

SEVENTH EDITION

Database System Concepts

Abraham Silberschatz
Henry F. Korth
S. Sudarshan

**Mc
Graw
Hill
Education**



DATABASE SYSTEM CONCEPTS

SEVENTH EDITION

Abraham Silberschatz

Yale University

Henry F. Korth

Lehigh University

S. Sudarshan

Indian Institute of Technology, Bombay





DATABASE SYSTEM CONCEPTS, SEVENTH EDITION

Published by McGraw-Hill Education, 2 Penn Plaza, New York, NY 10121. Copyright © 2020 by McGraw-Hill Education. All rights reserved. Printed in the United States of America. Previous editions © 2011, 2006, and 2002. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of McGraw-Hill Education, including, but not limited to, in any network or other electronic storage or transmission, or broadcast for distance learning.

Some ancillaries, including electronic and print components, may not be available to customers outside the United States.

This book is printed on acid-free paper.

1 2 3 4 5 6 7 8 9 LCR 21 20 19

ISBN 978-0-07-802215-9 (bound edition)

MHID 0-07-802215-0 (bound edition)

ISBN 978-1-260-51504-6 (loose-leaf edition)

MHID 1-260-51504-4 (loose-leaf edition)

Portfolio Manager: *Thomas Scaife Ph.D.*

Product Developers: *Tina Bower & Megan Platt*

Marketing Manager: *Shannon O'Donnell*

Content Project Managers: *Laura Bies & Sandra Schnee*

Buyer: *Susan K. Culbertson*

Design: *Egzon Shagiri*

Content Licensing Specialists: *Shawntel Schmitt & Lorraine Buczek*

Cover Image: © *Pavel Nesvadba/Shutterstock*

Compositor: *Aptara®, Inc.*

All credits appearing on page or at the end of the book are considered to be an extension of the copyright page.

Library of Congress Cataloging-in-Publication Data

Names: Silberschatz, Abraham, author. | Korth, Henry F., author. | Sudarshan, S., author.

Title: Database system concepts / Abraham Silberschatz, Yale University, Henry F. Korth, Lehigh University, S. Sudarshan, Indian Institute of Technology, Bombay.

Description: Seventh edition. | New York, NY: McGraw-Hill, [2020] | Includes bibliographical references.

Identifiers: LCCN 2018060474 | ISBN 9780078022159 (alk. paper) | ISBN 0078022150 (alk. paper)

Subjects: LCSH: Database management.

Classification: LCC QA76.9.D3 S5637 2020 | DDC 005.74—dc23 LC record available at <https://lcn.loc.gov/2018060474>

The Internet addresses listed in the text were accurate at the time of publication. The inclusion of a website does not indicate an endorsement by the authors or McGraw-Hill Education, and McGraw-Hill Education does not guarantee the accuracy of the information presented at these sites.

*To meine schatzi, Valerie
her parents and my dear friends, Steve and Mary Anne
and in memory of my parents, Joseph and Vera*

Avi Silberschatz

*To my wife, Joan
my children, Abigail and Joseph
my mother, Frances
and in memory of my father, Henry*

Hank Korth

*To my wife, Sita
my children, Madhur and Advaith
and my mother, Indira*

S. Sudarshan

About the Authors

Abraham (Avi) Silberschatz is the Sidney J. Weinberg Professor of Computer Science at Yale University. Prior to coming to Yale in 2003, he was the vice president of the Information Sciences Research Center at Bell Labs. He previously held an endowed professorship at the University of Texas at Austin, where he taught until 1993. Silberschatz is a fellow of the ACM, a fellow of the IEEE, and a member of the Connecticut Academy of Science and Engineering. He received the 2002 IEEE Taylor L. Booth Education Award, the 1998 ACM Karl V. Karlstrom Outstanding Educator Award, and the 1997 ACM SIGMOD Contribution Award. Silberschatz was awarded the Bell Laboratories President's Award three times, in 1998, 1999 and 2004. His writings have appeared in numerous journals, conferences, workshops, and book chapters. He has obtained over 48 patents and over 24 grants. He is an author of the textbook *Operating System Concepts*.

Henry F. (Hank) Korth is a Professor of Computer Science and Engineering and co-director of the Computer Science and Business program at Lehigh University. Prior to joining Lehigh, he was director of Database Principles Research at Bell Labs, a vice president of Panasonic Technologies, an associate professor at the University of Texas at Austin, and a research staff member at IBM Research. Korth is a fellow of the ACM and of the IEEE and a winner of the 10-Year Award at the VLDB Conference. His numerous research publications span a wide range of aspects of database systems, including transaction management in parallel and distributed systems, real-time systems, query processing, and the influence on these areas from modern computing architectures. Most recently, his research has addressed issues in the application of blockchains in enterprise databases.

S. Sudarshan is currently the Subrao M. Nilekani Chair Professor at the Indian Institute of Technology, Bombay. He received his Ph.D. at the University of Wisconsin in 1992, and he was a member of the technical staff at Bell Labs before joining IIT Bombay. Sudarshan is a fellow of the ACM. His research spans several areas of database systems, with a focus on query processing and query optimization. His paper on keyword search in databases published in 2002 won the IEEE ICDE Most Influential Paper Award in 2012, and his work on main-memory databases received the Bell Laboratories President's Award in 1999. His current research areas include testing and grading of SQL queries, optimization of database applications by rewriting of imperative code, and query optimization for parallel databases. He has published over 100 papers and obtained 15 patents.

