

## 4. Il modello Entita'/Relazione (E/R)

### 4.1. I modelli concettuali

Un **modello di dati** permette di descrivere i dati relativi ad un qualsiasi sistema informativo, e le relazioni intercorrenti tra essi, mediante determinate **strutture di rappresentazione**, il cui utilizzo e' sottoposto a regole definite nel modello stesso. Se la descrizione risultante e' indipendente dall'ambiente tecnologico nel quale i dati saranno trattati, il modello e' di tipo **concettuale**. In un modello concettuale i dati oggetto di descrizione possono essere dati elementari o disaggregati oppure dati aggregati (risultato di elaborazioni sui dati elementari o disaggregati). Le strutture di rappresentazione di un modello per i dati elementari o disaggregati sono in linea di principio notevolmente differenti da quelle di un modello per i dati aggregati, in quanto mentre nel primo caso le rappresentazioni sono direttamente riferibili al mondo reale (o meglio, agli oggetti del mondo reale rispetto ai quali i dati sono rilevati), nel caso di dati aggregati, oltre alle caratteristiche degli oggetti fonte dei dati, occorre poter anche rappresentare quelle delle elaborazioni eseguite per ottenerli (fig. 4.1).

L'uso dei modelli concettuali nelle metodologie di progettazione risponde all'esigenza di avere a disposizione strumenti di rappresentazione intermedi tra la realta' e la rappresentazione logico-fisica dei dati. Infatti, come gia' visto in precedenza, quest'ultima e' fortemente orientata alla struttura di memorizzazione utilizzata, e non possiede pertanto la capacita' di descrivere quegli aspetti della realta' non direttamente riconducibili ai dati. Classici modelli per la rappresentazione logico-fisica dei dati sono quelli gerarchico, reticolare e relazionale.

Come vedremo meglio in seguito, le strutture di rappresentazione si basano sul concetto di

**astrazione**. Un processo di astrazione e' un procedimento mentale attraverso il quale a partire da un insieme di oggetti considerati omogenei rispetto ad alcune proprieta' e' possibile giungere alla definizione di un nuovo oggetto. Partendo, ad esempio, dalle singole abitazioni, considerandone le proprieta' comuni rilevanti e tralasciando quelle non significative, e' possibile giungere alla definizione dell'oggetto ABITAZIONE: tale nuovo oggetto rappresenta l'insieme degli oggetti di partenza (le singole abitazioni).

I modelli concettuali possono essere valutati rispetto ad una serie di caratteristiche. Le piu' rilevanti appaiono:

- **l'espressivita'**, intesa come capacita' del modello di rappresentare esaurientemente le proprieta' degli oggetti di interesse;
- **la chiarezza**, cioe' la capacita' del modello di descrivere le proprieta' degli oggetti in modo familiare per l'utente e quindi a lui facilmente comprensibile;
- **l'univocita'**, cioe' la possibilita' di rappresentare gli oggetti simili della realta' di interesse con un unico tipo di struttura, cosi' che ogni struttura del modello abbia nello schema concettuale un suo specifico ruolo;
- **la formalita'**, cioe' la possibilita' di descrivere formalmente la sintassi e la semantica del modello in modo da non lasciare ambiguita' di interpretazione degli schemi concettuali;
- **la completezza**, intesa come capacita' del modello di esprimere tutte le proprieta' di interesse senza dover ricorrere a linguaggi non formali o al linguaggio naturale.

## 4.2. Le strutture di rappresentazione di un modello concettuale

In generale nei modelli concettuali identifichiamo tre tipi di strutture di rappresentazione:

1. strutture di classificazione;
2. strutture per la rappresentazione dei vincoli;
3. strutture per la rappresentazione delle operazioni.

Le strutture di classificazione permettono di rappresentare insiemi di oggetti della realtà attraverso l'individuazione delle proprietà comuni per giungere alla definizione di **CLASSI (DI OGGETTI)**. Si possono individuare differenti procedimenti mentali (astrazioni) attraverso i quali insiemi di oggetti vengono riassunti in un concetto unificante.

