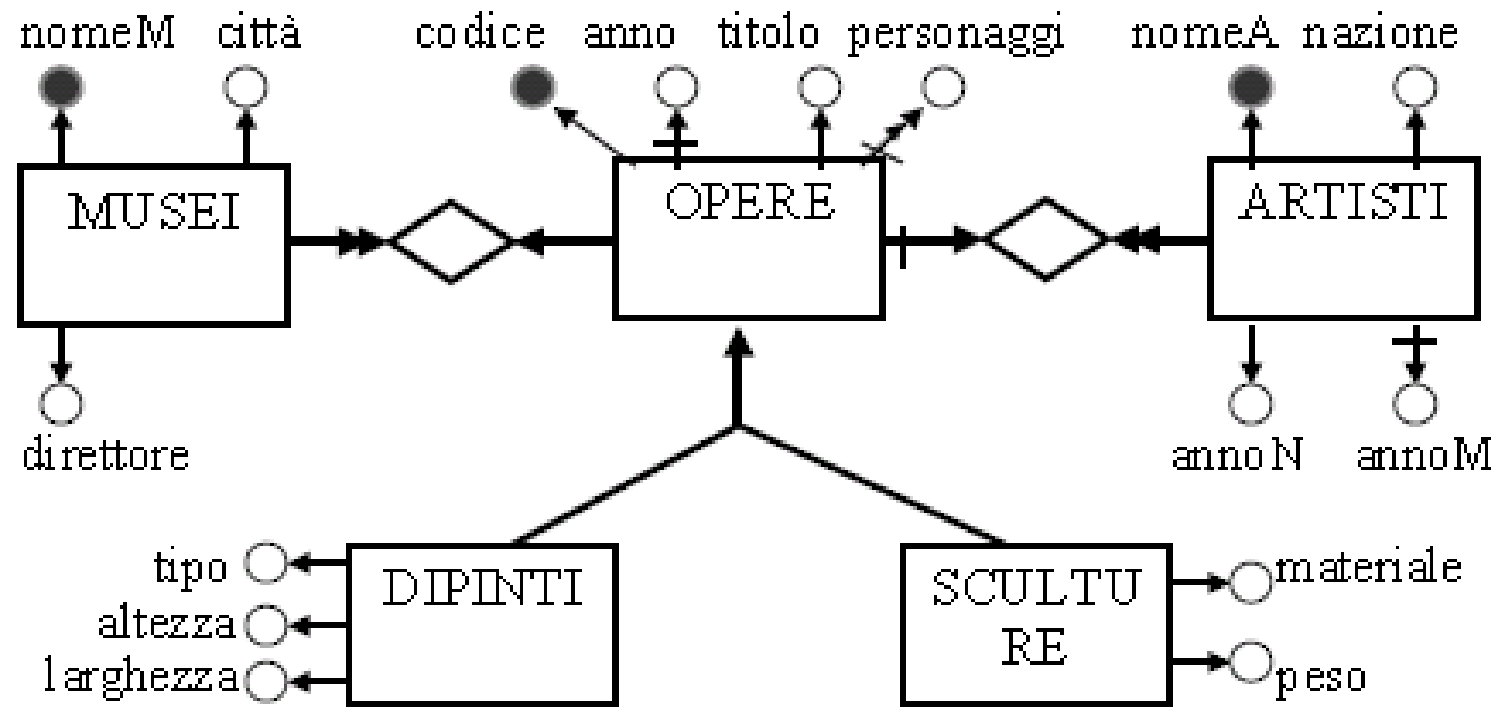
- 1. se può essere utile definirne di nuove per caratterizzare particolari sottoinsiemi di alcune classi**
- 2. se esistono classi che sono un sottoinsieme di altre e quindi possono essere ridefinite**
- 3. se esistono oggetti di classi che possono assumere nel tempo stati significativi per l'applicazione**

PROGETTAZIONE CONCETTUALE

Definizione delle proprietà degli elementi delle classi

Per ogni tipo di oggetto si elencano le proprietà interessanti, specificando, per ognuna di esse, il nome e il tipo. In questo passo va prestata molta attenzione alla possibilità se convenga introdurre nuove classi, o viceversa eliminarne alcune sostituendole da attributi.

