

I criteri visti hanno validità generale, sono cioè indipendenti dalla strategia di progettazione scelta. Come vedremo nel prossimo paragrafo infatti, in ogni strategia esiste prima o poi un momento in cui va presa la decisione sul costrutto da scegliere per rappresentare una certa specifica.

7.2.2 Pattern di progetto

Cominciamo da un caso semplice: quello in cui si individua nelle specifiche un concetto autonomo con proprietà associate, le chiare caratteristiche di una entità del modello E-R. Nel caso per esempio di un impiegato di cui sono di interesse un codice, il nome e l'azienda nel quale lavora, otteniamo il primo, semplice schema con una sola entità riportato in Figura 7.4.

È importante comprendere che, con questa soluzione, non stiamo rappresentando anche il concetto di azienda: qui l'azienda è solo un attributo, ovvero niente di più che una stringa che assegniamo a una occorrenza di impiegato. Per poter rappresentare esplicitamente il concetto di azienda dobbiamo reificare l'attributo, facendolo diventare un'entità. Otteniamo così lo schema in Figura 7.5.

Passiamo ora a dei semplici pattern che coinvolgono le relazioni. Un caso piuttosto frequente di uso di questo costrutto è quello in cui si vuole rappresentare il fatto che un'entità è *parte di* un'altra entità, come avviene negli schemi in Figura 7.6. Queste relazioni sono tipicamente uno a molti e si presentano in due forme. Nel primo caso, l'esistenza di una occorrenza dell'entità "parte" dipende dall'esistenza di una occorrenza dell'entità che la contiene (nell'esempio, la sala di un cinema multisala) e richiede un'identificazione esterna. Nel secondo, l'entità contenuta nell'altra (in questo esempio il tecnico di un team) ha esistenza autonoma, come indicato dalla partecipazione opzionale alla relazione.

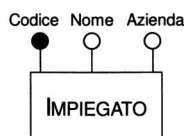


Figura 7.4 Un semplice pattern costituito da una sola entità.

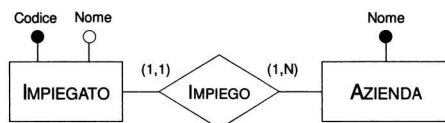


Figura 7.5 Reificazione dell'attributo Azienda in Figura 7.4.

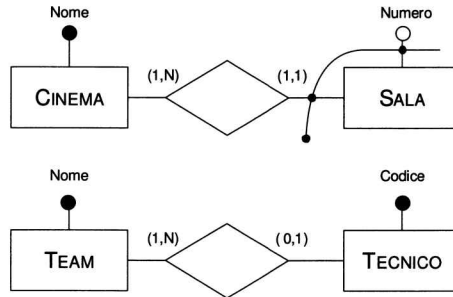


Figura 7.6 Relazioni di tipo "parte-di".

Un'altra situazione piuttosto comune è illustrata negli esempi in Figura 7.7 nei quali le occorrenze di un'entità della relazione sono *istanze di* occorrenze dell'altra entità.

Nel primo caso abbiamo un'entità che descrive il concetto astratto di volo presente sull'orario di una compagnia aerea, con un codice (per esempio AZ610), un'origine (per esempio Roma), una destinazione (per esempio New York) e un orario (per esempio 14:15), e un'altra entità che rappresenta il volo "reale", vale a dire l'istanza di un certo volo in un certo giorno (per esempio il volo AZ610 del 15/12/2013). È facile far confusione tra questi due concetti che però vanno tenuti ben distinti perché giocano ruoli diversi nell'applicazione. L'identificazione del volo reale avviene attraverso la data e, esternamente, il volo di cui è istanza (si assume quindi che lo stesso volo non possa essere ripetuto lo stesso giorno). Un caso analogo è l'altro schema in Figura 7.7 con il quale viene rappresentato il concetto di torneo sportivo (per esempio gli internazionali italiani di tennis) e una sua edizione (per esempio quella del 2014).

