

# Sommario

## 1 Introduzione alle Basi di Dati

### Base di dati

- Insieme organizzato di dati utilizzati per lo svolgimento di determinate attività (di un ente, azienda, ufficio, persona)
- Una risorsa integrata, condivisa fra i vari settori di una organizzazione

### Struttura di un sistema informatico

- Interfaccia utente
- Applicazioni
- Software di ambiente e di sistema
- Basi di dati
- Software di base
- Hardware centralizzato e distribuito
- Sistema di comunicazione di rete

**Ridondanza**   Informazioni ripetute

**Rischio di incoerenza**   Le versioni possono non coincidere

**Data Base Management System, Sistema di Gestione di Basi di Dati (DBMS)** Sistema (prodotto software) in grado di gestire collezioni di dati che siano (anche):

- **Grandi**, di dimensioni (molto) maggiori della memoria centrale dei sistemi di calcolo utilizzati
- **Persistenti**, con un periodo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano
- **Condivise**, utilizzate da applicazioni diverse

e garantendo:

- **Affidabilità**, resistenza a malfunzionamenti hardware e software
- **Sicurezza**, con una disciplina e un controllo degli accessi
- **Efficienza**, utilizzando al meglio le risorse di spazio e tempo del sistema

**Schema (di base di dati)** Uno schema descrive un insieme di concetti, cioè classi di dati

**Istanza (di base di dati)** Una istanza descrive un insieme di dati che fanno riferimento ai concetti dello schema

**Modello** Insieme di costrutti utilizzati per organizzare i dati di interesse in concetti e descriverne la dinamica, cioè come cambiano nel tempo

**Struttura di rappresentazione** Componenti fondamentali di un modello (*es. Struttura utilizzata: la relazione*)

**Modello logico**

- utilizzati nei DBMS esistenti in commercio per l'organizzazione dei dati
- utilizzati dai programmi
- **indipendenti dalle strutture fisiche**

esempi: Relazionale, reticolare, gerarchico, a oggetti

**Modello concettuale** permettono di rappresentare i dati in modo indipendente da ogni sistema DBMS

- cercano di descrivere direttamente i concetti del mondo reale
- sono utilizzati nelle fasi preliminari di progettazione

Il più noto è il modello Entità-Relazione

**Schema logico** Descrizione della base di dati con strutture di alto livello (ad esempio, la struttura di relazione)

**Schema fisico** Rappresentazione dello schema logico per mezzo di strutture fisiche di memorizzazione (files), blocchi in memoria, ecc.

**Schema esterno** Descrizione di parte della base di dati in un modello logico (“viste” parziali, di interesse per particolari gruppi di utenti)

**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 software

**Indipendenza logica** *Il livello esterno è indipendente da quello logico:*

- aggiunte o modifiche agli schemi esterni non richiedono modifiche al livello logico
- modifiche allo schema logico che lascino inalterato lo schema esterno sono trasparenti

**Data Description Language** Linguaggio di descrizione dei dati, con cui si descrivono le strutture degli schemi

**Data Manipulation Language** Linguaggio di manipolazione dei dati, con cui si esprimono le interrogazioni che ritrovano i dati dalla base di dati e le transazioni che li aggiornano (es. SQL)

## 2 Il modello Entità-Relazione

**Entità** Una entità è una classe di oggetti (fatti, persone, cose) che:

- sono di interesse per l'applicazione
- hanno esistenza autonoma
- hanno proprietà comuni

**Attributo di Entità** Un attributo di entità è una proprietà locale di un'entità, di interesse ai fini dell'applicazione, cioè una proprietà il cui valore in ogni istanza della entità dipende solamente dall'istanza della entità, e non da altri elementi dello schema

**Dominio di attributo** Si possono utilizzare come domini quelli tipicamente definiti nei linguaggi programmatici

**Relazione** Una relazione si definisce su due o più entità, e rappresenta un legame logico fra tali entità. Il numero di entità coinvolte in una relazione determina il suo grado (2, 3, 4, ...)