Contents

Chapter 1 Introduction

- 1.1 Database-System Applications 1
- 1.2 Purpose of Database Systems 5
- 1.3 View of Data 8
- 1.4 Database Languages 13
- 1.5 Database Design 17
- 1.6 Database Engine 18
- 1.7 Database and Application Architecture 21
- 1.8 Database Users and Administrators 24
- 1.9 History of Database Systems 25
- 1.10 Summary 29
 - Exercises 31
 - Further Reading 33

PART ONE ■ RELATIONAL LANGUAGES

Chapter 2 Introduction to the Relational Model

- 2.1 Structure of Relational Databases 37
- 2.2 Database Schema 41
- 2.3 Keys 43
- 2.4 Schema Diagrams 46
- 2.5 Relational Query Languages 47
- 2.6 The Relational Algebra 48
- 2.7 Summary 58
 - Exercises 60
 - Further Reading 63

Chapter 3 Introduction to SQL

- 3.1 Overview of the SQL Query Language 65
- 3.2 SQL Data Definition 66
- 3.3 Basic Structure of SQL Queries 71
- 3.4 Additional Basic Operations 79
- 3.5 Set Operations 85
- 3.6 Null Values 89
- 3.7 Aggregate Functions 91
- 3.8 Nested Subqueries 98
- 3.9 Modification of the Database 108
- 3.10 Summary 114
 - Exercises 115
 - Further Reading 124

Chapter 4 Intermediate SQL

- 4.1 Join Expressions 125
- 4.2 Views 137
- 4.3 Transactions 143
- 4.4 Integrity Constraints 145
- 4.5 SQL Data Types and Schemas 153
- 4.6 Index Definition in SQL 164
- 4.7 Authorization 165
- 4.8 Summary 173
 - Exercises 176
 - Further Reading 180

Chapter 5 Advanced SQL

- 5.1 Accessing SQL from a Programming Language 183
- 5.2 Functions and Procedures 198
- 5.3 Triggers 206
- 5.4 Recursive Queries 213
- 5.5 Advanced Aggregation Features 219
- 5.6 Summary 231
 - Exercises 232
 - Further Reading 238

PART TWO ■ DATABASE DESIGN

Chapter 6 Database Design Using the E-R Model

- 6.1 Overview of the Design Process 241
- 6.2 The Entity-Relationship Model 244
- 6.3 Complex Attributes 249
- 6.4 Mapping Cardinalities 252
- 6.5 Primary Key 256
- 6.6 Removing Redundant Attributes in Entity Sets 261
- 6.7 Reducing E-R Diagrams to Relational Schemas 264
- 6.8 Extended E-R Features 271
- 6.9 Entity-Relationship Design Issues 279
- 6.10 Alternative Notations for Modeling Data 285
- 6.11 Other Aspects of Database Design 291
- 6.12 Summary 292
 - Exercises 294
 - Further Reading 300

Chapter 7 Relational Database Design

- 7.1 Features of Good Relational Designs 303
- 7.2 Decomposition Using Functional Dependencies 308
- 7.3 Normal Forms 313
- 7.4 Functional-Dependency Theory 320
- 7.5 Algorithms for Decomposition Using Functional Dependencies 330
- 7.6 Decomposition Using Multivalued Dependencies 336
- 7.7 More Normal Forms 341
- 7.8 Atomic Domains and First Normal Form 342
- 7.9 Database-Design Process 343
- 7.10 Modeling Temporal Data 347
- 7.11 Summary 351
 - Exercises 353
 - Further Reading 360

PART THREE ■ APPLICATION DESIGN AND DEVELOPMENT

Chapter 8 Complex Data Types

- 8.1 Semi-structured Data 365
- 8.2 Object Orientation 376
- 8.3 Textual Data 382
- 8.4 Spatial Data 387
- 8.5 Summary 394
- Exercises 397
- Further Reading 401

Chapter 9 Application Development

- 9.1 Application Programs and User Interfaces 403
- 9.2 Web Fundamentals 405
- 9.3 Servlets 411
- 9.4 Alternative Server-Side Frameworks 416
- 9.5 Client-Side Code and Web Services 421
- 9.6 Application Architectures 429
- 9.7 Application Performance 434
- 9.8 Application Security 437
- 9.9 Encryption and Its Applications 447
- 9.10 Summary 453
- Exercises 455
- Further Reading 462

PART FOUR ■ BIG DATA ANALYTICS

Chapter 10 Big Data

- 10.1 Motivation 467
- 10.2 Big Data Storage Systems 472
- 10.3 The MapReduce Paradigm 483
- 10.4 Beyond MapReduce: Algebraic Operations 494
- 10.5 Streaming Data 500
- 10.6 Graph Databases 508
- 10.7 Summary 511
- Exercises 513
- Further Reading 516

Chapter 11 Data Analytics

- 11.1 Overview of Analytics 519
- 11.2 Data Warehousing 521
- 11.3 Online Analytical Processing 527
- 11.4 Data Mining 540
- 11.5 Summary 550
- Exercises 552
- Further Reading 555

PART FIVE ■ STORAGE MANAGEMENT AND INDEXING

Chapter 12 Physical Storage Systems

- 12.1 Overview of Physical Storage Media 559
- 12.2 Storage Interfaces 562
- 12.3 Magnetic Disks 563
- 12.4 Flash Memory 567
- 12.5 RAID 570
- 12.6 Disk-Block Access 577
- 12.7 Summary 580
- Exercises 582
- Further Reading 584

Chapter 13 Data Storage Structures

- 13.1 Database Storage Architecture 587
- 13.2 File Organization 588
- 13.3 Organization of Records in Files 595
- 13.4 Data-Dictionary Storage 602
- 13.5 Database Buffer 604
- 13.6 Column-Oriented Storage 611
- 13.7 Storage Organization in Main-Memory Databases 615
- 13.8 Summary 617
- Exercises 619
- Further Reading 621

Chapter 14 Indexing

- 14.1 Basic Concepts 623
- 14.2 Ordered Indices 625
- 14.3 B⁺-Tree Index Files 634
- 14.4 B⁺-Tree Extensions 650
- 14.5 Hash Indices 658
- 14.6 Multiple-Key Access 661
- 14.7 Creation of Indices 664
- 14.8 Write-Optimized Index Structures 665
- 14.9 Bitmap Indices 670
- 14.10 Indexing of Spatial and Temporal Data 672
- 14.11 Summary 677
- Exercises 679
- Further Reading 683

