

# Prefazione

Dopo dieci anni dalla pubblicazione del volume *Basi di dati: strutture ed algoritmi*, l'evoluzione della tecnologia delle basi di dati e la nuova organizzazione della didattica universitaria suggeriscono una sostanziale revisione del materiale.

In questi anni, nel settore si è verificato un interessante cambiamento: le applicazioni che anni fa avrebbero usato dati organizzati in archivi, oggi si sviluppano usando basi di dati, perché i sistemi per la loro gestione (DBMS) sono ormai *la tecnologia* per le applicazioni che usano dati permanenti. Fino a pochi anni fa, essi erano costosi e complessi, pertanto venivano usati comunemente solo da grandi aziende e richiedevano tecnici specializzati per il loro impiego. Oggigiorno, con il successo della tecnologia dei sistemi relazionali, dovuto sia alle loro prestazioni che alla semplicità d'uso, i DBMS sono ormai disponibili su calcolatori di ogni tipo e, sacrificando qualche funzionalità, fanno parte dell'insieme di strumenti minimi per la produttività individuale, in alcuni casi offerti gratuitamente. Questo fenomeno ha aumentato sia la produttività nello sviluppo delle applicazioni tradizionali, sia la varietà delle applicazioni, promuovendo una richiesta sempre più diffusa di tecnici con competenze approfondite sulla tecnologia delle basi di dati. Per questo motivo, mentre nel citato volume si partiva dall'organizzazione di dati in archivi per arrivare a strutture e ad algoritmi per implementare i DBMS, in questa nuova organizzazione del materiale l'attenzione è sin dall'inizio sull'implementazione dei DBMS.

L'altra ragione che ha suggerito una revisione del materiale è stata la recente riforma del sistema universitario che richiede una riorganizzazione della didattica universitaria, con corsi di minor durata e contenuti meno ampi di quelli del passato.

L'area delle basi di dati è di grande interesse per la formazione dei tecnici informatici e per avvicinarsi ad essa si possono guardare i DBMS da due punti di vista: quello dell'utilizzatore della tecnologia, interessato alle funzionalità dei DBMS, al disegno della base di dati e all'uso dell'SQL per recuperare i dati e sviluppare le applicazioni, e quello dell'amministratore di basi di dati complesse, o del realizzatore di strumenti per la gestione di dati, interessati a come funzionano i DBMS. Questi due punti di vista vengono bene affrontati in due corsi brevi che, secondo lo spirito della riforma, forniscono i prerequisiti per approfondire poi altri argomenti dell'area basi di dati. La presente edizione è mirata proprio ad un corso del secondo tipo, e si affianca ad un altro testo utile ad un corso del primo tipo (A. Albano, G. Ghelli e R. Orsini, *Basi di dati relazionali ed a oggetti*, Zanichelli, Bologna, 1997).

## XII Prefazione

### **Approccio**

Si parte con un'analisi dell'architettura dei DBMS relazionali per poi presentare le strutture e gli algoritmi principali per implementare i moduli per la gestione della memoria permanente, del buffer, delle strutture di memorizzazione, dei metodi di accesso, dell'affidabilità, della concorrenza e dell'ottimizzazione dell'esecuzione delle operazioni. Negli ultimi due capitoli, infine, si mostra come rivedere l'architettura, strutture e algoritmi per passare a DBMS relazionali distribuiti e paralleli ed a DBMS relazionali a oggetti. Ogni capitolo contiene numerosi esempi e termina con una serie di esercizi. Gli argomenti sono presentati soffermandosi sulle caratteristiche generali delle soluzioni per poi valutarne le prestazioni.

Aspetto originale del materiale è che, quando occorre esemplificare delle soluzioni, si mostrano quelle adottate nel sistema relazionale JRS, (*Java Relational System*) risultato di un progetto sviluppato in Java presso il Dipartimento di informatica dell'Università di Pisa dall'autore stesso con la collaborazione di diversi studenti attraverso le loro tesi di laurea.

L'obiettivo del progetto è la realizzazione di un sistema relazionale nel linguaggio Java da usare come un laboratorio per scopi didattici, al fine di sperimentare architetture, algoritmi e strutture per basi di dati, a partire da un sistema con un'architettura chiara e facilmente estendibile, per sperimentare algoritmi diversi per realizzare le funzionalità di un DBMS. La scelta del Java è motivata dalle ben note possibilità di avere un sistema portabile su sistemi diversi.

Materiale didattico aggiuntivo, il sistema JRS con la relativa documentazione e le soluzioni degli esercizi sono disponibili sul sito Web del libro (<http://www.di.unipi.it/~albano/costruireDBMS.html>).