La modalità più semplice per definire una **CLASSE** è quella di mettere in rilievo le proprietà in comune dell'insieme degli oggetti (**astrazione di classificazione**). È il caso di insiemi di persone che sono riassumibili nel concetto di persona, da cui la definizione della classe **PERSONA**. Vale la pena sottolineare che, come esiste una corrispondenza tra il concetto di persona e la relativa classe, è possibile stabilire una corrispondenza tra gli individui che formano l'insieme delle persone (da cui il concetto di **PERSONA**) e gli oggetti rappresentati attraverso la classe che chiameremo **istanze della classe** e che nel nostro esempio saranno le istanze della classe persona (fig. 4.2).

Un secondo procedimento di astrazione è quello mediante il quale si giunge alla definizione di un concetto (e quindi di una classe) a partire da altri concetti (altre classi) visti come componenti del primo (**astrazione di aggregazione**). Ad esempio, a partire dai concetti di **MATRICOLA**, **NOME**, **COGNOME** e **SESSO** si può definire la classe **DIPENDENTE**.

Un terzo meccanismo è quello mediante il quale si può giungere alla definizione di una nuova classe sulla base delle caratteristiche comuni individuabili tra altre classi (**astrazione di generalizzazione**). È il caso della definizione della classe **PERSONA** come generalizzazione delle classi **UOMO** e **DONNA**.

È importante, a questo punto, sottolineare che lo **SCHEMA CONCETTUALE**, con le relative **CLASSI**, rappresenta una descrizione della realtà che si può considerare stabile nel tempo. I cambiamenti che avvengono nella realtà sono registrati solo dagli insiemi di istanze delle classi dello schema, rimanendo quindi invariata la concettualizzazione della realtà descritta attraverso le **CLASSI**. Ad esempio in un'azienda si può considerare, perlomeno in un periodo ragionevole, immutabile la concettualizzazione descritta dalla classe **IMPIEGATO**; ciò che varierà saranno gli individui che ne fanno parte e le relative occorrenze di una eventuale base di dati che li registra insieme alle loro caratteristiche.

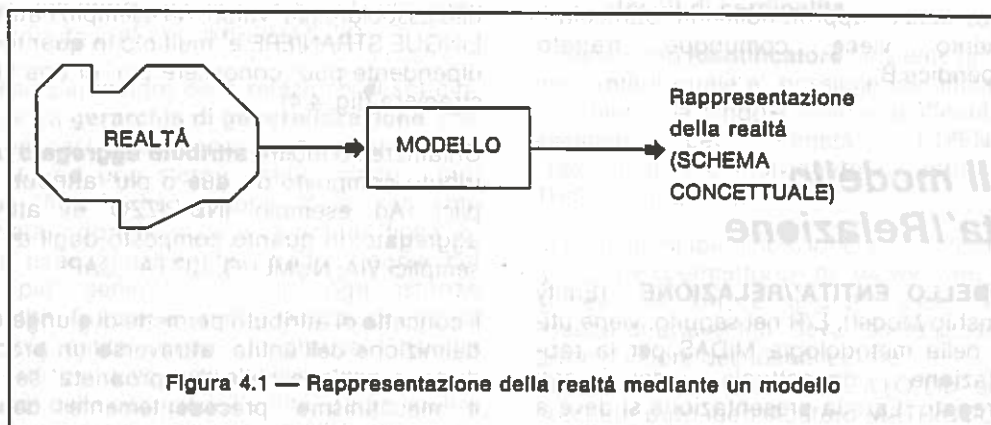
Le regole che gli oggetti definiti nelle classi debbono soddisfare per rappresentare in modo corretto la realtà d'interesse sono espresse, nello schema concettuale, mediante **VINCOLI DI INTEGRITÀ**. Distinguiamo tra vincoli impliciti, propri delle strutture di classificazione di ogni modello, e vincoli espliciti che vengono espressi nel modello da specifiche strutture di rappresentazione simboliche o testuali.