PART SIX ■ QUERY PROCESSING AND OPTIMIZATION

Chapter 15 Query Processing

- 15.1 Overview 689
- 15.2 Measures of Query Cost 692
- 15.3 Selection Operation 695
- 15.4 Sorting 701
- 15.5 Join Operation 704
- 15.6 Other Operations 719
- 15.7 Evaluation of Expressions 724
- 15.8 Query Processing in Memory 731
- 15.9 Summary 734
- Exercises 736
- Further Reading 740

Chapter 16 Query Optimization

- 16.1 Overview 743
- 16.2 Transformation of Relational Expressions 747
- 16.3 Estimating Statistics of Expression Results 757
- 16.4 Choice of Evaluation Plans 766
- 16.5 Materialized Views 778
- 16.6 Advanced Topics in Query Optimization 783
- 16.7 Summary 787
- Exercises 789
- Further Reading 794

PART SEVEN ■ TRANSACTION MANAGEMENT

Chapter 17 Transactions

- 17.1 Transaction Concept 799
- 17.2 A Simple Transaction Model 801
- 17.3 Storage Structure 804
- 17.4 Transaction Atomicity and Durability 805
- 17.5 Transaction Isolation 807
- 17.6 Serializability 812
- 17.7 Transaction Isolation and Atomicity 819
- 17.8 Transaction Isolation Levels 821
- 17.9 Implementation of Isolation Levels 823
- 17.10 Transactions as SQL Statements 826
- 17.11 Summary 828
- Exercises 831
- Further Reading 834

Chapter 18 Concurrency Control

- 18.1 Lock-Based Protocols 835
- 18.2 Deadlock Handling 849
- 18.3 Multiple Granularity 853
- 18.4 Insert Operations, Delete Operations, and Predicate Reads 857
- 18.5 Timestamp-Based Protocols 861
- 18.6 Validation-Based Protocols 866
- 18.7 Multiversion Schemes 869
- 18.8 Snapshot Isolation 872
- 18.9 Weak Levels of Consistency in Practice 880
- 18.10 Advanced Topics in Concurrency Control 883
- 18.11 Summary 894
- Exercises 899
- Further Reading 904

Chapter 19 Recovery System

- 19.1 Failure Classification 907
- 19.2 Storage 908
- 19.3 Recovery and Atomicity 912
- 19.4 Recovery Algorithm 922
- 19.5 Buffer Management 926
- 19.6 Failure with Loss of Non-Volatile Storage 930
- 19.7 High Availability Using Remote Backup Systems 931
- 19.8 Early Lock Release and Logical Undo Operations 935
- 19.9 ARIES 941
- 19.10 Recovery in Main-Memory Databases 947
- 19.11 Summary 948
- Exercises 952
- Further Reading 956

PART EIGHT ■ PARALLEL AND DISTRIBUTED DATABASES

Chapter 20 Database-System Architectures

- 20.1 Overview 961
- 20.2 Centralized Database Systems 962
- 20.3 Server System Architectures 963
- 20.4 Parallel Systems 970
- 20.5 Distributed Systems 986
- 20.6 Transaction Processing in Parallel and Distributed Systems 989
- 20.7 Cloud-Based Services 990
- 20.8 Summary 995
- Exercises 998
- Further Reading 1001

Chapter 21 Parallel and Distributed Storage

- 21.1 Overview 1003
- 21.2 Data Partitioning 1004
- 21.3 Dealing with Skew in Partitioning 1007
- 21.4 Replication 1013
- 21.5 Parallel Indexing 1017
- 21.6 Distributed File Systems 1019
- 21.7 Parallel Key-Value Stores 1023
- 21.8 Summary 1032
- Exercises 1033
- Further Reading 1036

Chapter 22 Parallel and Distributed Query Processing

- 22.1 Overview 1039
- 22.2 Parallel Sort 1041
- 22.3 Parallel Join 1043
- 22.4 Other Operations 1048
- 22.5 Parallel Evaluation of Query Plans 1052
- 22.6 Query Processing on Shared-Memory Architectures 1061
- 22.7 Query Optimization for Parallel Execution 1064
- 22.8 Parallel Processing of Streaming Data 1070
- 22.9 Distributed Query Processing 1076
- 22.10 Summary 1086
- Exercises 1089
- Further Reading 1093

Chapter 23 Parallel and Distributed Transaction Processing

- 23.1 Distributed Transactions 1098
- 23.2 Commit Protocols 1100
- 23.3 Concurrency Control in Distributed Databases 1111
- 23.4 Replication 1121
- 23.5 Extended Concurrency Control Protocols 1129
- 23.6 Replication with Weak Degrees of Consistency 1133
- 23.7 Coordinator Selection 1146
- 23.8 Consensus in Distributed Systems 1150
- 23.9 Summary 1162
- Exercises 1165
- Further Reading 1168

PART NINE ■ ADVANCED TOPICS

Chapter 24 Advanced Indexing Techniques

- | | | | |
|--|------|-------------------|------|
| 24.1 Bloom Filter | 1175 | 24.5 Hash Indices | 1190 |
| 24.2 Log-Structured Merge Tree and
Variants | 1176 | 24.6 Summary | 1203 |
| 24.3 Bitmap Indices | 1182 | Exercises | 1205 |
| 24.4 Indexing of Spatial Data | 1186 | Further Reading | 1206 |

Chapter 25 Advanced Application Development

- | | | | |
|---|------|------------------------------------|------|
| 25.1 Performance Tuning | 1210 | 25.5 Distributed Directory Systems | 1240 |
| 25.2 Performance Benchmarks | 1230 | 25.6 Summary | 1243 |
| 25.3 Other Issues in Application
Development | 1234 | Exercises | 1245 |
| 25.4 Standardization | 1237 | Further Reading | 1248 |