### **Udienza**

Questo libro è stato pensato per studenti dei corsi di laurea specialistica in informatica delle facoltà di scienze e ingegneria. Per come sono presentati gli argomenti, prima soffermandosi sugli aspetti generali dei problemi e delle soluzioni e poi sugli aspetti più tecnici e specialistici, il materiale può essere anche usato in corsi introduttivi trascurando gli approfondimenti. I capitoli sui dati multidimensionali, sui DBMS relazionali distribuiti e paralleli, sui DBMS relazionali a oggetti e le sezioni degli altri capitoli contrassegnate da un asterisco riguardano argomenti che sono stati aggiunti per dare una visione più completa sull'implementazione dei DBMS, ma sono adatti a corsi più avanzati perché non trovano spazio in un unico corso della laurea specialistica.

### **Organizzazione del testo**

Nel primo capitolo viene presentata l'architettura dei DBMS relazionali e le caratteristiche dei moduli che la compongono. Vengono inoltre richiamate le nozioni fondamentali del modello relazionale e del linguaggio SQL, per fissare la terminologia e la notazione usata nel resto del libro.

Il Capitolo 2 si sofferma sulle caratteristiche dei gestori della memoria permanente e del buffer, si mostra inoltre come si memorizzano i dati in memoria permanente usando file.

Il Capitolo 3 è dedicato alle più semplici organizzazioni dei dati permanenti, quella seriale e sequenziale, e all'ordinamento di file. Viene anche mostrato come si affronta il problema della valutazione delle prestazioni delle soluzioni.

Il Capitolo 4 tratta le organizzazioni basate su tecniche *hash*, statiche e dinamiche, per recuperare con pochi accessi un record di una collezione conoscendone la chiave.

Il Capitolo 5 continua l'analisi delle organizzazioni per chiave prendendo in considerazione le tecniche basate su indici organizzati con strutture ad albero.

Il Capitolo 6 è dedicato alle organizzazioni per grandi insiemi di record, al fine di agevolare il recupero di piccoli sottoinsiemi caratterizzati da condizioni logiche sui valori di attributi non chiavi. Vengono anche presentate le strutture per la gestione di testi.

Il Capitolo 7 tratta le organizzazioni per dati multidimensionali.

Il Capitolo 8 è dedicato ai metodi di accesso del gestore della macchina fisica e agli operatori forniti dal sistema JRS per realizzare gli operatori fisici dei piani d'accesso.

I Capitoli 9 e 10 sono dedicati al gestore delle transazioni e della loro esecuzione concorrente. Vengono presentate sia le strutture che gli algoritmi per proteggere i dati da malfunzionamenti e interferenze indesiderate in caso di accessi concorrenti.

I Capitoli 11 e 12 sono dedicati agli algoritmi per la realizzazione degli operatori dell'algebra relazionale e dell'ottimizzatore dell'esecuzione dei comandi SQL.

Il Capitolo 13 tratta le soluzioni dei principali problemi da risolvere per realizzare DBMS relazionali che utilizzano la distribuzione dei dati e il parallelismo.

Il Capitolo 14 infine tratta le soluzioni dei principali problemi da risolvere per realizzare DBMS relazionali che supportano il paradigma ad oggetti.

## Ringraziamenti

Essendo il materiale basato su quello della prima edizione, voglio innanzitutto ricordare coloro che mi avevano aiutato in quell'occasione in modi diversi: Paolo Ciaccia, Gualtiero Leoni, Gianni Mainetto, Luigi Mancini, Renzo Orsini, Barbara Pernici, Pasquale Savino e Paolo Tiberio. A questo elenco aggiungo Giorgio Ghelli, che con me ha insegnato per diversi anni questi argomenti ed ha collaborato a preparare la raccolta delle soluzioni degli esercizi, Dario Colazzo, Giovanni Conforti, Paolo Ferragina, Carlo Sartiani, Luca Ventura e gli studenti che con le loro tesi di laurea hanno partecipato con convinzione alla realizzazione del sistema JRS: Lorenzo Brandimarte, Giovanna Colucci, Patrizia Dedato, Stefano Fantechi, Stefano Dinelli, Martina Filippeschi, Simone Marchi, Cinzia Partigliani e Marco Sbaffi. Un ringraziamento particolare va a Stefano Fantechi, che mi ha consentito di usare parti della sua tesi nella stesura dei capitoli sulla gestione delle transazioni, e a Paola Soldi che, pur non conoscendo la materia, ha fatto un'attenta lettura del materiale e mi ha dato numerosi suggerimenti per migliorarne la chiarezza.

#### XIV Prefazione

Infine, l'ultimo e più sentito ringraziamento è per Giovanna, Laura e Andrea che con pazienza si sono rassegnati ad attendere che completassi l'opera.

A. A.