

## **Capitolo Secondo**

# **La progettazione concettuale dei dati**

---

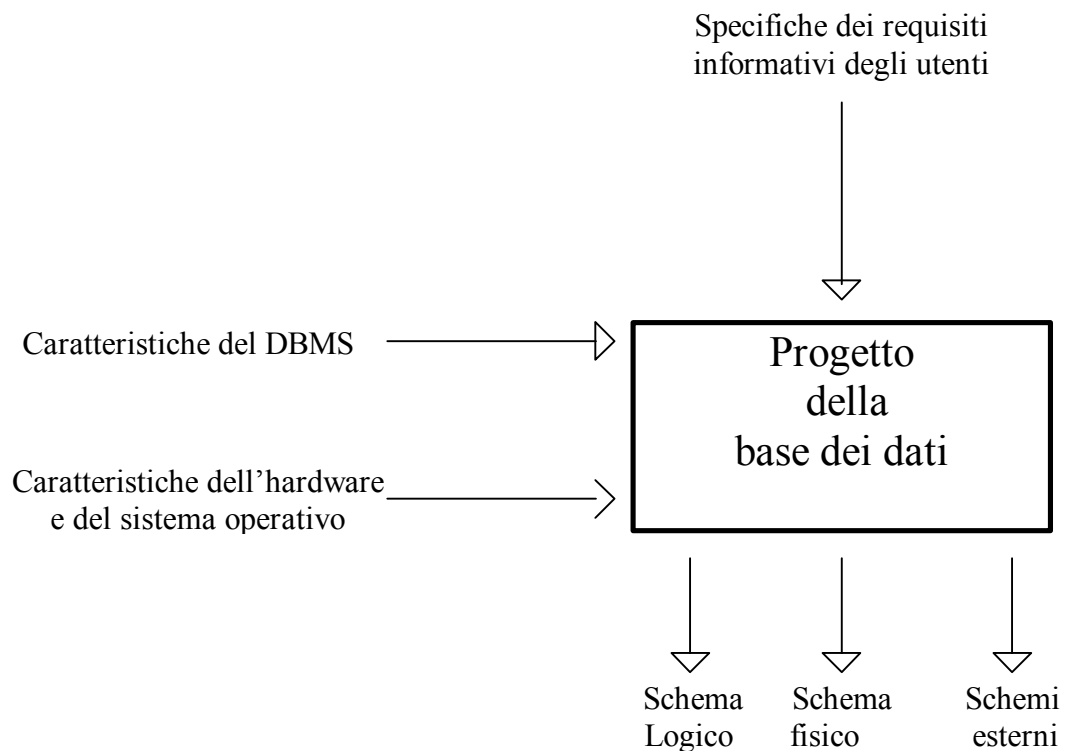
## **2.1. Introduzione**

La progettazione concettuale dei dati non segue dei canoni ben definiti, ma le metodologie che possono essere attuate sono molteplici. Quella che qui verrà analizzata è la metodologia elaborata da Batini ed Altri ed ha come base il modello Entità/Relazioni di P. Chen. Si cerca in quest'opera del Batini di proporre un approccio metodologico che si basi appunto sul modello E/R, che sia il più possibile flessibile e che quindi possa trovare applicazione a prescindere dal singolo caso e dal modello dei dati utilizzati. Per questo si farà la distinzione tra procedimento metodologico e metodologia proprio per sottolineare il fatto che la metodologia che viene utilizzata può trovare impiego in situazioni diverse.

## **2.2. Le metodologie di progetto di una base di dati**

Si può vedere il progetto di una base di dati come un processo di trasformazione che ha come elementi di partenza le specifiche esigenze informative degli utenti (requisiti) e i vincoli realizzativi.

I requisiti riguardano tutto ciò che gli utenti sperano che un sistema informatico realizzi per soddisfare i bisogni informativi e riguarda anche le modalità di funzionamento dello stesso (ad esempio caratterizzazione delle procedure in batch o in tempo reale, tempi di risposta ecc.). I vincoli riguardano invece le componenti hardware e software e il DBMS esistente per cui si dovrà valutare il progetto tenendo conto di questi vincoli.