Chapter 26 Blockchain Databases

- | | | | |
|--|------|------------------------------|------|
| 26.1 Overview | 1252 | 26.6 Smart Contracts | 1269 |
| 26.2 Blockchain Properties | 1254 | 26.7 Performance Enhancement | 1274 |
| 26.3 Achieving Blockchain Properties via
Cryptographic Hash Functions | 1259 | 26.8 Emerging Applications | 1276 |
| 26.4 Consensus | 1263 | 26.9 Summary | 1279 |
| 26.5 Data Management in a Blockchain | 1267 | Exercises | 1280 |
| | | Further Reading | 1282 |

PART TEN ■ APPENDIX A

Appendix A Detailed University Schema 1287

Index 1299

PART ELEVEN ■ ONLINE CHAPTERS

Chapter 27 Formal Relational Query Languages

Chapter 28 Advanced Relational Database Design

Chapter 29 Object-Based Databases

Chapter 30 XML

Chapter 31 Information Retrieval

Chapter 32 PostgreSQL

Preface

Database management has evolved from a specialized computer application to a central component of virtually all enterprises, and, as a result, knowledge about database systems has become an essential part of an education in computer science. In this text, we present the fundamental concepts of database management. These concepts include aspects of database design, database languages, and database-system implementation.

This text is intended for a first course in databases at the junior or senior undergraduate, or first-year graduate, level. In addition to basic material for a first course, the text contains advanced material that can be used for course supplements, or as introductory material for an advanced course.

We assume only a familiarity with basic data structures, computer organization, and a high-level programming language such as Java, C, C++, or Python. We present concepts as intuitive descriptions, many of which are based on our running example of a university. Important theoretical results are covered, but formal proofs are omitted. In place of proofs, figures and examples are used to suggest why a result is true. Formal descriptions and proofs of theoretical results may be found in research papers and advanced texts that are referenced in the bibliographical notes.

The fundamental concepts and algorithms covered in the book are often based on those used in existing commercial or experimental database systems. Our aim is to present these concepts and algorithms in a general setting that is not tied to one particular database system, though we do provide references to specific systems where appropriate.

In this, the seventh edition of *Database System Concepts*, we have retained the overall style of the prior editions while evolving the content and organization to reflect the changes that are occurring in the way databases are designed, managed, and used. One such major change is the extensive use of “Big Data” systems. We have also taken into account trends in the teaching of database concepts and made adaptations to facilitate these trends where appropriate.

Among the notable changes in this edition are:

- Extensive coverage of Big Data systems, from the user perspective (Chapter 10), as well as from an internal perspective (Chapter 20 through Chapter 23), with extensive additions and modifications compared to the sixth edition.
- A new chapter entitled “Blockchain Databases” (Chapter 26) that introduces blockchain technology and its growing role in enterprise applications. An important focus in this chapter is the interaction between blockchain systems and database systems.
- Updates to all chapters covering database internals (Chapter 12 through Chapter 19) to reflect current-generation technology, such as solid-state disks, main-memory databases, multi-core systems, and column-stores.
- Enhanced coverage of semi-structured data management using JSON, RDF, and SPARQL (Section 8.1).
- Updated coverage of temporal data (in Section 7.10), data analytics (Chapter 11), and advanced indexing techniques such as write-optimized indices (Section 14.8 and Section 24.2).
- Reorganization and update of chapters to better support courses with a significant hands-on component (which we strongly recommend for any database course), including use of current-generation application development tools and Big Data systems such as Apache Hadoop and Spark.

These and other updates have arisen from the many comments and suggestions we have received from readers of the sixth edition, our students at Yale University, Lehigh University, and IIT Bombay, and our own observations and analyses of developments in database technology.

Content of This Book

The text is organized in eleven major parts.

- **Overview** (Chapter 1). Chapter 1 provides a general overview of the nature and purpose of database systems. We explain how the concept of a database system has developed, what the common features of database systems are, what a database system does for the user, and how a database system interfaces with operating systems. We also introduce an example database application: a university organization consisting of multiple departments, instructors, students, and courses. This application is used as a running example throughout the book. This chapter is motivational, historical, and explanatory in nature.

- **Part 1: Relational Model and SQL** (Chapter 2 through Chapter 5). Chapter 2 introduces the relational model of data, covering basic concepts such as the structure of relational databases, database schemas, keys, schema diagrams, relational query languages, relational operations, and the relational algebra. Chapter 3, Chapter 4, and Chapter 5 focus on the most influential of the user-oriented relational languages: SQL. The chapters in this part describe data manipulation: queries, updates, insertions, and deletions, assuming a schema design has been provided. Although data-definition syntax is covered in detail here, schema design issues are deferred to Part 2.
- **Part 2: Database Design** (Chapter 6 and Chapter 7). Chapter 6 provides an overview of the database-design process and a detailed description of the entity-relationship data model. The entity-relationship data model provides a high-level view of the issues in database design and of the problems encountered in capturing the semantics of realistic applications within the constraints of a data model. UML class-diagram notation is also covered in this chapter. Chapter 7 introduces relational database design. The theory of functional dependencies and normalization is covered, with emphasis on the motivation and intuitive understanding of each normal form. This chapter begins with an overview of relational design and relies on an intuitive understanding of logical implication of functional dependencies. This allows the concept of normalization to be introduced prior to full coverage of functional-dependency theory, which is presented later in the chapter. Instructors may choose to use only this initial coverage without loss of continuity. Instructors covering the entire chapter will benefit from students having a good understanding of normalization concepts to motivate them to learn some of the challenging concepts of functional-dependency theory. The chapter ends with a section on modeling of temporal data.
- **Part 3: Application Design and Development** (Chapter 8 and Chapter 9). Chapter 8 discusses several complex data types that are particularly important for application design and development, including semi-structured data, object-based data, textual data, and spatial data. Although the popularity of XML in a database context has been diminishing, we retain an introduction to XML, while adding coverage of JSON, RDF, and SPARQL. Chapter 9 discusses tools and technologies that are used to build interactive web-based and mobile database applications. This chapter includes detailed coverage on both the server side and the client side. Among the topics covered are servlets, JSP, Django, JavaScript, and web services. Also discussed are application architecture, object-relational mapping systems including Hibernate and Django, performance (including caching using memcached and Redis), and the unique challenges in ensuring web-application security.
- **Part 4: Big Data Analytics** (Chapter 10 and Chapter 11). Chapter 10 provides an overview of large-scale data-analytic applications, with a focus on how those applications place distinct demands on data management compared with the de-