L'utilizzo di strutture di rappresentazione delle **OPERAZIONI** permette di descrivere i cambiamenti che avvengono nella realtà tramite le modalità con cui cambiano le occorrenze della base di dati che ne rappresentano la descrizione. Esempi di operazioni di questo tipo sono le operazioni di "inserimento", "variazione" e "cancellazione" che vanno sotto il nome di "operazioni di manipolazione".

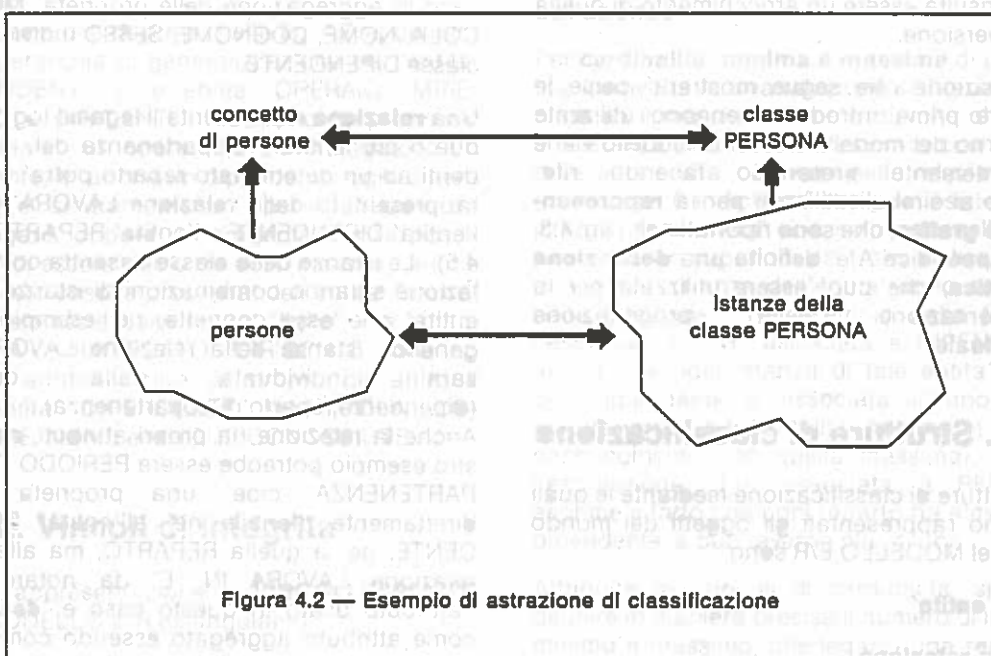
Oltre al tipo di operazioni è utile descrivere le strutture di controllo che permettono di definire il modo con cui si interviene sulle occorrenze della base di dati. Le strutture di controllo tipiche sono:

- la **sequenza**, l'esecuzione una dopo l'altra delle operazioni che seguono;
- l'**alternativa**, l'esecuzione di una o più operazioni condizionate dal verificarsi o meno di una certa condizione;
- il **ciclo**, la ripetizione ciclica, sotto condizione, di una certa operazione.

Il presente manuale non ha, tra le proprie finalità, quella approfondire le tematiche relative a tale argomento, che è proprio più dell'analisi funzionale che di quella dei dati. Di conseguenza in questa sede la trattazione delle operazioni sarà limitata a cenni sulla loro



**Figura 4.1 — Rappresentazione della realtà mediante un modello**



**Figura 4.2 — Esempio di astrazione di classificazione**



tipologia, senza approfondimenti particolari; l'argomento viene comunque trattato nell'Appendice B.