Esempio



PROGETTAZIONE LOGICA

Scopo della terza fase della metodologia, la progettazione logica, è di tradurre lo schema concettuale nello schema logico espresso nel modello dei dati del sistema scelto per la realizzazione della base di dati.

Esempio

MUSEI (NomeM, Città, Indirizzo, Direttore)

ARTISTI (NomeA, Nazionalità, DataN, DataM:optional)

OPERE (Codice, Anno, Titolo, NomeM*, NomeA*)

PERSONAGGI (Personaggio, Codice*)

DIPINTI (Codice*, Tipo, Larghezza, Altezza)

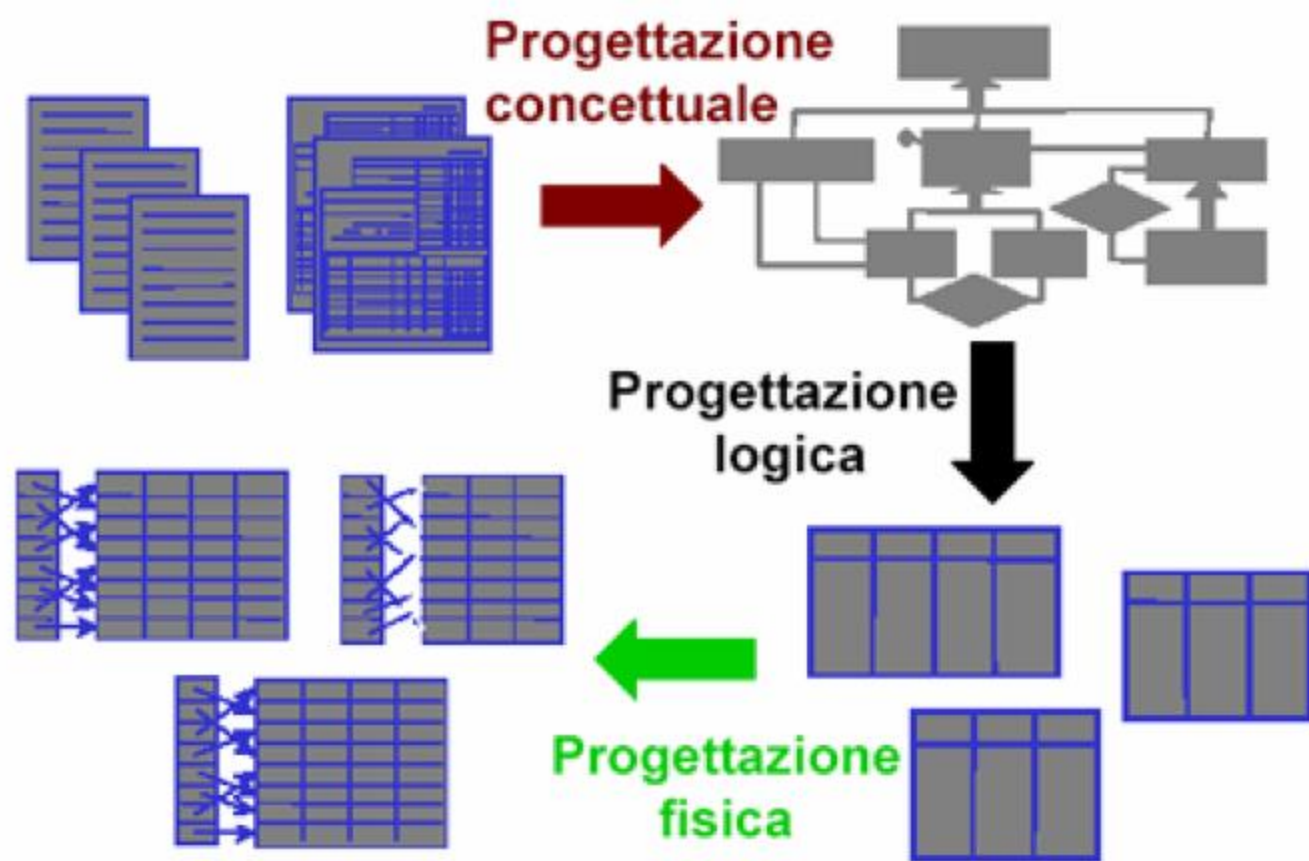
SCULTURE (Codice*, Materiale, Altezza, Peso)