- singolare
- sostantivi invece che verbi (se possibile e naturale)

**Attributo di relazione**

- Un attributo di relazione è una proprietà locale di una relazione
- Un attributo della relazione  $R$  tra le entità  $E_1, E_2, \dots, E_n$  descrive una proprietà che non è di  $E_1$ , non è di  $E_2, \dots$ , non è di  $E_n$ , ma del legame logico tra  $E_1, E_2, \dots, E_n$  rappresentato da  $R$
- è una funzione che associa ad ogni istanza di relazione un valore appartenente ad un insieme detto dominio dell'attributo

**Vincolo di cardinalità** Un vincolo di cardinalità tra una entità E e una relazione R esprime un limite minimo (cardinalità minima) ed un limite massimo (cardinalità massima) di istanze della relazione R a cui può partecipare ogni istanza dell'entità E

**Relazione IS-A** Tra due classi rappresentate da due entità nello schema concettuale sussista la relazione IS-A (o relazione di sottoinsieme), e cioè che ogni istanza di una sia anche istanza dell'altra.

La relazione IS-A nel modello ER si può definire tra due entità, che si dicono “entità padre” ed “entità figlia” (o sottoentità, cioè quella che rappresenta un sottoinsieme della entità padre).

**Generalizzazione** Talvolta, però, l'entità padre può generalizzare diverse sottoentità rispetto ad un unico criterio. In questo caso si parla di generalizzazione.

- **Completa:** l'unione delle istanze delle sottoentità è uguale all'insieme delle istanze dell'entità padre (freccia nera)
- **Non completa** (freccia vuota)

**Attributi composti** Un attributo può anche essere definito su un dominio complesso. Di particolare interesse è il caso di dominio di tipo “record”, cioè composto di domini (anche detti coampi) elementari. Un attributo il cui dominio è di tipo record si dice composto.

**Relazioni n-arie** Una relazione di grado maggiore di 2 si dice n-aria

**Identificatori** Un identificatore di una entità è un insieme di proprietà (attributi o relazioni) che permettono di identificare univocamente le istanze di un'entità.

Tipi:

- **Interno**, ossia formato solo da attributi di E
- **Esterno**, ossia formato da attributi di E e da relazioni che coinvolgono E, oppure solo da relazioni che coinvolgono E

### 3 Modello Relazionale

**Modello logico dei dati** Un insieme di strutture di rappresentazione utilizzabili per descrivere un insieme di dati, o schema logico, che, a sua volta, descrive una realtà di interesse

**Struttura di rappresentazione** Relazione  $\rightarrow$  Descrizione nel modello logico

es. *Gara(Numero, Sede, SiglaAutomobile, Marca, Guidatore)*

#### Schema logico

- **Relationship**: che rappresenta una classe di fatti, nel modello Entity-Relationship
- **Relazione matematica**: come nella teoria degli insiemi
- **Relazione**: secondo il modello relazionale dei dati

#### Relazione matematica

- $D_1, \dots, D_n$  ( $n$  insiemi anche non distinti, detti anche domini)
- Prodotto cartesiano  $D_1 \times \dots \times D_n$ :
  - L'insieme di tutte le  $n$ -uple  $(d_1, \dots, d_n)$  tali che  $d_1 \in D_1, \dots, d_n \in D_n$
- relazione matematica su  $D_1, \dots, D_n$ :
  - un sottoinsieme di  $D_1 \times \dots \times D_n$ .
- $D_1, \dots, D_n$  sono i domini della relazione

Proprietà:

- una relazione matematica è un insieme di  $n$ -uple al loro interno ordinate
- una relazione è un **insieme**, quindi:
  - non c'è ordinamento fra le diverse  $n$ -uple
  - le  $n$ -uple sono distinte (non ce ne possono essere due uguali)
  - ciascuna  $n$ -upla è al suo interno ordinata: l' $i$ -esimo valore proviene dall' $i$ -esimo dominio

