

Figura 8.28 Rappresentazione grafica di una traduzione dello schema in Figura 8.20.

A tale riguardo, è possibile adottare un semplice formalismo grafico che permette di rappresentare sia le relazioni con i relativi attributi sia i vincoli di integrità referenziale esistenti tra le varie relazioni. Un esempio di questo tipo di rappresentazione, di facile comprensione, viene dato in Figura 8.27, con riferimento alla traduzione dello schema in Figura 8.17. In questi diagrammi le chiavi delle relazioni sono rappresentate in grassetto, le frecce indicano vincoli di integrità referenziale e la presenza di asterischi sui nomi di attributo indica la possibilità di avere valori nulli.

Si può osservare che, con questo formalismo, si riesce a mantenere traccia delle associazioni dello schema E-R originale. Questo può risultare utile per individuare, in maniera immediata, i *cammini di join*, ovvero le operazioni di join necessarie per ricostruire l'informazione rappresentata dalle associazioni originarie, nel caso dell'esempio, le informazioni sui progetti ai quali gli impiegati partecipano, attraverso il join tra IMPIEGATO, PARTECIPAZIONE e PROGETTO.

Un altro esempio di questi tipi di rappresentazione viene fornito in Figura 8.28, con riferimento alla traduzione dello schema in Figura 8.20.

È interessante osservare come, con questo tipo di rappresentazione, sia possibile rappresentare esplicitamente anche le associazioni dello schema Entità-Relazione di partenza alle quali, nello schema relazionale equivalente, non corrisponde nessuna relazione (l'associazione CONTRATTO nell'esempio in questione).

Come esempio finale, in Figura 8.29 viene riportata la rappresentazione dello schema relazionale ottenuto nel Paragrafo 8.4.5. I legami logici tra le varie relazioni possono essere ora facilmente identificati.

8.5 Un esempio di progettazione logica

Riprendiamo l'esempio presentato nel capitolo precedente relativo alla base di dati della società di formazione, il cui schema concettuale viene riportato, per comodità,

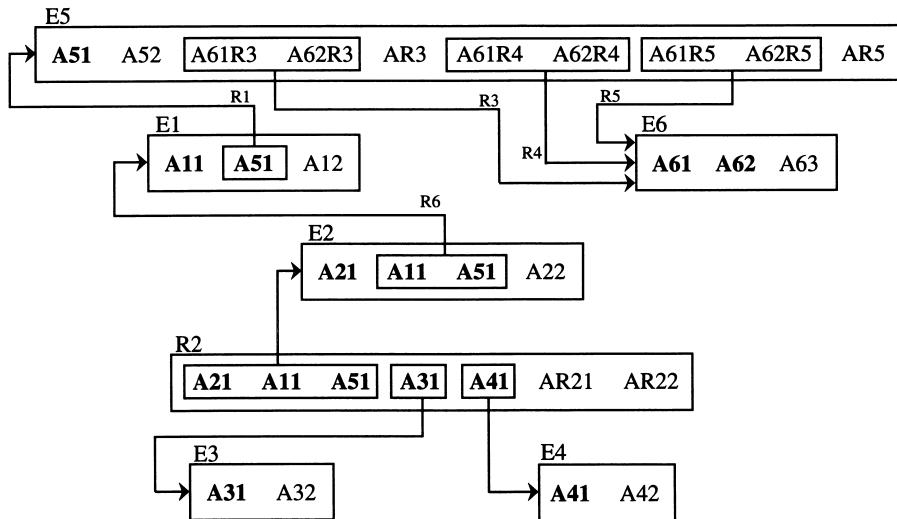


Figura 8.29 Rappresentazione grafica dello schema ottenuto nel Paragrafo 8.4.5.

