

# Appunti del corso di basi di dati I

Riccardo Cereghino

8 marzo 2020



# Indice

<b>1</b>	<b>DBMS - Data Base Managament System</b>	<b>9</b>
1.1	Sistema di gestione di basi di dati . . . . .	9
1.1.1	Obiettivi e servizi di un DBMS . . . . .	10
1.1.2	Modelli dei dati . . . . .	11
1.1.3	Modello relazionale . . . . .	12
1.1.4	Schemi ed istanze . . . . .	12
1.1.5	Livelli nella rappresentazione dei dati . . . . .	12
1.1.6	Linguaggi di un DBMS . . . . .	13



## Elenco delle figure



## Elenco delle tabelle





# Capitolo 1

## DBMS - Data Base Managment System

### 1.1 Sistema di gestione di basi di dati

I sistemi di gestione di basi di dati si pongono di risolvere i problemi legati alla gestione e salvataggio di dati, quali:

- **ridondanza ed inconsistenza dei dati**, ovvero la duplicazione dei dati su file multipli, che comporta anche un pericolo di inconsistenza;
- **difficoltà nell'accesso ai dati**, la mancanza di una descrizione di alto livello e centralizzata dei dati ne rende estremamente difficoltoso l'utilizzo al fine di rispondere a nuove esigenze applicative, a questo proposito i *DBMS* forniscono linguaggi per facilitare l'accesso ai dati, integrazione con i linguaggi di programmazione e funzionalità reattive;
- **problemi nell'accesso concorrente ai dati**, operazioni multiple sullo stesso dato possono causare problemi di concorrenza, i *DBMS* mettono a disposizione il concetto di *transazione*, che garantisce la consistenza dei dati in presenza di transazioni concorrenti;

- **problemi di integrità dei dati**, i *DBMS* forniscono la possibilità di stabilire dei vincoli al salvataggio ed alla modifica dei dati, preservandone l'integrità.

### 1.1.1 Obiettivi e servizi di un DBMS

I DBMS implementano una serie di servizi per offrire l'utilizzo della base di dati e lo sviluppo di applicazioni con cui si interfacciano, alcuni sono invocabili dagli utenti, altri sono utilizzati per il funzionamento interno.

I DBMS adottano un'*architettura client-server*, per cui le funzionalità del sistema sono realizzate da due moduli distinti: il **client** che gestisce l'interazione tra utente e DBMS ed il **server** che si occupa della memorizzazione e gestione dei dati.

#### Principali servizi offerti da un DBMS

- **Descrizione dei dati**: per specificare i dati da memorizzare nella base di dati;
- **manipolazione dei dati, per**:
  - accedere ai dati;
  - inserire nuovi dati;
  - modificare dati esistenti;
  - cancellare dati esistenti;
- **controllo di integrità**: per evitare di memorizzare dati non corretti;
- **strutture di memorizzazione**: per rappresentare in memoria secondaria i costrutti del modello dei dati;
- **ottimizzazione di interrogazioni**: per determinare la strategia più efficiente per accedere ai dati;
- **protezione dei dati**: per proteggere i dati da accessi non autorizzati;
- **ripristino della base di dati**, per evitare che errori e malfunzionamenti:

- determinino una base di dati inconsistente;
- provochino perdite di dati;
- **controllo della concorrenza:** per evitare che accessi concorrenti alla base di dati provochino inconsistenze dei dati.

### 1.1.2 Modelli dei dati

Una delle caratteristiche principali di un DBMS è di poter disporre di una rappresentazione logica ad alto livello dei dati, i *modelli di dati* astraggono i dati presenti nel sistema per essere utilizzati più semplicemente da utenti ed applicazioni.

#### Concetti di base

Un *modello di dati* è un formalismo che racchiude tre componenti fondamentali:

- un insieme di strutture dati;
- un linguaggio per specificare, aggiornare e vincolare le strutture dati previste dal modello;
- un linguaggio per manipolare i dati.

**Struttura dati** Ad modello di dati corrisponde una struttura dati, composta da:

- **entità:** insieme di oggetti della realtà applicativa di interesse, aventi caratteristiche comuni;
- **istanza di entità:** singolo oggetto della realtà applicativa di interesse, modellato da una certa entità;
- **attributo:** proprietà significativa di un entità, ai fini della descrizione della realtà applicativa di interesse (ogni entità è caratterizzata da uno o più attributi), un attributo di un entità assume uno o più valori per ciascuna delle istanze dell'entità, detti *valori dell'attributo*, in un insieme di possibili valori, detto *dominio dell'attributo*;

- **associazione:** corrispondenza tra un certo numero di entità, anche le associazioni possono avere degli attributi che corrispondono alle proprietà delle associazioni;
- **istanza di associazione:** corrispondenza tra le istanze di un certo numero di entità.