mands of traditional database applications. The chapter then discusses how those demands are addressed. Among the topics covered are Big Data storage systems including distributed file systems, key-value stores and NoSQL systems, MapReduce, Apache Spark, streaming data, and graph databases. The connection of these systems and concepts with database concepts introduced earlier is emphasized. Chapter 11 discusses the structure and use of systems designed for large-scale data analysis. After first explaining the concepts of data analytics, business intelligence, and decision support, the chapter discusses the structure of a data warehouse and the process of gathering data into a warehouse. The chapter next covers usage of warehouse data in OLAP applications followed by a survey of data-mining algorithms and techniques.

- **Part 5: Storage Management and Indexing** (Chapter 12 through Chapter 14). Chapter 12 deals with storage devices and how the properties of those devices influence database physical organization and performance. Chapter 13 deals with data-storage structures, including file organization and buffer management. A variety of data-access techniques are presented in Chapter 14. Multilevel index-based access is described, culminating in detailed coverage of B⁺-trees. The chapter then covers index structures for applications where the B⁺-tree structure is less appropriate, including write-optimized indices such as LSM trees and buffer trees, bitmap indices, and the indexing of spatial data using k-d trees, quadrees and R-trees.
- **Part 6: Query Processing and Optimization** (Chapter 15 and Chapter 16). Chapter 15 and Chapter 16 address query-evaluation algorithms and query optimization. Chapter 15 focuses on algorithms for the implementation of database operations, particularly the wide range of join algorithms, which are designed to work on very large data that may not fit in main-memory. Query processing techniques for main-memory databases are also covered in this chapter. Chapter 16 covers query optimization, starting by showing how query plans can be transformed to other equivalent plans by using transformation rules. The chapter then describes how to generate estimates of query execution costs, and how to efficiently find query execution plans with the lowest cost.
- **Part 7: Transaction Management** (Chapter 17 through Chapter 19). Chapter 17 focuses on the fundamentals of a transaction-processing system: atomicity, consistency, isolation, and durability. It provides an overview of the methods used to ensure these properties, including log-based recovery and concurrency control using locking, timestamp-based techniques, and snapshot isolation. Courses requiring only a survey of the transaction concept can use Chapter 17 on its own without the other chapters in this part; those chapters provide significantly greater depth. Chapter 18 focuses on concurrency control and presents several techniques for ensuring serializability, including locking, timestamping, and optimistic (validation) techniques. Multiversion concurrency control techniques, including the widely used snapshot isolation technique, and an extension of the technique that

guarantees serializability, are also covered. This chapter also includes discussion of weak levels of consistency, concurrency on index structures, concurrency in main-memory database systems, long-duration transactions, operation-level concurrency, and real-time transaction processing. Chapter 19 covers the primary techniques for ensuring correct transaction execution despite system crashes and storage failures. These techniques include logs, checkpoints, and database dumps, as well as high availability using remote backup systems. Recovery with early lock release, and the widely used ARIES algorithm are also presented. This chapter includes discussion of recovery in main-memory database systems and the use of NVRAM.

- **Part 8: Parallel and Distributed Databases** (Chapter 20 through Chapter 23). Chapter 20 covers computer-system architecture, and describes the influence of the underlying computer system on the database system. We discuss centralized systems, client-server systems, parallel and distributed architectures, and cloud-based systems in this chapter. The remaining three chapters in this part address distinct aspects of parallel and distributed databases, with Chapter 21 covering storage and indexing, Chapter 22 covering query processing, and Chapter 23 covering transaction management. Chapter 21 includes discussion of partitioning and data skew, replication, parallel indexing, distributed file systems (including the Hadoop file system), and parallel key-value stores. Chapter 22 includes discussion of parallelism both among multiple queries and within a single query. It covers parallel and distributed sort and join, MapReduce, pipelining, the Volcano exchange-operator model, thread-level parallelism, streaming data, and the optimization of geographically distributed queries. Chapter 23 includes discussion of traditional distributed consensus such as two-phase commit and more sophisticated solutions including Paxos and Raft. It covers a variety of algorithms for distributed concurrency control, including replica management and weaker degrees of consistency. The trade-offs implied by the CAP theorem are discussed along with the means of detecting inconsistency using version vectors and Merkle trees.
- **Part 9: Advanced Topics** (Chapter 24 through Chapter 26). Chapter 24 expands upon the coverage of indexing in Chapter 14 with detailed coverage of the LSM tree and its variants, bitmap indices, spatial indexing, and dynamic hashing techniques. Chapter 25 expands upon the coverage of Chapter 9 with a discussion of performance tuning, benchmarking, testing, and migration from legacy systems, standardization, and distributed directory systems. Chapter 26 looks at blockchain technology from a database perspective. It describes blockchain data structures and the use of cryptographic hash functions and public-key encryption to ensure the blockchain properties of anonymity, irrefutability, and tamper resistance. It describes and compares the distributed consensus algorithms used to ensure decentralization, including proof-of-work, proof-of-stake, and Byzantine consensus. Much of the chapter focuses on the features that make blockchain an important database concept, including the role of permissioned blockchains, the encoding

of business logic and agreements in smart contracts, and interoperability across blockchains. Techniques for achieving database-scale transaction-processing performance are discussed. A final section surveys current and contemplated enterprise blockchain applications.