Progettazione fisica

Consiste nel completamento, o modifica, dello schema logico in funzione della organizzazione fisica dei dati ed i meccanismi per operare su di essi.

Realizzazione

Consiste nel costruire la rappresentazione dei fatti specificati nel progetto (schema dei dati, procedure), utilizzando strumenti informatici (linguaggi di programmazione, SGBD)

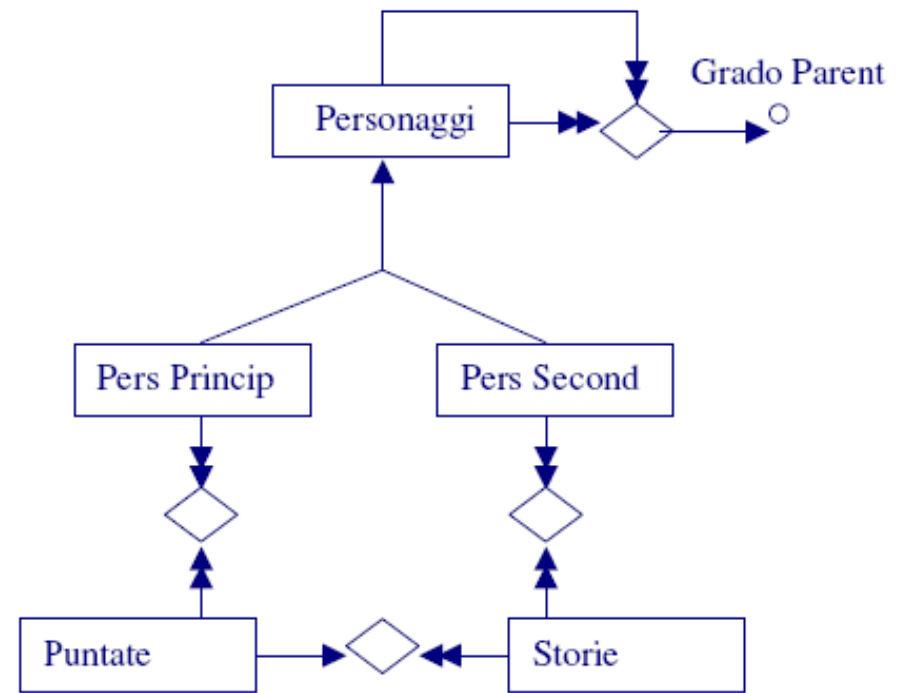


9.Fumetti

Una rivista periodica di fumetti vuole memorizzare informazioni relative a tutte le storie che ha pubblicato nel passato, ed ai relativi personaggi.

Di una storia interessa il titolo, che la identifica, ed interessano informazioni relative alle puntate in cui è stata divisa: per ogni puntata interessa il numero di pagine, il numero d'ordine all'interno della storia (prima, seconda...) ed il numero della rivista su cui è stata pubblicata.

I personaggi si dividono in principali e secondari. Per tutti i personaggi interessa il nome, che li identifica. Per i personaggi secondari interessa ricordare le storie in cui sono apparsi, mentre per quelli principali si vogliono memorizzare precisamente le puntate di apparizione. Se due personaggi sono parenti, se ne memorizza la relazione di parentela (ovvero, il fatto che sono parenti ed anche il grado di parentela).