### 4.3. Il modello Entita'/Relazione

Il **MODELLO ENTITA'/RELAZIONE** (Entity Relationship Model), E/R nel seguito, viene utilizzato, nella metodologia MIDAS per la rappresentazione concettuale dei dati disaggregati. La sua presentazione si deve a P.P. Chen nel 1976, mentre la versione qui descritta risulta essere un arricchimento di quella prima versione.

L'esposizione che segue mostrerà come le strutture prima introdotte vengono utilizzate all'interno del modello E/R. Tale modello viene frequentemente presentato facendo riferimento ai simboli utilizzati per la **rappresentazione grafica**, che sono riportati nella fig. 4.3; nella Appendice A è definita una **descrizione linguistica**, che può essere utilizzata per la documentazione della progettazione concettuale.

#### 4.3.1. Strutture di classificazione

Le strutture di classificazione mediante le quali vengono rappresentati gli oggetti del mondo reale nel MODELLO E/R sono:

- l'**entità** ;
- la **relazione** .

Per entità si intendono classi di oggetti con caratteristiche comuni; è il caso dell'entità **DIPENDENTE** che individua tutte le persone che hanno un rapporto di lavoro all'interno di un'azienda (fig. 4.4). I singoli dipendenti rappresenteranno le istanze di quella entità.

Le proprietà elementari di una entità sono dette **attributi**; alla precedente entità **DIPENDENTE** si possono associare come attributi la **MATRICOLA**, il **NOME**, il **COGNOME**, il **SESSO**. I singoli attributi possono assumere diversi valori entro un dominio predefinito: il numero di matricola potrà essere un codice alfanumerico definito dall'azienda, il sesso avrà valori **MASCHIO** o **FEMMINA**, ecc. (fig. 4.4).

Chiameremo **attributo multiplo** quell'attributo al quale per ogni istanza dell'entità è possi-

bile associare più valori. Ad esempio l'attributo **LINGUE STRANIERE** è multiplo in quanto ogni dipendente può conoscere più di una lingua straniera (fig. 4.4).

Chiameremo inoltre **attributo aggregato** un attributo composto da due o più attributi semplici. Ad esempio **INDIRIZZO** è attributo aggregato, in quanto composto dagli attributi semplici **VIA**, **NUMERO**, **CITTA'**, **CAP**.

Il concetto di attributo permette di giungere alla definizione dell'entità attraverso un processo di aggregazione delle sue proprietà secondo il meccanismo precedentemente descritto come astrazione di aggregazione; nel nostro caso la aggregazione delle proprietà **MATRICOLA**, **NOME**, **COGNOME**, **SESSO** individua la classe **DIPENDENTE**.

Una **relazione** rappresenta il legame logico tra due o più entità; l'appartenenza dei dipendenti ad un determinato reparto potrà essere rappresentato dalla relazione **LAVORA IN** tra l'entità **DIPENDENTE** e l'entità **REPARTO** (fig. 4.5). Le istanze della classe descritta come relazione saranno combinazioni di istanze delle entità che essa connette; nell'esempio una generica istanza della relazione **LAVORA IN** sarà individuata dalla coppia (dipendente, reparto di appartenenza).

Anche la relazione ha propri attributi, nel nostro esempio potrebbe essere **PERIODO DI APPARTENENZA**, cioè una proprietà non direttamente riferibile né all'entità **DIPENDENTE**, né a quella **REPARTO**, ma alla loro relazione **LAVORA IN**. È da notare che l'attributo usato in questo caso è definibile come attributo aggregato essendo composto dai due attributi **DATA INIZIO DI APPARTENENZA** e **DATA FINE DI APPARTENENZA** (fig. 4.5).