- **Part 10: Appendix.** Appendix A presents details of our university schema, including the full schema, DDL, and all the tables.
- **Part 11: Online Chapters** (Chapter 27 through Chapter 32) available online at db-book.com. We provide six chapters that cover material that is of historical nature or is advanced; these chapters are available only online. Chapter 27 covers “pure” query languages: the tuple and domain relational calculus and Datalog, which has a syntax modeled after the Prolog language. Chapter 28 covers advanced topics in relational database design, including the theory of multivalued dependencies and fourth normal form, as well as higher normal forms. Chapter 29 covers object-based databases and more complex data types such as array, and multiset types, as well as tables that are not in 1NF. Chapter 30 expands on the coverage in Chapter 8 of XML. Chapter 31 covers information retrieval, which deals with querying of unstructured textual data. Chapter 32 provides an overview of the PostgreSQL database system, and is targeted at courses focusing on database internals. The chapter is likely to be particularly useful for supporting student projects that work with the open-source code base of the PostgreSQL database.

At the end of each chapter we provide references in a section titled *Further Reading*. This section is intentionally abbreviated and provides references that allow students to continue their study of the material covered in the chapter or to learn about new developments in the area covered by the chapter. On occasion, the further reading section includes original source papers that have become classics of which everyone should be aware. Detailed bibliographical notes for each chapter are available online, and provide references for readers who wish to go into further depth on any of the topics covered in the chapter.

The Seventh Edition

The production of this seventh edition has been guided by the many comments and suggestions we received concerning the earlier editions, by our own observations while teaching at Yale University, Lehigh University, and IIT Bombay, and by our analysis of the directions in which database technology is evolving.

We provided a list of the major new features of this edition earlier in this preface; these include coverage of extensive coverage of Big Data, updates to all chapters to reflect current generation hardware technology, semi-structured data management, advanced indexing techniques, and a new chapter on blockchain databases. Beyond these major changes, we revised the material in each chapter, bringing the older material

up-to-date, adding discussions on recent developments in database technology, and improving descriptions of topics that students found difficult to understand. We have also added new exercises and updated references.

For instructors who previously used the sixth edition, we list the more significant changes below:

- Relational algebra has been moved into Chapter 2, to help students better understand relational operations that form the basis of query languages such as SQL. Deeper coverage of relational algebra also aids in understanding the algebraic operators needed for discussion later of query processing and optimization. The two variants of the relational calculus are now in an online chapter, since we believe they are now of value only to more theoretically oriented courses, and can be omitted by most database courses.
- The SQL chapters now include more details of database-system specific SQL variations, to aid students carrying out practical assignments. Connections between SQL and the multiset relational algebra are also covered in more detail. Chapter 4 now covers all the material concerning joins, whereas previously natural join was in the preceding chapter. Coverage of sequences used to generate unique key values, and coverage of row-level security have also been added to this chapter. Recent extensions to the JDBC API that are particularly useful are now covered in Chapter 5; coverage of OLAP has been moved from this chapter to Chapter 11.
- Chapter 6 has been modified to cover E-R diagrams along with E-R concepts, instead of first covering the concepts and then introducing E-R diagrams as was done in earlier editions. We believe this will help students better comprehend the E-R model.
- Chapter 7 now has improved coverage of temporal data modeling, including SQL:2011 temporal database features.
- Chapter 8 is a new chapter that covers complex data types, including semi-structured data, such as XML, JSON, RDF, and SPARQL, object-based data, textual data, and spatial data. Object-based databases, XML, and information retrieval on textual data were covered in detail in the sixth edition; these topics have been abbreviated and covered in Chapter 8, while the original chapters from the sixth edition have now been made available online.
- Chapter 9 has been significantly updated to reflect modern application development tools and techniques, including extended coverage of JavaScript and JavaScript libraries for building dynamic web interfaces, application development in Python using the Django framework, coverage of web services, and disconnection operations using HTML5. Object-relation mapping using Django has been added, as also discussion of techniques for developing high-performance applications that can handle large transaction loads.

- Chapter 10 is a new chapter on Big Data, covering Big Data concepts and tools from a user perspective. Big Data storage systems, the MapReduce paradigm, Apache Hadoop and Apache Spark, and streaming and graph databases are covered in this chapter. The goal is to enable readers to use Big Data systems, with only a summary coverage of what happens behind the scenes. Big Data internals are covered in detail in later chapters.
- The chapter on storage and file structure has been split into two chapters. Chapter 12 which covers storage has been updated with new technology, including expanded coverage of flash memory, column-oriented storage, and storage organization in main-memory databases. Chapter 13, which covers data storage structures has been expanded, and now covers details such as free-space maps, partitioning, and most importantly column-oriented storage.
- Chapter 14 on indexing now covers write-optimized index structures including the LSM tree and its variants, and the buffer tree, which are seeing increasing usage. Spatial indices are now covered briefly in this chapter. More detailed coverage of LSM trees and spatial indices is provided in Chapter 24, which covers advanced indexing techniques. Bitmap indices are now covered in brief in Chapter 14, while more detailed coverage has been moved to Chapter 24. Dynamic hashing techniques have been moved into Chapter 24, since they are of limited practical importance today.
- Chapter 15 on query processing has significantly expanded coverage of pipelining in query processing, new material on query processing in main-memory, including query compilation, as well as brief coverage of spatial joins. Chapter 16 on query optimization has more examples of equivalence rules for operators such as outer joins and aggregates, has updated material on statistics for cost estimation, an improved presentation of the join-order optimization algorithm. Techniques for decorrelating nested subqueries using semijoin and antijoin operations have also been added.
- Chapter 18 on concurrency control has new material on concurrency control in main-memory. Chapter 19 on recovery now gives more importance to high availability using remote backup systems.
- Our coverage of parallel and distributed databases has been completely revamped. Because of the evolution of these two areas into a continuum from low-level parallelism to geographically distributed systems, we now present these topics together.
 - Chapter 20 on database architectures has been significantly updated from the earlier edition, including new material on practical interconnection networks like the tree-like (or fat-tree) architecture, and significantly expanded and updated material on shared-memory architectures and cache coherency. There is an entirely new section on cloud-based services, covering virtual machines and containers, platform-as-a-service, software-as-a-service, and elasticity.