**Figura 2.1. Esempio di una base di dati**

Il risultato finale della progettazione sarà la realizzazione dei tre schemi con i quali contraddistinguere solitamente un DBMS. In una metodologia di progettazione si possono riscontrare distinte fasi che interagiscono fra loro consentendo il completo svolgimento dell'attività di progettazione di una base di dati. Si possono distinguere queste fasi :

1. *Analisi dei requisiti.* In questa fase s'individuano i requisiti del sistema informatico secondo quelli che sono i bisogni informativi degli utenti. In effetti è proprio in base a questi ultimi che si deve costruire la base di dati. I requisiti descrivono sia i dati che le operazioni che si prevede di eseguire su di essi. Per rendere possibile la descrizione di questi requisiti si utilizzeranno particolari moduli cartacei, interviste, documentazione di archivi esistenti. Questa fase può essere condotta separatamente per ogni

settore aziendale, per cui per ogni settore corrisponde un insieme di requisiti.

2. *Progettazione concettuale.* L'obiettivo della progettazione concettuale è quello di tradurre i requisiti del sistema informatico in una descrizione formale (cioè in linguaggio non ambiguo), integrata (cioè riferito all'intero ambiente analizzato) e indipendente dai vincoli dell'ambiente tecnologico esistente.
3. *Progettazione logica.* In questa fase si traduce lo schema concettuale nelle strutture proprie del DBMS scelto per la realizzazione (ad esempio il modello relazionale).
4. *Progettazione fisica.* Questa fase è dipendente dal particolare DBMS scelto e si cerca di realizzare una migliore efficienza del prodotto. Si specificeranno quali sono i parametri fisici di memorizzazione e ricerca dei dati tenuto conto sia dell'hardware che del software con cui si opera.

### 2.3. Vantaggi delle progettazione concettuale

Alcuni tra i vantaggi di quest'approccio metodologico sono :

- ? *Migliore comunicazione fra utenti e progettisti.* Avviene uno scambio di informazioni fra progettisti e utenti per interpretare le richieste di questi ultimi. Spesso però gli utenti sono "schiavi" delle soluzioni proposte dai progettisti. Con modelli come l'E/R l'utente ha la possibilità di verificare, pur non essendo esperto, se il progetto realizzato corrisponde o no ai suoi bisogni informativi.
- ? *Indipendenza della metodologia dagli aspetti tecnologici.* Si può così puntare l'attenzione in maniera accurata sui dati e sulle loro relazioni potendo valutare successivamente tutte le possibili alternative tecnologiche.
- ? *Realizzazione di uno schema aziendale dei dati.* Attraverso le tecniche di progettazione concettuale si attua una sorta di censimento del patrimonio informativo aziendale tramite la produzione dello schema concettuale globale integrato di tutti i dati dell'azienda.

#### **2.4. Nuovi ruoli professionali per la progettazione di basi di dati**

L'applicazione di una metodologia di progettazione di basi di dati ha delle conseguenze sull'organizzazione del lavoro creando nuove figure come l'amministratore della base dei dati (DBA) e l'amministratore dei dati (DA). Il primo ha come compito quello di progettare lo schema logico e lo schema fisico della base dei dati, di curare la loro ottimizzazione e di gestirne la manutenzione. La sua è una competenza professionale di tipo informatico. Il secondo ha il compito di progettare lo schema concettuale dei dati partendo dalle esigenze informative dell'utente. Non deve avere necessariamente competenze informatiche ma la sua conoscenza deve essere rivolta essenzialmente all'ambiente aziendale. Svolge prevalentemente una funzione di supporto, fornendo il suo contributo all'analisi dei sistemi informativi che verranno automatizzati realizzando sia lo schema concettuale dei singoli settori aziendali che una loro integrazione nello schema concettuale aziendale.