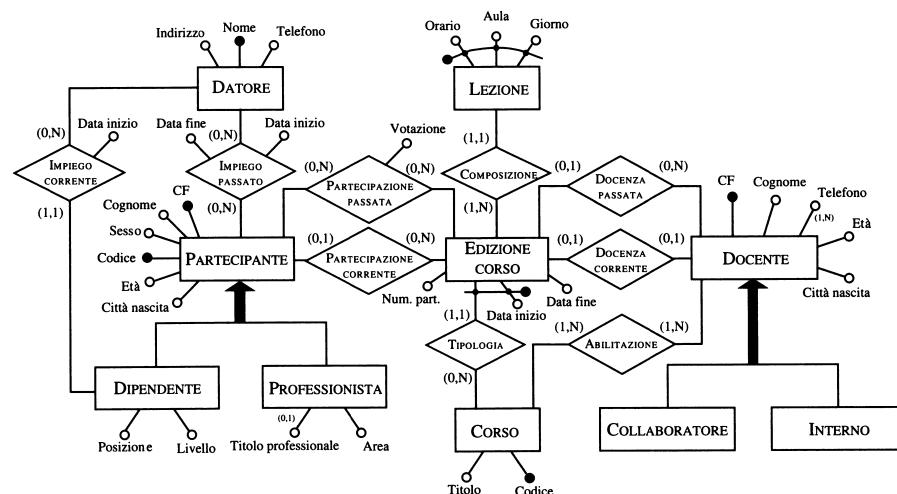


Figura 8.30 Lo schema E-R di una società di formazione.

in Figura 8.30. Le varie ristrutturazioni che discuteremo sono riportate nello schema finale in Figura 8.33.

Sui dati descritti da questo schema erano state previste le seguenti operazioni.

Operazione 1: inserisci un nuovo partecipante indicando tutti i suoi dati.

Operazione 2: assegna un partecipante a una edizione di corso.

Operazione 3: inserisci un nuovo docente indicando tutti i suoi dati e i corsi che può insegnare.

Operazione 4: assegna un docente abilitato a una edizione di un corso.

Operazione 5: stampa tutte le informazioni sulle edizioni passate di un corso con titolo, orari delle lezioni e numero dei partecipanti.

Operazione 6: stampa tutti i corsi offerti, con informazioni sui docenti che possono insegnarli.

Operazione 7: per ogni docente, trova i partecipanti a tutti i corsi da lui insegnati.

Operazione 8: effettua una statistica su tutti i partecipanti a un corso con tutte le informazioni su di essi, sulla edizione alla quale hanno partecipato e la rispettiva votazione.

8.5.1 Fase di ristrutturazione

Supponiamo che i dati di carico siano quelli riportati in Figura 8.31. Eseguiamo, sulla base di questi dati, i vari passi della ristrutturazione.

Analisi delle ridondanze C'è un solo dato ridondante nello schema: l'attributo **Numero di partecipanti** in **EDIZIONE CORSO** che può essere derivato dalle associazioni **PARTECIPAZIONE CORRENTE** e **PARTECIPAZIONE PASSATA**. Questo dato richiede un quantitativo di memoria pari a $4 \times 1000 = 4000$ byte, avendo assunto che sono necessari 4 byte per ogni occorrenza di **EDIZIONE CORSO** per memorizzare il numero di partecipanti. Le operazioni coinvolte con questo dato sono la 2, la 5 e la 8. L'ultima di queste può essere trascurata perché si tratta di una operazione non frequente ed eseguita in modalità batch. Proviamo a valutare il costo delle Operazioni 2 e 5 in caso di presenza e in assenza di dato ridondante. Possiamo dedurre dalla tavola dei volumi che ogni edizione di corso ha, in media, 8 lezioni e 10 partecipanti. Da questi dati sono facilmente calcolabili le tavole degli accessi riportate in Figura 8.32.