### 1.1.3 Modello relazionale

Il modello relazionale è basato su di una singola struttura dati, la relazione. Una relazione viene solitamente rappresentata come una tabella con righe (dette tuple) e colonne contenenti dati di tipo specificato.

Una tupla tipicamente rappresenta un'istanza dell'entità modellata dalla tabella a cui la tupla appartiene, mentre la colonna di una tabella rappresenta un particolare attributo dell'entità modellata dalla tabella stessa.

Attributi con la proprietà di identificare univocamente le tuple di una relazione vengono chiamati *chiavi*, mentre i corrispondenti attributi nell'altra relazione, prendono il nome di *chiavi esterne*.

### 1.1.4 Schemi ed istanze

In un DBMS individuiamo lo *schema della base di dati* e l' *istanza della base di dati*.

Lo schema fornisce una descrizione dei dati per tipo, l'istanza è l'insieme delle tuple contenute nella base di dati.

### 1.1.5 Livelli nella rappresentazione dei dati

Uno degli scopi di un DBMS è di fornire una rappresentazione ad alto livello di un insieme di dati nascondendone l'effettiva memorizzazione, per cui la base di dati fornisce tre livelli diversi di visualizzazione/astrazione:

- **livello fisico:** il livello più basso con cui viene definito lo schema fisico della base di dati, precisando come i dati sono effettivamente memorizzati tramite strutture di memorizzazione;
- **livello logico:** il secondo livello di astrazione in cui viene descritto lo *schema logico*, ovvero quali sono i dati memorizzati nella base di dati, eventuali associazione tra di essi ed eventuali vincoli di integrità semantica e di autorizzazione, l'intera base di dati è descritta tramite un numero limitato di strutture dati che costituiscono il modello dei dati;
- **livello esterno:** è il livello di astrazione più alto, descrive una porzione dell'intero schema della base di dati, possono essere definite più viste di una stessa base di dati.

La presenza di questi livelli di astrazione assicura alcune importanti proprietà ai dati: l'**indipendenza fisica** e l'**indipendenza logica**.

Il concetto di indipendenza fisica esprime la possibilità di interagire con il livello logico senza apportare modifiche al livello fisico.

Il concetto di indipendenza logica consente la possibilità di interagire con il livello esterno senza apportare modifiche al livello logico.

### 1.1.6 Linguaggi di un DBMS

I DBMS mettono a disposizione un insieme di linguaggi che permettono agli utenti di interagire con il DBMS per descrivere e manipolare i dati di interesse e specificare vincoli, tra questi i linguaggi più rilevanti sono:

**DDL - Data Definition Language** Il linguaggio di definizione dei dati consente di specificare ed aggiornare lo schema di una base di dati, in particolare il DDL concretizza il modello dei dati fornendo la notazione che permette di specificarne le strutture dati.

Il DDL deve supportare la specifica del nome della base di dati, come pure di tutte le unità logiche elementari della base di dati e di eventuali vincoli di integrità

semantica e di autorizzazione prevedendo comandi per l'aggiornamento delle strutture dati previste dal modello.

Per quanto concerne i linguaggio di manipolazione dei dati si distingue tra *linguaggi procedurali (operazionali)* e *linguaggi non procedurali (dichiarativi)* per esempio SQL.

**DML - Data Manipulation Language** Una base di dati, organizzata logicamente tramite il modello dei dati e definita tramite il DDL è accessibile agli utenti ed alle applicazioni tramite il DML. Le operazioni fornite da questo linguaggio servono per gestire le istanze della base di dati e sono fondamentalmente quattro:

- **inserimento:** per l'immissione di nuovi dati;
- **ricerca:** per il ritrovamento dei dati in interesse, detta interrogazione (*query*);
- **cancellazione:** per l'eliminazione di dati obsoleti;
- **modifica:** per variare dati esistenti.

**SDL - Storage Definition Language** La corrispondenza tra le strutture logiche e le strutture di memorizzazione deve essere opportunamente definita. Nella maggior parte dei DBMS attuali la definizione di tale corrispondenza è eseguita automaticamente dal DBMS stesso una volta che lo schema logico è definito, tuttavia l'utente può influenzare le scelte operate dal DBMS tramite i comandi del linguaggio di definizione delle strutture di memorizzazione.