#### **2.5. I modelli dei dati**

Nella fase di progettazione concettuale si possono distinguere tre aspetti rilevanti :

1. Le strutture di classificazione, che permettono di definire le classi di oggetti dello schema concettuale. Le classi permettono di descrivere le caratteristiche comuni d'insiemi di oggetti della realtà.
2. Le strutture per la rappresentazione dei vincoli di integrità, che permettono di descrivere le regole relative alle classi dello schema concettuale per rappresentare la realtà di interesse in modo corretto secondo il modello concettuale prescelto.
3. Le strutture per la descrizione delle operazioni che esprimono l'evolversi della realtà d'interesse.

### **2.5.1 Le strutture di classificazione**

Il punto di partenza è la realtà come viene percepita. Le strutture di classificazione permettono di rappresentare gli oggetti della realtà raggruppandoli in classi in base a proprietà comuni. Questo procedimento può essere effettuato grazie all'astrazione, che è un processo mentale attraverso il quale si evidenzia alcune proprietà e caratteristiche in un insieme d'oggetti che si ritiene importante escludendo le altre che sono ritenute irrilevanti, arrivando poi, a definire un nuovo oggetto che contiene in se tutte le proprietà considerate. Ad esempio si può considerare l'astrazione che porta alla definizione del concetto di bicicletta per cui noi si astrae dalle singole parti meccaniche (manubrio, sella, ecc..) e si considera l'insieme di tali componenti come un unico oggetto.

Nel progetto concettuale le astrazioni considerate sono :

- ✍ La classificazione
- ✍ L'aggregazione
- ✍ La generalizzazione

La classificazione è l'astrazione che conduce alla definizione di una classe a partire da un insieme di oggetti (istanze della classe) di cui si cerca di definire le proprietà comuni. Il risultato è un nuovo oggetto come elemento unificatore delle istanze che dovranno essere atomiche ovvero non ulteriormente scomponibili. Ad esempio, si può applicare l'astrazione di classificazione all'insieme delle persone esistenti nella nostra realtà percepita e da questa possiamo, tramite tale procedimento di astrazione, definire il concetto di persona a cui si attribuiscono certe caratteristiche astraendo da quelle particolari dei singoli.