Da queste risulta:

- dato ridondante presente: per l'Operazione 2 abbiamo $2 \times 50 = 100$ accessi in lettura e altrettanti in scrittura al giorno mentre, per l'Operazione 5, abbiamo $19 \times 10 = 190$ accessi in lettura al giorno, per un totale di 490 accessi giornalieri (avendo contato doppie le operazioni di scrittura);
- dato ridondante assente: per l'Operazione 2 abbiamo 50 accessi in lettura e altrettanti in scrittura al giorno, mentre, per l'Operazione 5, abbiamo $29 \times 10 = 290$ accessi in lettura al giorno, per un totale di 440 accessi giornalieri (avendo contato doppie le operazioni di scrittura).

Abbiamo quindi, in presenza di ridondanza, degli svantaggi sia in termini di memoria sia di efficienza. Decidiamo quindi di eliminare l'attributo ridondante **Numero di partecipanti** dalla relazione **EDIZIONE CORSO**.

Eliminazione delle gerarchie Nello schema sono presenti due gerarchie: quella relativa ai docenti e quella relativa ai partecipanti. Per i docenti si può notare che le

Tavola dei volumi

Concetto	Tipo	Volume
Lezione	E	8000
Edizione corso	E	1000
Corso	E	200
Docente	E	300
Collaboratore	E	250
Interno	E	50
Partecipante	E	5000
Dipendente	E	4000
Professionista	E	1000
Datore	E	8000
Part. passata	R	10 000
Part. corrente	R	500
Composizione	R	8000
Tipologia	R	1000
Doc. passata	R	900
Doc. corrente	R	100
Abilitazione	R	500
Impiego corrente	R	4000
Impiego passato	R	1000

Tavola delle operazioni

Operazione	Tipo	Frequenza
Op. 1	I	40/giorno
Op. 2	I	50/giorno
Op. 3	I	2/giorno
Op. 4	I	15/giorno
Op. 5	I	10/giorno
Op. 6	I	20/giorno
Op. 7	I	5/sett.
Op. 8	B	10/mese

Figura 8.31 Tavole dei volumi e delle operazioni per lo schema in Figura 8.30.

operazioni che li riguardano, cioè la 3, la 4, la 6 e la 7, non fanno distinzioni tra collaboratori esterni e dipendenti interni della società. Tra l'altro, le entità corrispondenti non hanno attributi specifici che li distinguono. Decidiamo quindi di accorpare le entità figlie della generalizzazione nel genitore aggiungendo un attributo **Tipo** all'entità **DOCENTE** che ha un dominio costituito dai simboli *C* (per Collaboratore) e *I* (per Interno).

Per quanto riguarda i partecipanti, osserviamo che anche in questo caso le operazioni che coinvolgono questo dato (la 1, la 2 e la 8) non fanno sostanziali differenze tra i vari tipi di occorrenze. Possiamo però osservare dallo schema che i professionisti e i dipendenti hanno degli attributi che li distinguono gli uni dagli altri. Risulta quindi

Tavole degli accessi in presenza di ridondanza				Tavole degli accessi in assenza di ridondanza			
Operazione 2				Operazione 2			
Concetto	Costr.	Acc.	Tipo	Concetto	Costr.	Acc.	Tipo
Partecipante	E	1	L	Partecipante	E	1	L
Par. corrente	R	1	S	Par. corrente	R	1	S
Ediz. corso	E	1	L				
Ediz. corso	E	1	S				

Operazione 5				Operazione 5			
Concetto	Costr.	Acc.	Tipo	Concetto	Costr.	Acc.	Tipo
Ed. corso	E	1	L	Ediz. corso	E	1	L
Tipologia	R	1	L	Tipologia	R	1	L
Corso	E	1	L	Corso	E	1	L
Composiz.	R	8	L	Composiz.	R	8	L
Lezione	E	8	L	Lezione	E	8	L
				Par. corrente	R	10	L

Figura 8.32 Tavole degli accessi per lo schema in Figura 8.30.

preferibile lasciare le entità DIPENDENTE e PROFESSIONISTA e aggiungere due associazioni uno a uno tra queste entità e l'entità PARTECIPANTE. In questa maniera, si evita di avere attributi con possibili valori nulli sull'entità genitore della generalizzazione e riduciamo le dimensioni delle relazioni. Il risultato di queste ristrutturazioni e di altre che discuteremo più avanti si può vedere nello schema in Figura 8.33.