Consideriamo ora il caso in cui si utilizza una relazione, tipicamente molti a molti, per descrivere un concetto che lega altri due concetti, come avviene nell'esempio in Figura 7.8 nel quale l'esame è rappresentato da una relazione tra lo studente

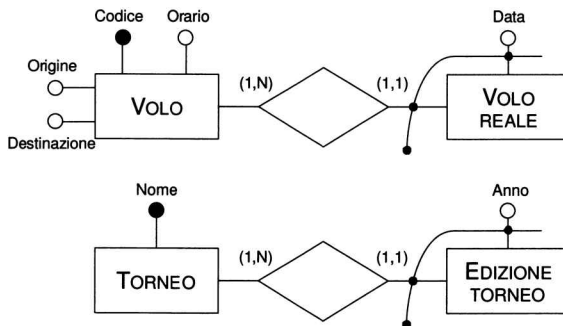


Figura 7.7 Relazione di tipo "istanza-di".

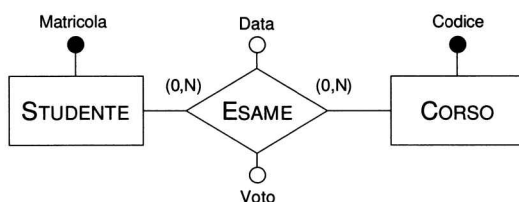


Figura 7.8 Relazione che rappresenta un concetto che lega altri concetti.

e il corso. Come già discusso nel Paragrafo 6.2.1, questa soluzione è valida solo se ogni studente può sostenere una sola volta un certo esame perché, per definizione, una occorrenza della relazione ESAME è un insieme di coppie studente–corso, senza duplicati.

Il fatto che questo concetto abbia degli attributi associati non cambia la situazione, ci suggerisce piuttosto che, soprattutto nel caso in cui uno studente può sostenere più volte lo stesso esame, la soluzione corretta è lo schema in Figura 7.9, nel quale abbiamo reificato la relazione ESAME di Figura 7.8 rappresentandola come entità. In questo caso, l'identificazione di un esame avviene attraverso lo studente, il corso e la data dell'esame.

Una soluzione alternativa che non richiede un'identificazione esterna complessa è costituita dallo schema in Figura 7.10, nel quale è stato introdotto un codice identificativo. Questa scelta semplifica le cose ma bisogna tenere conto del fatto che il codice è un concetto nuovo, non presente nelle specifiche e che quindi dovrà essere opportunamente gestito dal sistema informativo in via di sviluppo. Torneremo a parlare in termini generali di questo aspetto nel capitolo dedicato alla progettazione logica, quando affronteremo il problema della scelta degli identificatori nella traduzione verso il modello relazionale.

Lo schema in Figura 7.11 rappresenta un altro pattern piuttosto comune. Anche qui il concetto di partita può essere inizialmente visto inizialmente come una relazione ricorsiva sull'entità SQUADRA. Ma se, come spesso accade, in un torneo due squadre si incontrano più volte, è necessario reificare la relazione binaria e ottenere

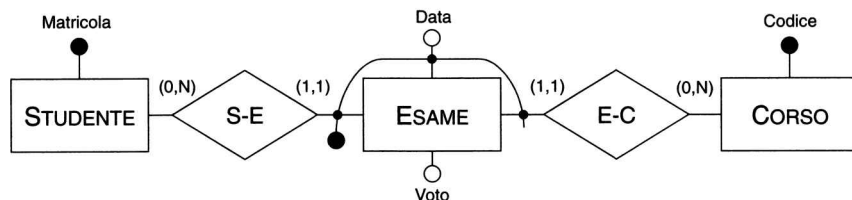


Figura 7.9 Reificazione della relazione in Figura 7.8.

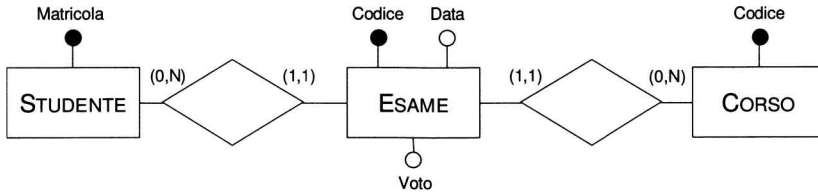


Figura 7.10 Introduzione di un codice nell'entità Esame in Figura 7.9.