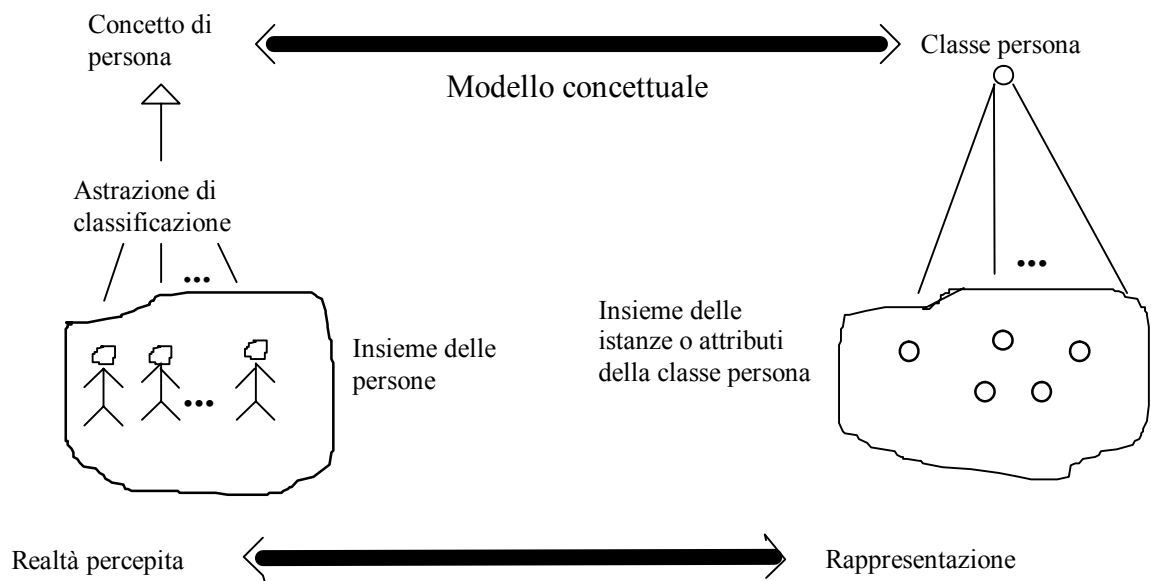


Figura 2.2. Esempio d'astrazione di classificazione

L'aggregazione, invece, è l'astrazione attraverso la quale si giunge alla definizione di un concetto a partire da altri concetti che ne costituiscono le proprietà. Si costituisce così una classe che è a sua volta aggregazione di classi. Ad esempio, il concetto di esame è una astrazione di aggregazione applicata ai concetti di corso studente e voto.

La generalizzazione, infine, è l'astrazione attraverso cui si giunge alla definizione di un concetto partendo da altri concetti definiti dall'astrazione di classificazione e che hanno caratteristiche comuni. Si definisce quindi una nuova classe in cui entrano tutte le istanze delle classi definite tramite classificazione. Un esempio di astrazione di generalizzazione è quello per cui partendo dai concetti di studente e d'impiegato si può arrivare al concetto di persona.

### 2.5.2. Strutture per la rappresentazione dei vincoli d'integrità

Le strutture per la rappresentazione dei vincoli d'integrità permettono di rappresentare nello schema concettuale le regole che gli oggetti della realtà rispettano e che allo stesso modo gli oggetti della rappresentazione devono

rispettare. Tali regole vengono dette vincoli d'integrità e sono diverse secondo il modello dei dati prescelto.

Una prima distinzione che può essere fatta è fra vincoli impliciti e vincoli espliciti di cui si è già parlato nel primo capitolo.

Un'altra distinzione che si ha è quella fra vincoli statici e vincoli dinamici. I vincoli statici sono delle regole che devono essere osservate dagli oggetti della realtà in ogni istante. Quindi queste regole esprimono delle condizioni che devono essere sempre soddisfatte affinché la base dei dati sia una rappresentazione plausibile della realtà. Ad esempio, nella realtà esiste una regola che associa ad ogni impiegato uno ed un solo codice fiscale. I vincoli dinamici sono invece regole che devono essere rispettate da ogni evoluzione della realtà determinata da transazioni rappresentabili tramite operazioni, le quali a loro volta sono esprimibili tramite un linguaggio di manipolazione. Un esempio di questi vincoli è dato da uno studente che non può iscriversi ad un anno di corso precedente a quello in cui è iscritto. Comunque la distinzione tra vincoli dinamici e vincoli statici è relativa e dipende dallo schema concettuale scelto e non solo dal tipo di regole che il vincolo rappresenta.

### **2.5.3. Le strutture per la rappresentazione delle operazioni**

Le strutture per la rappresentazione delle operazioni consentono :

- ? Operazioni di interrogazione. Descrivono i procedimenti che hanno lo scopo di interrogare le occorrenze della base dei dati.
- ? Operazioni di manipolazione. Descrivono le modalità con cui le occorrenze della base dei dati evolvono per adeguarsi alle modificazioni avvenute nella realtà d'interesse. Permettono quindi :
  - l'aggiunta di un oggetto come istanza di una classe
  - l'eliminazione di un oggetto delle istanze di una classe.



## 2.6. Il modello Entità - Relazioni

Questo modello è stato sviluppato da P.P.Chen nel 1976 ed è una delle più diffuse metodologie di progettazione. La sua forza risiede nell'estrema semplicità e sinteticità che lo rende comprensibile anche ai non esperti d'informatica dando allo stesso tempo una buona visione d'insieme dei sistemi aziendali e dei collegamenti esistenti tra loro. In effetti, il risultato dell'applicazione di questa metodologia di progettazione è una rappresentazione della struttura di una base di dati che si serve di simboli grafici.