Partizionamento/accorpamento di concetti Dall'analisi dei dati e delle operazioni si possono individuare diverse ristrutturazioni di questo tipo. La prima riguarda l'entità EDIZIONE DI CORSO: si può osservare che l'Operazione 5 riguarda solo una frazione delle edizioni, quelle passate, e che le associazioni DOCENZA PASSATA e PARTECIPAZIONE PASSATA fanno riferimento solo a queste edizioni di corso. Si potrebbe quindi pensare, per rendere più efficiente l'operazione suddetta, di decomporre orizzontalmente l'entità in maniera da distinguere le edizioni correnti da quelle passate. L'inconveniente di questa scelta però è che le associazioni COMPOSIZIONE e TIPOLOGIA andrebbero duplicate; inoltre, le Operazioni 7 e 8, che non fanno grosse distinzioni tra le edizioni correnti e quelle passate, risulterebbero più costose perché richiedono la visita di due entità distinte. Decidiamo quindi di non partizionare tale entità.

Due altre possibili ristrutturazioni che si può pensare di effettuare, proprio in conseguenza a quanto detto sulle edizioni dei corsi, sono l'accorpamento delle as-

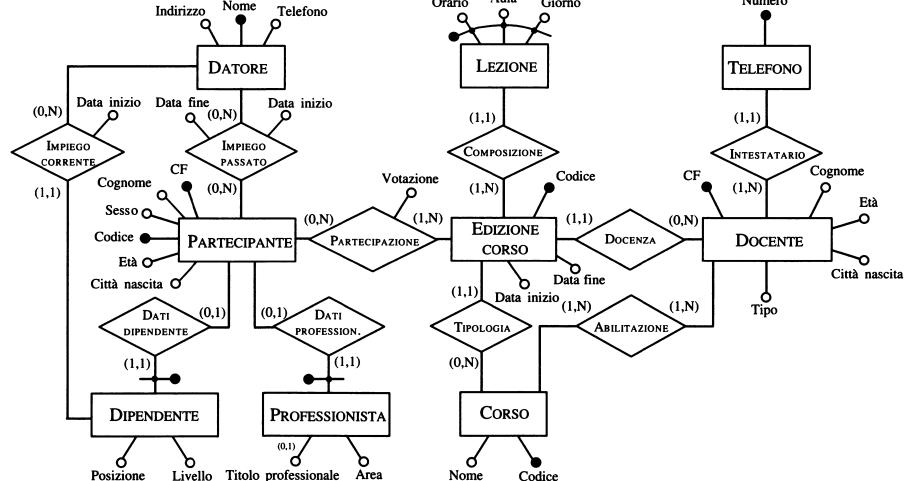


Figura 8.33 Lo schema E-R di Figura 8.30 dopo la fase di ristrutturazione.

sociazioni **DOCENZA PASSATA** e **DOCENZA CORRENTE** e delle associazioni analoghe **PARTECIPAZIONE PASSATA** e **PARTECIPAZIONE CORRENTE**. Si tratta infatti, in entrambi i casi, di due concetti simili (l'unica differenza è di carattere temporale) tra i quali alcune operazioni non fanno differenza (la 7 e la 8). Il loro accorpamento produrrebbe un altro beneficio: non sarebbe necessario trasferire occorrenze da un'associazione a un'altra quando un'edizione di corso termina. Per le partecipazioni ai corsi, un inconveniente è la presenza dell'attributo **Votazione** che non si applica alle partecipazioni correnti e quindi provocherebbe la presenza di valori nulli. Del resto, la tavola dei volumi ci dice che il numero medio di occorrenze dell'entità **PARTECIPAZIONE CORRENTE** è 500 e quindi, supponendo di aver bisogno di 4 byte per memorizzare la votazione, lo spreco di memoria sarebbe di soli 2 kilobyte. Decidiamo quindi di accoppare le due coppie di relazioni come descritto in Figura 8.33. Va aggiunto il vincolo non esprimibile dallo schema che un docente non può insegnare più di una edizione di corso nello stesso periodo e, analogamente, il vincolo che un partecipante non può seguire più di un corso nello stesso periodo.