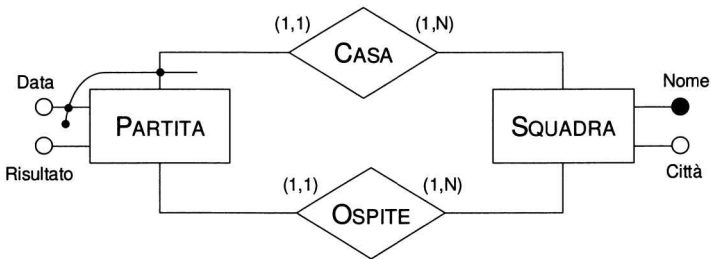


Figura 7.11 Reificazione di relazione ricorsiva.

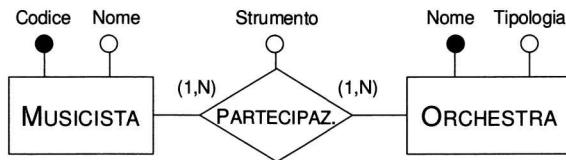


Figura 7.12 Relazione molti a molti con attributo.

lo schema in figura. L'identificazione dell'entità PARTITA coinvolge solo la data e la squadra che gioca in casa perché qui evidentemente si assume che una squadra non possa giocare due partite nello stesso giorno.

Consideriamo ora la relazione molti a molti in Figura 7.12 che rappresenta la partecipazione di un musicista a un'orchestra con un certo strumento. In base a quanto sopra esposto, se il musicista può suonare strumenti diversi ma suona, per ogni orchestra, sempre lo stesso strumento, lo schema è corretto ed è sufficiente una relazione con attributo **Strumento**. Il difetto di questo schema è semmai un altro e ha a che fare con quanto detto per lo schema in Figura 7.5: non stiamo rappresentando

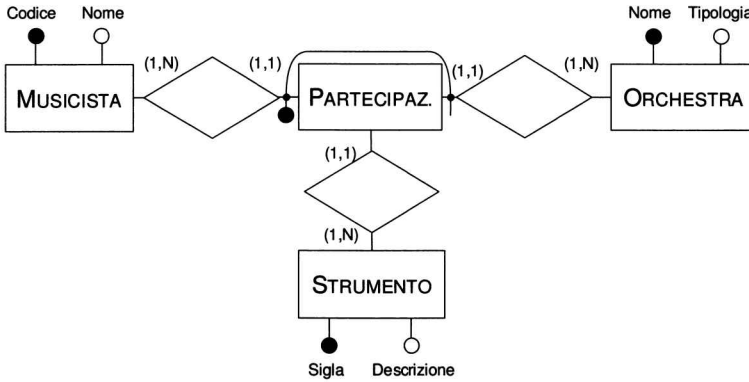


Figura 7.13 Reificazione di attributo di relazione.

esplicitamente il concetto di strumento, che qui è solo una stringa. Se lo strumento è un concetto rilevante per l'applicazione, dobbiamo reificare l'attributo della relazione e, per poterlo fare, dobbiamo reificare anche la relazione. Si ottiene in questo modo lo schema in Figura 7.13.

Passiamo ora ad alcuni pattern che coinvolgono le generalizzazioni. Un primo esempio di uso comune di questo costrutto è quello riportato in Figura 7.14, nel quale si vuole rappresentare un caso particolare di un altro, nell'esempio il sottoinsieme degli impiegati che sono dei manager. Si noti come sia possibile specializzare in questo modo i vari ruoli all'interno di un progetto (l'altra entità dello schema): la gestione è a carico solo dei manager.

In questo schema è ragionevole assumere che un manager può gestire solo un progetto al quale partecipa. Questo implica che che ogni coppia manager-progetto che compare tra le occorrenze della relazione GESTIONE deve comparire anche tra le occorrenze della relazione PARTECIPAZIONE. Questo vincolo però non può essere espresso direttamente sullo schema con un apposito costrutto e va quindi aggiunta una regola alla documentazione dello schema, come abbiamo descritto nel Paragrafo 6.3 del capitolo precedente.