- Chapter 21 covers parallel and distributed storage; while a few parts of this chapter were present in the sixth edition, such as partitioning techniques, everything else in this chapter is new.
- Chapter 22 covers parallel and distributed query processing. Again only a few sections of this chapter, such as parallel algorithms for sorting, join, and a few other relational operations, were present in the sixth edition, almost everything else in this chapter is new.
- Chapter 23 covers parallel and distributed transaction processing. A few parts of this chapter, such as the sections on 2PC, persistent messaging, and concurrency control in distributed databases, are new but almost everything else in this chapter is new.

As in the sixth edition, we facilitate the following of our running example by listing the database schema and the sample relation instances for our university database together in Appendix A as well as where they are used in the various regular chapters. In addition, we provide, on our web site db-book.com, SQL data-definition statements for the entire example, along with SQL statements to create our example relation instances. This encourages students to run example queries directly on a database system and to experiment with modifying those queries. All topics not listed above are updated from the sixth edition, though their overall organization is relatively unchanged.

End of Chapter Material

Each chapter has a list of review terms, in addition to a summary, which can help readers review key topics covered in the chapter.

As in the sixth edition, the exercises are divided into two sets: **practice exercises** and **exercises**. The solutions for the practice exercises are publicly available on the web site of the book. Students are encouraged to solve the practice exercises on their own and later use the solutions on the web site to check their own solutions. Solutions to the other exercises are available only to instructors (see “Instructor’s Note,” below, for information on how to get the solutions).

Many chapters have a tools section at the end of the chapter that provides information on software tools related to the topic of the chapter; some of these tools can be used for laboratory exercises. SQL DDL and sample data for the university database and other relations used in the exercises are available on the web site of the book and can be used for laboratory exercises.

Instructor’s Note

It is possible to design courses by using various subsets of the chapters. Some of the chapters can also be covered in an order different from their order in the book. We outline some of the possibilities here:

- Chapter 5 (Advanced SQL). This chapter can be skipped or deferred to later without loss of continuity. We expect most courses will cover at least Section 5.1.1 early, as JDBC is likely to be a useful tool in student projects.
- Chapter 6 (E-R Model). This chapter can be covered ahead of Chapter 3, Chapter 4, and Chapter 5 if you so desire, since Chapter 6 does not have any dependency on SQL. However, for courses with a programming emphasis, a richer variety of laboratory exercises is possible after studying SQL, and we recommend that SQL be covered before database design for such courses.
- Chapter 15 (Query Processing) and Chapter 16 (Query Optimization). These chapters can be omitted from an introductory course without affecting coverage of any other chapter.
- Part 7 (Transaction Management). Our coverage consists of an overview (Chapter 17) followed by chapters with details. You might choose to use Chapter 17 while omitting Chapter 18 and Chapter 19, if you defer these latter chapters to an advanced course.
- Part 8 (Parallel and Distributed Databases). Our coverage consists of an overview (Chapter 20), followed by chapters on the topics of storage, query processing, and transactions. You might choose to use Chapter 20 while omitting Chapter 21 through Chapter 23 if you defer these latter chapters to an advanced course.
- Part 11 (Online chapters). Chapter 27 (Formal-Relational Query Languages). This chapter can be covered immediately after Chapter 2, ahead of SQL. Alternatively, this chapter may be omitted from an introductory course. The five other online chapters (Advanced Relational Database Design, Object-Based Databases, XML, Information Retrieval, and PostgreSQL) can be used as self-study material or omitted from an introductory course.

Model course syllabi, based on the text, can be found on the web site of the book.

Web Site and Teaching Supplements

A web site for the book is available at the URL: db-book.com. The web site contains:

- Slides covering all the chapters of the book.
- Answers to the practice exercises.
- The six online chapters.
- Laboratory material, including SQL DDL and sample data for the university schema and other relations used in exercises, and instructions for setting up and using various database systems and tools.
- An up-to-date errata list.

The following additional material is available only to faculty:

- An instructor's manual containing solutions to all exercises in the book.
- A question bank containing extra exercises.

For more information about how to get a copy of the instructor's manual and the question bank, please send an email message to sem@mheducation.com. In the United States, you may call 800-338-3987. The McGraw-Hill web site for this book is www.mhhe.com/silberschatz.

Contacting Us

We have endeavored to eliminate typos, bugs, and the like from the text. But, as in new releases of software, bugs almost surely remain; an up-to-date errata list is accessible from the book's web site. We would appreciate it if you would notify us of any errors or omissions in the book that are not on the current list of errata.

We would be glad to receive suggestions on improvements to the book. We also welcome any contributions to the book web site that could be of use to other readers, such as programming exercises, project suggestions, online labs and tutorials, and teaching tips.

Email should be addressed to db-book-authors@cs.yale.edu. Any other correspondence should be sent to Avi Silberschatz, Department of Computer Science, Yale University, 51 Prospect Street, P.O. Box 208285, New Haven, CT 06520-8285 USA.

Acknowledgments

Many people have helped us with this seventh edition, as well as with the previous six editions from which it is derived, and we are indebted to all of them.

