

Basi di dati

Elia Ronchetti

@ulerich

Marzo 2022

Indice

1	Introduzione - Che cos'è un DB e un DBMS	3
1.1	Perché creare un Database	3
1.2	Base di Dati - DB - Data Base	4
1.2.1	Modello dei dati	6
1.3	Schemi e Istanze	6
1.4	Modelli concettuali	6
1.5	Modelli logici - Modello Relazionale	7
1.5.1	Modello Relazionale	7
1.5.2	Linguaggi per basi di dati	7
1.5.3	Creazione di un database	7
1.6	Vantaggi e svantaggi dei DBMS	8
2	Modello Entity Relationship - ER	9
2.1	Fasi del ciclo di vita	9
2.2	La progettazione di basi di dati	10
2.2.1	Progettazione concettuale	10
2.2.2	Vantaggi della progettazione concettuale	11
2.3	Modello Entità Relazione	11
2.4	I Costrutti del modello ER	11
2.4.1	Entità	12

Capitolo 1

Introduzione - Che cos'è un DB e un DBMS

Che cos'è un Data Base Una collezione di dati utilizzati per rappresentare le informazioni di interesse di un sistema informativo

Che cos'è un DBMS? Un DBMS (Data Base Management System) è un insieme di programmi che permettono di creare, usare e gestire una base di dati, è quindi un software general purpose che facilita il processo di definizione, costruzione e manipolazione del database per varie applicazioni.

1.1 Perchè creare un Database

Un soggetto, come per esempio un'azienda, ha molti dati da manipolare

- Persone
- Denaro
- Materiali
- Informazioni

Per gestire questi dati è necessario un sistema che li organizzi e li gestisca in modo efficiente e sicuro. Questo sistema è detto **Sistema Informativo**, cioè un componente di una organizzazione che gestisce le informazioni di interesse, con i seguenti scopi:

- Acquisizione/Memorizzazione
- Aggiornamento

4 *CAPITOLO 1. INTRODUZIONE - CHE COS'È UN DB E UN DBMS*

- Interrogazione
- Elaborazione

Il **Sistema Informatico** è invece la porzione automatizzata del Sistema informativo, la parte quindi che gestisce informazioni tramite tecnologia informatica.

Il Sistema Informatico ha i seguenti obiettivi:

- Garantisce che i dati siano conservati in modo permanente sui dispositivi di memorizzazione
- Permette un rapido Aggiornamento dei dati
- Rende i dati accessibili alle interrogazioni degli utenti
- Può essere distribuito sul territorio

Gestione delle informazioni Nei sistemi informatici le informazioni vengono rappresentate in modo essenziale attraverso i dati. I Dati hanno bisogno di essere interpretati, ma costituiscono una precisa rappresentazione di forme più ricche di informazioni e conoscenza, inoltre sono più stabili nel tempo rispetto ad altre componenti (come processi, tecnologie, ruoli umani) e restano gli stessi nella migrazione da un sistema al successivo.

1.2 **Base di Dati - DB - Data Base**

Data Base - DB Collezione di dati utilizzati per rappresentare le informazioni di interesse di un sistema informativo

Altra definizione di DB Insieme di archivi in cui ogni dato è rappresentato logicamente una sola volta e può essere utilizzato da un insieme di applicazioni da diversi utenti secondo opportuni criteri di riservatezza.

Data Base Management System - DBMS Sistema software capace di gestire collezioni di dati che siano grandi, condivise e persistenti, assicurando la loro affidabilità e privacy.

Elenco caratteristiche DBMS Sistema che garantisce collezioni di dati:

- grandi
- persistenti
- condivise

Garantendo:

- Privatezza - Meccanismi di autorizzazione (come ACL)
- Affidabilità - Resistenza malfunzionamenti hardware e software (tramite tecniche come la gestione delle transazioni)
- Efficienza
- Efficacia

Transazione → Insieme di operazioni da considerare indivisibile (atomico), la sequenza di operazioni sulla base di dati viene eseguita per intero o per niente.

L'effetto di transazioni concorrenti deve essere coerente (ad esempio "equivalente" all'esecuzione separata).

I risultati delle transazioni sono permanenti, la conclusione di una transazione corrisponde a un impegno (in inglese commitment) a mantenere traccia del risultato in modo definitivo.