Infine, bisogna eliminare l'attributo multivalore **Telefono** associato all'entità **DOCENTE**. Per far questo, introduciamo una nuova entità **TELEFONO** legata da una associazione uno a molti con l'entità **DOCENTE**, che viene privata del relativo attributo.

È interessante osservare che le decisioni prese in questa fase ribaltano, in qualche maniera, decisioni prese in fase di progettazione concettuale. Questo però non deve sorprenderci: l'obiettivo della progettazione concettuale è solo quello di rappresentare nella maniera migliore la realtà d'interesse, mentre nella progettazione logica dobbiamo cercare di ottimizzare le prestazioni ed è quasi inevitabile dover rivedere le decisioni prese.

Scelta degli identificatori principali Solo l'entità PARTECIPANTE presenta due identificatori: il codice fiscale e il codice interno. Tra i due è certamente preferibile scegliere il secondo. Infatti, un codice fiscale richiede 16 byte di memoria mentre un codice interno, che serve a distinguere al più 5000 occorrenze (vedi tavola dei volumi), richiede non più di 2 byte.

C'è in effetti un'altra considerazione di carattere pragmatico da fare sugli identificatori e che riguarda l'entità EDIZIONE CORSO. Questa entità è identificata dall'attributo Data inizio e dall'entità CORSO. Ne risulta un identificatore piuttosto pesante che, in una rappresentazione relazionale, deve essere usato per rappresentare due associazioni (PARTECIPAZIONE e DOCENZA) con molte occorrenze. Si può osservare però che ogni corso ha un codice e che, in media, il numero di edizioni di un corso è pari a cinque. Questo significa che è sufficiente aggiungere un intero di una cifra al codice di un corso per avere un identificatore delle edizioni dei corsi, operazione che può essere fatta durante la creazione di una nuova edizione in maniera piuttosto efficiente e sicura. Da questa discussione risulta che è conveniente definire un nuovo identificatore per le edizioni dei corsi che rimpiazza l'identificatore esterno precedente. Questo è un esempio di analisi e ristrutturazione che non rientra in nessuna delle categorie generali viste ma che, nei casi pratici, capita di incontrare.

Abbiamo con questo terminato la fase di ristrutturazione dello schema E-R originale. Lo schema risultante è quello in Figura 8.33.

8.5.2 Traduzione verso il relazionale

Seguendo la strategia di traduzione descritta in questo capitolo, lo schema E-R in Figura 8.33 può essere tradotto nel seguente schema relazionale.

```

EDIZIONECORSO(Codice, DataInizio, DataFine, Corso, Docente)
    LEZIONE(Ora, Aula, Giorno, EdizioneCorso)
    DOCENTE(CF, Cognome, Età, CittàNascita, Tipo)
    TELEFONO(Numero, Docente), CORSO(Codice, Nome)
        ABILITAZIONE(Corso, Docente)
    PARTECIPANTE(Codice, CF, Cognome, Età, CittàNascita, Sesso)
    PARTECIPAZIONE(Partecipante, EdizioneCorso, Votazione*)
        DATORE(Nome, Telefono, Indirizzo)
    IMPIEGOPASSATO(Partecipante, Datore, DataInizio, DataFine)
        PROFESSIONISTA(Partecipante, Area, Titolo*)
    DIPENDENTE(Partecipante, Livello, Posizione, Datore, DataInizio)

```

Lo schema logico ottenuto va naturalmente completato con una documentazione di supporto che descriva, tra l'altro, tutti i vincoli di integrità referenziale che sussistono tra le varie relazioni. Questo può essere fatto usando la notazione grafica introdotta nel Paragrafo 8.4.7.