Come tutti i modelli, l'E/R fornisce una rappresentazione semplificata della realtà in cui si cerca di cogliere gli aspetti che più interessano. Spesso viene visto come schema provvisorio o schema di lavoro sul quale impostare discussioni con utenti di vari reparti per verificarne la rispondenza al mondo reale e allo stesso tempo migliorare la percezione che le stesse persone hanno della loro realtà in cui lavorano. La sua estensione può riguardare l'intera azienda come la singola transazione.

Il modello è composto di due elementi base : entità e relazioni.

? *Un'entità è una cosa, un concetto, un'organizzazione o un evento d'interesse per l'organizzazione. Un tipo d'entità è una classificazione d'entità che soddisfano determinati criteri<sup>1</sup>. Un'entità rappresenta un insieme d'oggetti della realtà cui s'individuano proprietà comuni. E questa realtà può essere modellata in maniera diversa secondo l'organizzazione in questione. Si può pensare ad un'entità come ad una classe definita da un'astrazione di classificazione. Nella realtà scolastica si potrà trovare, ad esempio, un'entità studente considerando le singole proprietà che interessano come nome, cognome, facoltà cui è iscritto, esami sostenuti e votazione. Altra entità può essere il corso caratterizzato dal suo codice, descrizione, numero d'ore annue ecc.. Solitamente i nomi delle entità si indicano al singolare, quindi avremo entità *studente* ed entità *corso*. Le entità , quindi,*

---

<sup>1</sup> P.P. Chen, L'approccio entità relazioni - Bit 3/1991 traduzione da byte

sono delle astrazioni della realtà mentre le sue concrete materializzazioni si dicono occorrenze. Ad esempio le occorrenze dell'entità studente saranno Mario Rossi, Paolo Verdi e i loro specifici valori per i vari attributi. Un'entità può avere perfettamente senso anche se in un particolare momento non vi sono occorrenze. Ad esempio, l'entità ordine per un'azienda, può avere perfettamente senso anche se in un certo momento non dovesse esistere nessuno specifico ordine in essere.

? Una relazione (o associazione) è un'interazione fra entità. Un tipo di relazione è una classificazione di relazioni basata su determinati criteri. Normalmente, i nomi corrispondono ad entità, mentre i verbi corrispondono a relazioni <sup>2</sup>. Una relazione, quindi, è una classe che può vedersi come astrazione di aggregazione delle entità in esse coinvolte.

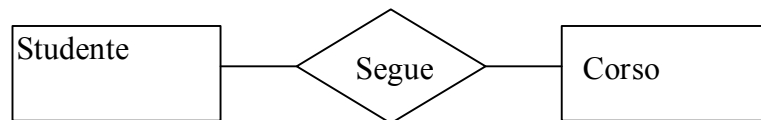


Figura 2.3. Esempio di schema E/R

La relazione può interessare più di due entità però, in generale, si preferisce spezzare le eventuali associazioni n-arie che legano n entità in singole associazioni binarie, per unire le entità due a due<sup>3</sup>.

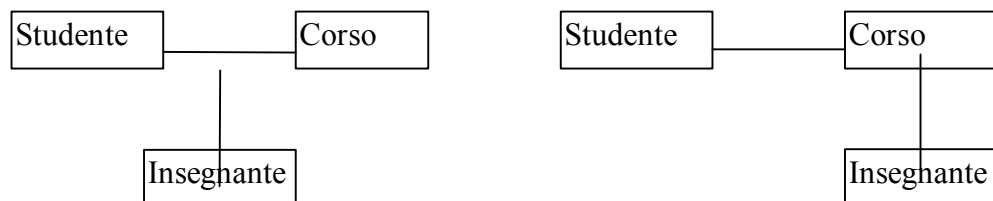


Figura 2.4. Schema E/R di tre entità legate da un'associazione e schema E/R di tre entità legate da due associazioni

<sup>2</sup> P.P. Chen, L'approccio entità relazioni - Bit 3/1991 traduzione da byte

<sup>3</sup> Questo tipo di logica che tende a vedere le associazioni in maniera binaria, tipica della struttura relazionale, viene meno nell'ambito dell'analisi multidimensionale in cui occorre prendere in considerazione più "dimensioni" contemporaneamente ( Radel ).