I DBMS devono essere efficienti cercando di utilizzare al meglio le risorse di spazio di memoria e tempo.

Efficacia intesa come resa produttiva delle attività dei loro utilizzatori.

Caratteristiche di un DB

- Ridondanza minima e controllata
- Consistenza delle informazioni
- Dati disponibili per utenze diverse e concorrenti
- Dati controllati e protetti (da malfunzionamenti hardware e software)
- Indipendenza dei dati dal programma

Riassumendo, un DBMS è un prodotto software in grado di gestire collezioni di dati che siano:

- Grandi
- Persistenti
- Condivise

E che garantiscano

- Affidabilità
- Privacy
- Efficienza

I DBMS permettono inoltre ai dati di essere indipendenti dalla propria rappresentazione fisica.

1.2.1 Modello dei dati

Insieme di costrutti per organizzare i dati di interesse e descriverne la dinamica. Sono componenti fondamentali che permettono la strutturazione dei dati. Per esempio il modello relazionale prevede il costruttore relazione, che permette di definire insiemi di record omogenei.

1.3 Schemi e Istanze

In ogni base di dati esistono:

- Lo schema, sostanzialmente invariante nel tempo, che ne descrive la struttura, il significato (aspetto intensionale). Costituisce quindi la parte astratta delle proprietà.
- L'istanza, che sono i valori attuali e possono cambiare anche molto rapidamente (aspetto estensionale). Costituisce quindi l'aspetto concreto che varia nel tempo.

1.4 Modelli concettuali

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

1.5 Modelli logici - Modello Relazionale

Sono i modelli adottati nei DBMS esistenti per l'organizzazione dei dati e sono utilizzati dai programmi, sono indipendenti dalle strutture fisiche. L'esempio più diffuso e che noi tratteremo è quello del **modello Relazionale**.

1.5.1 Modello Relazionale

I dati vengono strutturati in tabelle, in particolare un DBMS relazione può essere pensato come un insieme di tabelle, dove ogni tabella mantiene informazioni di tipo omogeneo. Diverse tabelle sono collegate (in relazione) fra loro grazie alla presenza di un campo comune che permette di mettere in relazione i dati delle due tabelle.

In questo caso lo schema è la componente intensionale e descrive la struttura della tabella (ed è stabile nel tempo)

Mentre l'istanza è la componente estensionale e descrive i valori attuali, cioè i dati (ed è variabile nel tempo).

1.5.2 Linguaggi per basi di dati

Ci sono i DDL (Data Definition Language) che permettono di definire il DB. Mentre i DML (Data Manipulation Language) permettono di manipolare i dati, interrogando e aggiornando delle basi di dati. Alcuni linguaggi, come SQL (Structured Query Language) hanno funzioni di entrambe le categorie.

1.5.3 Creazione di un database

Le tre fasi

- Definizione - DDL
- Creazione/Popolazione - DDL
- Manipolazione - DML

Query È fondamentale poter interrogare un DB, attraverso per esempio delle **query**. L'efficacia della query dipende da:

- Conoscenza del contenuto del DB
- Esperienza del linguaggio di interrogazione

1.6 Vantaggi e svantaggi dei DBMS

Pro

- Permettono di considerare i dati come risorsa comune di un'organizzazione, a disposizione di molteplici applicazioni e utenti
- Offrono modello della parte di mondo di interesse che è unificato e preciso, utilizzabile in applicazioni attuali e future
- Controllo centralizzato dei dati, riduce ridondanze e incosistenze
- Indipendenza dei dati: favorisce sviluppi di applicazioni flessibili e facilmente modificabili

Contro

- Costosi, complessi, hanno specifici requisiti in termini di software e hardware
- Difficile separare, tra tutti i servizi offerti da un DBMS, quelli effettivamente utilizzati da quelli inutili
- Inadatti alla gestione di applicazioni con pochi utenti

Capitolo 2

Modello Entity Relationship - ER

In questa parte si studierà la come progettare una base di dati a livello concettuale e logico, partendo dai requisiti di utente. Per capirne l'importanza è utile analizzare il ciclo di vita di un sistema informativo

2.1 Fasi del ciclo di vita

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

Il ciclo di vita segue un modello a spirale. Per garantire prodotti di buona qualità è fondamentale seguire una metodologia di progetto.