**Relazioni** Una relazione (o tabella) nel modello relazionale rappresenta una relazione (in matematica) in cui:

- i valori di ogni colonna (attributo) sono fra loro omogenei (es. Juve, Lazio, ecc.)
- le righe (n-ple o tuple) sono diverse fra loro
- le intestazioni delle colonne (nomi di attributi) sono diverse tra loro

In una relazione nel modello relazionale:

- l'ordinamento tra le righe è irrilevante
- l'ordinamento tra le colonne è irrilevante

**Modello relazionale basato su valori** Il modello relazionale è basato su valori. I riferimenti fra dati in relazioni diverse sono rappresentati per mezzo di valori dei domini che compaiono nelle ennuple.

Vantaggi:

- l'utente finale vede gli stessi dati dei programmatori, che sono appunto costituiti dalle tabelle
- indipendenza della rappresentazione dalle strutture fisiche (ad es. indirizzi di memoria), che possono cambiare dinamicamente
- i dati sono portabili più facilmente da un sistema ad un altro, perché non dipendono dalle caratteristiche fisiche

**Schema di relazione** Un nome  $R$  con un insieme di attributi  $A_1, \dots, A_n$ :

$$R(A_1, \dots, A_n)$$

**Schema di base di dati** Insieme di schemi di relazione:

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

## Ennupla / n-upla

- Una ennupla (o n-pla) su un insieme di attributi  $X$  è una funzione che associa a ciascun attributo  $A$  in  $X$  un valore del dominio di  $A$
- Esempio  $\rightarrow$  Codice:01 Titolo:Analisi Docente:Rossi
- Il simbolo  $t[A]$  denota il valore della ennupla  $t$  sull'attributo  $A$
- Esempio  $t[\text{Codice}]$  per la ennupla sopra è 01

**Istanza di relazione** Istanza di relazione (o relazione) su uno schema  $R(X)$ : insieme  $r$  di ennuple su  $X$

**Istanza di base di dati** Istanza di base di dati (o base di dati) su uno schema  $R = \{R_1(X_1), \dots, R_n(X_n)\}$  : insieme di relazioni  $r = r_1, \dots, r_n$  (con  $r_i$  relazione su  $R_i$ )

**Vincolo di integrità** Proprietà che deve essere soddisfatta da tutte le istanze di uno schema che rappresentano info corrette per l'applicazione  
Un vincolo di integrità è una funzione booleana (o predicato) che associa ad ogni istanza  $r$ :

- Il valore vero se la istanza è corretta (rappresentazione della realtà)
- Il valore falso se la istanza è scorretta (rappresentazione della realtà)

**Vincolo intrarelazionale** Definiti all'interno di una relazione

**Vincolo di valore** Vincoli su valori (o di dominio), cioè su singoli valori di attributi

**Vincolo di ennupla** Vincoli di ennupla, cioè definiti sulle ennuple di una relazione

- Esprimono condizioni sui valori di ciascuna ennupla in una relazione
- Esempio: non ci possono essere due ennuple con lo stesso numero di matricola



**Vincolo di relazione** Vincoli relativi all'insieme di ennuple di una relazione

**Vincolo di chiave** Abbiamo necessità di individuare informazioni che ci permettano di rappresentare ogni oggetto di interesse con una ennupla differente e identificarlo quando se ne abbia necessita

- **Chiave:** Insieme di attributi che identificano le ennuple di una relazione
- un insieme  $K$  di attributi è **superchiave** per una relazione  $r$  se  $r$  non contiene due ennuple distinte  $t_1$  e  $t_2$  con  $t_1[K] = t_2[K]$
- $K$  è anche chiave per  $r$  se è una superchiave minimale per  $r$  (cioè non contiene un'altra superchiave)

**Vincolo interrelazionale** Definiti tra due o più relazioni

**Valori nulli**

- **valore sconosciuto:** non lo conosciamo ma esiste
- **valore inesistente:** non esiste
- **valore senza informazione:** non si può dire niente (non lo conosciamo ovvero non esiste)