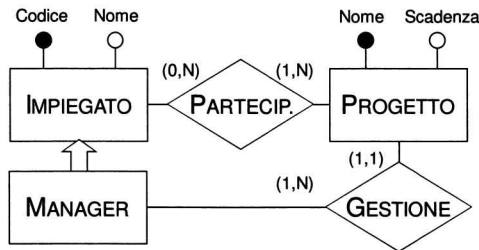


Figura 7.14 Caso particolare di entità.

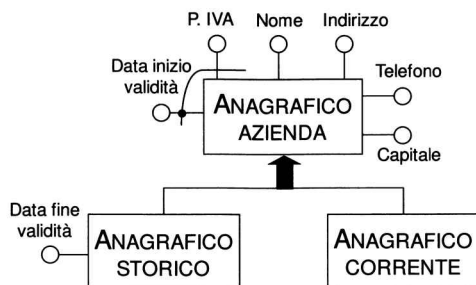


Figura 7.15 Storicizzazione di entità.

L'esempio appena presentato è un caso di sottoinsieme (la generalizzazione ha una sola entità figlia). Il pattern però può essere facilmente generalizzato al caso in cui ci siano più casi particolari da considerare: in questo caso la generalizzazione avrebbe più entità figlie, magari su più livelli, e le varie proprietà (attributi e partecipazione a relazioni) andrebbero distribuite, a seconda della loro specificità, tra le varie entità partecipanti alla generalizzazione.

Lo schema in Figura 7.15 mostra un altro caso di uso comune del costrutto di generalizzazione. Si tratta di uno schema nel quale si vuole gestire la “storicizzazione” di un concetto, nel caso particolare, di un'entità. Nell'esempio vogliamo memorizzare le informazioni correnti di un'azienda, tenendo però traccia dei dati che sono variati. Come suggerito dallo schema, una soluzione piuttosto efficace consiste nell'utilizzare allo scopo due entità con gli stessi attributi: una rappresenta il concetto di interesse con le informazioni aggiornate, l'altra lo “storico”. Le proprietà di queste entità vengono messe a fattor comune mediante una generalizzazione la cui entità genitore rappresenta tutte le informazioni anagrafiche delle aziende, sia quelle correnti sia quelle passate. Vengono inoltre introdotti degli attributi per definire l'intervallo di validità dei dati (data inizio e data fine). L'identificazione si ottiene aggiungendo all'identificatore “naturale” (in questo caso, la partita iva) la data di inizio di validità delle informazioni, ovvero il momento in cui esse sono state introdotte: questo istante diventerà anche la data di fine validità delle informazioni che vengono soppiantate.

Un caso analogo è mostrato nello schema in Figura 7.16. In questo caso si vuole storicizzare un concetto rappresentato da una relazione tra entità, nell'esempio gli impieghi presenti e passati di una persona, in altre parole, il suo curriculum lavorativo. Come nel caso precedente, una possibile soluzione consiste nel rappresentare separatamente i dati correnti e i dati storici e introdurre opportuni attributi per specificare gli intervalli di validità delle informazioni. Notare le differenti cardinalità delle partecipazioni dell'entità PERSONA alle due relazioni. Un'analisi attenta di questo ultimo schema ci fa comprendere che, per i motivi già discussi in precedenza, qui non possiamo rappresentare il fatto che una persona possa aver lavorato, in periodi diversi, per la stessa azienda. Avremmo infatti in questo caso due occorrenze identiche della relazione IMPIEGO PASSATO. La soluzione in questo caso è, ancora una volta, la reificazione delle relazioni. Otteniamo così lo schema in Figura 7.17 che consente

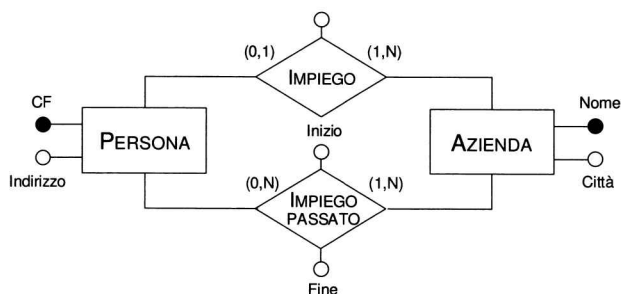


Figura 7.16 Storizzazione di relazione.