Metodologia è un'articolazione in fasi/passi di guida ad una attività di progettazione. Avere una metodologia di progetto:

- Permette di suddividere la progettazione in fasi
- Fornisce una strategia da seguire

- Fornisce modelli di riferimento (linguaggi) per descrivere la realtà che stiamo progettando

Serve per garantire:

- Generalità rispetto ai problemi da affrontare
- Qualità in termini di correttezza, completezza ed efficienza
- Facilità d'uso

La metodologia si basa su un principio semplice ma efficace:

Separazione netta tra decisioni relative a:

- Cosa rappresentare
- Come farlo

2.2 La progettazione di basi di dati

La progettazione si divide in 3 fasi:

- Progettazione concettuale
- Progettazione logica
- Progettazione fisica

Ognuna delle fasi si basa su un modello, che permette di generare una rappresentazione formale (schema) della base di dati ad un dato livello di astrazione (concettuale, logico, fisico).

2.2.1 Progettazione concettuale

Traduce i requisiti del sistema informatico in una descrizione formalizzata, integrata delle esigenze aziendali, espressa in modo **indipendente** dalle scelte implementative.

- Formale - Espressa con un linguaggio non ambiguo e capace di descrivere il sistema analizzato
- Integrata - Deve essere in grado di descrivere nella globalità l'ambiente analizzato
- Indipendente dall'ambiente tecnologico

Nel nostro caso:

- Schema concettuale - **Modello ER**
- Schema logico - **Modello relazionale**

2.2.2 Vantaggi della progettazione concettuale

Permette una descrizione dei dati indipendente dagli aspetti tecnologici con un livello di astrazione intermedio fra utente e sistema. Prevala l'aspetto intensionale.

Si tratta di una rappresentazione prevalentemente grafica. Utile per la documentazione.

2.3 Modello Entità Relazione

Il modello ER è un modello grafico semi-formale per la rappresentazione di schemi concettuali. Si è ormai affermato come standard nelle metodologie di progetto e nei sistemi Software di ausilio alla progettazione.

2.4 I Costrutti del modello ER

- Entità
- Relazione
- Attributo semplice
- Attributo composto
- Cardinalità
- Cardinalità di un Attributo
- Identificatore interno
- Identificatore esterno
- Generalizzazione
- Sottoinsieme

2.4.1 Entità

Classe di oggetti (fatti, persone, cose) della applicazione di interesse con proprietà comuni e con esistenza autonoma e della quale si vogliono registrare fatti specifici.

Rappresentazione grafica di entità



Definita come sostantivo al singolare (es. studente, classe, docente, ecc.) A livello estensionale un'entità è costituita da un insieme di oggetti che sono chiamati le sue istanze.

Istanza Occorrenza di un'entità, è l'oggetto della classe che entità rappresenta. Nello schema concettuale rappresentiamo le entità, non le singole istanze.

Riassumendo:

- Conoscenza Astratta \rightarrow Entità
- Conoscenza Concreta \rightarrow Istanza di entità

Attributi

Un attributo di un'entità è una proprietà locale di un'entità di interesse ai fini dell'applicazione. Associa ad ogni istanza di un'entità un valore appartenente a un insieme detto dominio dell'attributo (es. int, string, char, ecc.).

Viene definito quando si vuole rappresentare una proprietà locale delle istanze dell'entità E.

Una proprietà di un oggetto si dice locale quando in ogni istanza dello schema il valore di tale proprietà dipende solamente dall'oggetto stesso e non ha alcun rapporto con altri elementi dell'istanza dello schema.

Attributi composti Si ottengono raggruppando attributi di una medesima entità o relazione che presentano affinità nel loro significato o uso.

Esempio: Indirizzo è composto da Via, Numero, Cap.

Relazione-Associazione

Ogni relazione ha un nome che la identifica univocamente nello schema.

Convenzioni

- Singolare
- Sostantivi invece che verbi

A livello estensionale una relazione R tra le entità E ed F è costituita da un insieme di coppie (x, y) tali che x è una istanza di E , ed y è un'istanza di F . Ogni coppia è detta istanza della relazione R .