Due entità si trovano in una **relazione di sottoinsieme** se ogni istanza dell'entità dipendente (detta entità "figlia") è anche istanza dell'entità superiore (detta entità "padre"), mentre non è vero il contrario: cioè non tutte le istanze dell'entità "padre" sono anche istanze dell'entità "figlia". Ne consegue che tutte le proprietà (attributi, relazioni, vincoli, ecc.) della entità "padre" sono ereditate da quella definita "figlia". Quanto detto è esemplificato dalla figura 4.6 in cui è rappresentata la relazione tra l'entità **DIPENDENTE** e l'entità **IMPIEGATO**: è evidente che non tutti i dipendenti sono identificabili nella sola classe **IMPIEGATO**, che potrà quindi avere delle sue proprietà non generalizzabili nella entità di

ordine superiore; e' il caso, ad esempio, dell'attributo LIVELLO IMPIEGATIZIO.

Un caso particolare della relazione di sottinsieme e' la **gerarchia di generalizzazione**, che si ha nel caso in cui e' possibile definire, relativamente ad una stessa entita' "padre", piu' entita' "figlie" che godono della seguente proprieta': ogni istanza di un'entita' "figlia" e' anche istanza dell'entita' "padre" (come nel caso piu' generale); in piu' ogni istanza dell'entita' "padre" e' anche istanza di una ed una sola entita' "figlia". In pratica l'insieme delle istanze delle entita' figlie coincide con l'insieme delle istanze dell'entita' padre; inoltre le entita' figlie non hanno tra di loro istanze comuni.

Un esempio e' descritto nella fig. 4.7: e' definita una gerarchia di generalizzazione tra l'entita' DIPENDENTE e le entita' OPERAIO, IMPIEGATO e DIRIGENTE. Ricorrono le proprieta' enunciate: un operaio (o un impiegato o un dirigente) e' anche un dipendente; ogni dipendente e' sicuramente un operaio, oppure un impiegato, oppure un dirigente, senza possibilita' di essere due o tre cose insieme. Anche in questo caso le proprieta' dell'entita' "padre" sono ereditate dalle entita' "figlie". Le entita' OPERAIO, IMPIEGATO, DIRIGENTE avranno come attributi, oltre ai propri, gli attributi dell'entita' DIPENDENTE: MATRICOLA, COGNOME, NOME e LINGUE CONOSCIUTE.

### 4.3.2. Vincoli di integrita'

Nella rappresentazione dei **VINCOLI ESPliciti** del MODELLO E/R distinguiamo:

#### - vincoli di identificazione ;

#### - vincoli di cardinalita' .

Chiameremo **identificatore** l'insieme di attributi mediante il quale e' possibile identificare univocamente le singole istanze dell'entita'. Ad esempio per l'entita' DIPENDENTE l'identificatore e' individuato dall'attributo MATRICOLA (fig. 4.4).

Per vincoli di identificazione si intendono le regole che permettono di individuare univocamente le istanze di una classe: in ogni istanza di una entita' deve essere presente un diverso valore dell'identificatore.

Nel caso dell'entita' IMPIEGATO cio' significa che l'attributo identificatore MATRICOLA deve assumere valori differenti per ogni impiegato dell'azienda.

Per **cardinalita' minima e massima** di una entita' coinvolta in una relazione si intende il minimo ed il massimo numero di volte che ogni sua istanza puo' essere associata con istanze delle altre entita' coinvolte nella relazione. La cardinalita' 0 indichera' che puo' esistere una istanza della entita' non coinvolta nella relazione. Valori superiori, da 1 a n, indicano che una o piu' istanze dell'entita' sono coinvolte nella relazione. Nell'esempio della fig. 4.5, l'espressione "1,1" associata a DIPENDENTE indica che ogni istanza di tale entita' (cioe' ogni dipendente) e' associata ad uno ed un solo reparto (la cardinalita' minima in questo caso coincide con quella massima), mentre l'espressione "1,n" associata a REPARTO esprime il fatto che ogni reparto ha almeno un dipendente, e puo' averne piu' di uno.

Attribuire dei vincoli di cardinalita' significa definire in maniera precisa il numero di istanze, minimo e massimo, che legano una relazione con un'entita'.