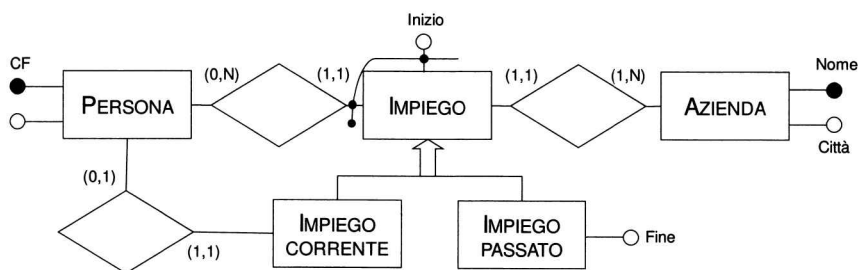


Figura 7.17 Reificazione delle relazioni in Figura 7.16.

di utilizzare una generalizzazione. Anche in questo caso risulta necessario l'inserimento di un vincolo esterno allo schema che impone che tutte le occorrenze della relazione tra PERSONA e IMPIEGO CORRENTE compaiano anche tra le occorrenze della relazione tra PERSONA e IMPIEGO.

L'ultimo esempio di uso comune del costrutto di generalizzazione è quello riportato in Figura 7.18. In questo schema vogliamo rappresentare il fatto che un certo concetto subisce una evoluzione nel tempo che può essere diversa per le diverse occorrenze del concetto. Nell'esempio abbiamo dei progetti che vengono proposti con l'obiettivo di ottenere un finanziamento. Solo alcuni di questi vengono accettati e, per questi, vanno aggiunte ulteriori informazioni quali la data di inizio ufficiale del progetto e il finanziamento effettivamente assegnato. Come si vede dallo schema in figura, il costrutto di generalizzazione si presta bene a modellare questa situazione.

Consideriamo infine la relazione ternaria in Figura 7.19. Come già accennato nel Capitolo 6.2.2, negli schemi E-R le relazioni ternarie si incontrano raramente e relazioni che coinvolgono più di tre entità sono fortemente sconsigliate, perché tipicamente cercano di rappresentare, con un unico costrutto, concetti tra loro indipendenti. Questo aspetto verrà chiarito maggiormente nel Capitolo 9 dedicato alla normalizzazione, una tecnica sistematica che consente di analizzare queste situazioni. Nell'esempio in figura è stata scelta correttamente una relazione ternaria perché si vuole modellare il caso in cui un operatore può effettuare operazioni che consistono in attività diverse svolte in

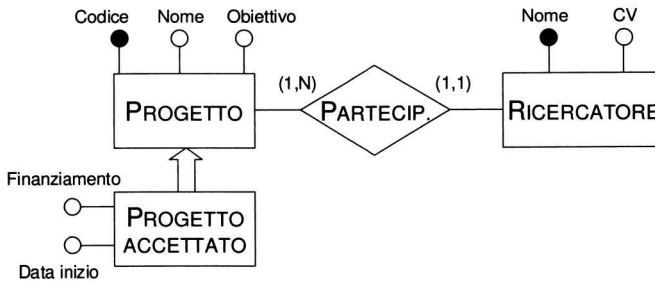


Figura 7.18 Evoluzione di un concetto.

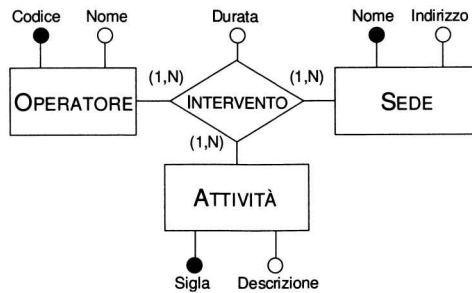


Figura 7.19 Relazione ternaria.

sedi diverse. Inoltre in ogni sede possono operare operatori diversi svolgendo attività diverse. Infine le attività possono essere svolte da operatori diversi e in sedi diverse.