ESEMPI - I

Si vogliono trattare informazioni su attori e registi di film.

Di un attore o un regista interessano:

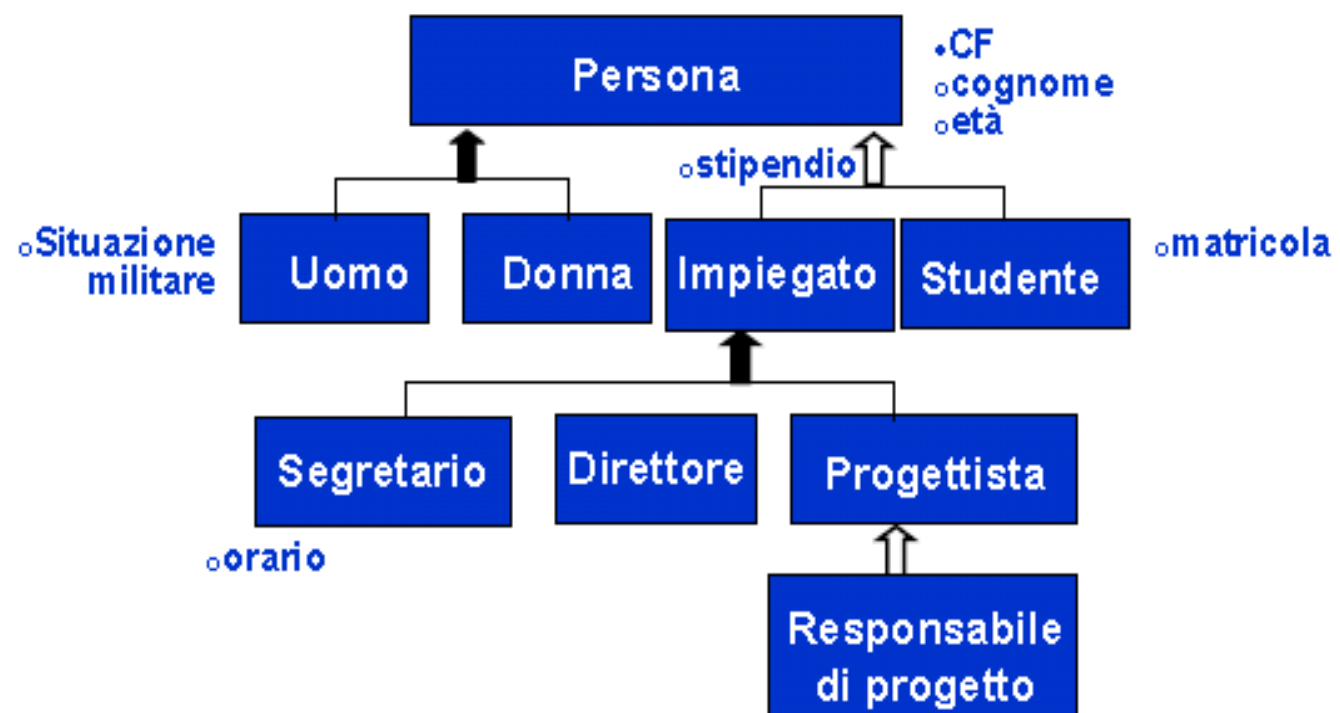
- **il nome, che lo identifica,**
- **l'anno di nascita**
- **la nazionalità**
- **un attore può essere anche un regista**

Di un film interessano

- **il titolo**
- **l'anno di produzione**
- **gli attori**
- **il regista**
- **il produttore**

Esempio - II

- 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)



Esempio - III

- Un impiegato ha codice, cognome ed uno stipendio; può partecipare ad un progetto definito da un budget e da un nome; l'impiegato può dirigere un solo dipartimento definito da un nome e dal telefono ma può afferire in date differenti a più dipartimenti; il dipartimento ha una sede composta da un indirizzo e da una città