Anche questa relazione, come tutte le altre, può essere reificata e questa operazione si rende necessaria quanto la realtà da modellare è diversa da quella appena descritta. In Figura 7.20 viene riportata la reificazione della relazione ternaria in Figura 7.19. Il nuovo schema modella esattamente la situazione dello schema originario perché la nuova entità risulta identificata da tutte le entità originarie. Cambiando opportunamente l'identificazione siamo però in grado di modellare con questo pattern altre situazioni per le quali la relazione ternaria non sarebbe corretta.

In particolare, se in ogni sede, ogni operatore svolge sempre la stessa attività, l'entità INTERVENTO sarebbe identificata solo dalle entità SEDE e OPERATORE. Se viceversa in ogni sede, ogni attività viene svolta sempre dallo stesso operatore, l'entità INTERVENTO sarebbe identificata solo dalle entità ATTIVITÀ e SEDE. Se infine ogni operatore svolge ogni attività in una sola sede, l'entità INTERVENTO sarebbe identificata solo dalle entità OPERATORE e SEDE.

Infine, lo schema in Figura 7.21 descrive nel modo migliore la situazione in cui la sola entità ATTIVITÀ è identificante, succede cioè che ogni attività viene svolta in una sola sede da un solo operatore.

In questo caso lo schema si semplifica perché il legame tra attività e la sede si può rappresentare separatamente da quello tra l'attività e l'operatore.

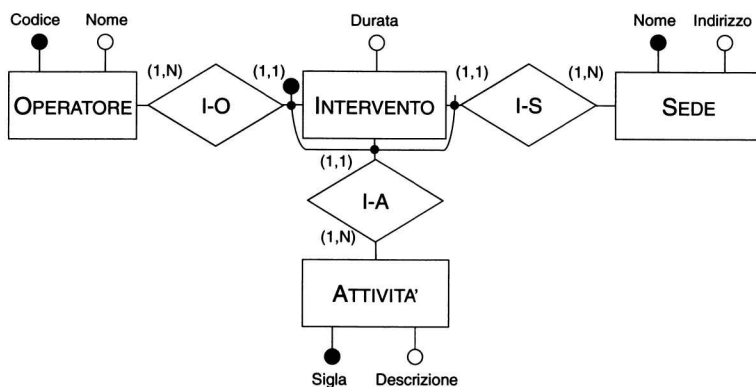


Figura 7.20 Reificazione della relazione ternaria in Figura 7.19.

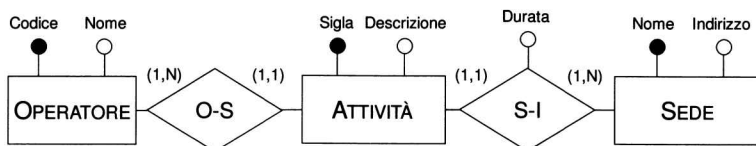


Figura 7.21 Semplificazione dello schema in Figura 7.20.

7.3 Strategie di progetto

Lo sviluppo di uno schema concettuale a partire dalle sue specifiche può essere considerato a tutti gli effetti un processo di ingegnerizzazione e, come tale, risultano a esso applicabili le strategie di progetto utilizzate anche in altre discipline. Vediamo quali sono queste strategie con specifico riferimento alla modellazione di una base di dati.

7.3.1 Strategia top-down

In questa strategia, lo schema concettuale viene prodotto mediante una serie di raffinamenti successivi a partire da uno schema iniziale che descrive tutte le specifiche con pochi concetti molto astratti. Lo schema viene poi via via raffinato mediante opportune trasformazioni che aumentano il dettaglio dei vari concetti presenti. Questo procedimento viene descritto graficamente in Figura 7.22 dove vengono rappresentati i diversi piani di raffinamento del processo: ognuno di questi piani contiene uno schema che descrive le medesime informazioni a un diverso livello di dettaglio. Con questa strategia quindi, tutti gli aspetti presenti nello schema finale sono presenti, in linea di principio, a ogni livello di raffinamento.