Seventh Edition

- Ioannis Alagiannis and Renata Borovica-Gajic for writing Chapter 32 on the PostgreSQL database, which is available online. The chapter is a complete rewrite of the PostgreSQL chapter in the 6th edition, which was authored by Anastasia Ailamaki, Sailesh Krishnamurthy, Spiros Papadimitriou, Bianca Schroeder, Karl Schnaitter, and Gavin Sherry.
- Judi Paige for her help in generating figures, presentation slides, and with handling the copy-editing material.
- Mark Wogahn for making sure that the software to produce the book, including LaTeX macros and fonts, worked properly.

- Sriram Srinivasan for discussions and feedback that have immensely benefited the chapters on parallel and distributed databases.
- N. L. Sarda for his insightful feedback on the sixth edition, and on some sections of the seventh edition.
- Bikash Chandra and Venkatesh Emani for their help with updates to the application development chapter, including creation of sample code.
- Students at IIT Bombay, particularly Ashish Mithole, for their feedback on draft versions of the chapters on parallel and distributed databases.
- Students at Yale, Lehigh, and IIT Bombay, for their comments on the sixth edition.
- Jeffrey Anthony, partner and CTO, Synaptic; and Lehigh students Corey Caplan (now co-founder, Leavitt Innovations); Gregory Cheng; Timothy LaRowe; and Aaron Rotem for comments and suggestions that have benefited the new blockchain chapter.

Previous Editions

- Hakan Jakobsson (Oracle), for writing the chapter on the Oracle database system in the sixth edition; Sriram Padmanabhan (IBM), for writing the chapter describing the IBM DB2 database system in the sixth edition; and Sameet Agarwal, José A. Blakeley, Thierry D’Hers, Gerald Hinson, Dirk Myers, Vaqar Pirzada, Bill Ramos, Balaji Rathakrishnan, Michael Rys, Florian Waas, and Michael Zwilling for writing the chapter describing the Microsoft SQL Server database system in the sixth edition; and in particular José Blakeley, who sadly is no longer amongst us, for coordinating and editing the chapter; and César Galindo-Legaria, Goetz Graefe, Kalen Delaney, and Thomas Casey for their contributions to the previous edition of the Microsoft SQL Server chapter. These chapters, however, are not part of the seventh edition.
- Anastasia Ailamaki, Sailesh Krishnamurthy, Spiros Papadimitriou, Bianca Schroeder, Karl Schnaitter, and Gavin Sherry for writing the chapter on PostgreSQL in the sixth edition.
- Daniel Abadi for reviewing the table of contents of the fifth edition and helping with the new organization.
- Steve Dolins, University of Florida; Rolando Fernanez, George Washington University; Frantisek Franek, McMaster University; Latifur Khan, University of Texas at Dallas; Sanjay Madria, Missouri University of Science and Technology; Aris Ouksel, University of Illinois; and Richard Snodgrass, University of Waterloo; who served as reviewers of the book and whose comments helped us greatly in formulating the sixth edition.

- Judi Paige for her help in generating figures and presentation slides.
- Mark Wogahn for making sure that the software to produce the book, including LaTeX macros and fonts, worked properly.
- N. L. Sarda for feedback that helped us improve several chapters. Vikram Pudi for motivating us to replace the earlier bank schema; and Shetal Shah for feedback on several chapters.
- Students at Yale, Lehigh, and IIT Bombay, for their comments on the fifth edition, as well as on preprints of the sixth edition.
- Chen Li and Sharad Mehrotra for providing material on JDBC and security for the fifth edition.
- Marilyn Turnamian and Nandprasad Joshi provided secretarial assistance for the fifth edition, and Marilyn also prepared an early draft of the cover design for the fifth edition.
- Lyn Dupré copyedited the third edition and Sara Strandtman edited the text of the third edition.
- Nilesh Dalvi, Sumit Sanghai, Gaurav Bhalotia, Arvind Hulgeri K. V. Raghavan, Prateek Kapadia, Sara Strandtman, Greg Speegle, and Dawn Bezviner helped to prepare the instructor's manual for earlier editions.
- The idea of using ships as part of the cover concept was originally suggested to us by Bruce Stephan.
- The following people offered suggestions and comments for the fifth and earlier editions of the book. R. B. Abhyankar, Hani Abu-Salem, Jamel R. Alsabbagh, Raj Ashar, Don Batory, Phil Bernhard, Christian Breimann, Gavin M. Bierman, Janek Bogucki, Haran Boral, Paul Bourgeois, Phil Bohannon, Robert Brazile, Yuri Breitbart, Ramzi Bualuan, Michael Carey, Soumen Chakrabarti, Tom Chappell, Zhengxin Chen, Y. C. Chin, Jan Chomicki, Laurens Damen, Prasanna Dhandapani, Qin Ding, Valentin Dinu, J. Edwards, Christos Faloutsos, Homma Farian, Alan Fekete, Frantisek Franek, Shashi Gadia, Hector Garcia-Molina, Goetz Graefe, Jim Gray, Le Gruenwald, Eitan M. Gurari, William Hankley, Bruce Hillyer, Ron Hitchens, Chad Hogg, Arvind Hulgeri, Yannis Ioannidis, Zheng Jiaping, Randy M. Kaplan, Graham J. L. Kemp, Rami Khouri, Hyoung-Joo Kim, Won Kim, Henry Korth (father of Henry F.), Carol Kroll, Hae Choon Lee, Sang-Won Lee, Irwin Levinstein, Mark Llewellyn, Gary Lindstrom, Ling Liu, Dave Maier, Keith Marzullo, Marty Maskarinec, Fletcher Mattox, Sharad Mehrotra, Jim Melton, Alberto Mendelzon, Ami Motro, Bhagirath Narahari, Yiu-Kai Dennis Ng, Thanh-Duy Nguyen, Anil Nigam, Cyril Orji, Meral Ozsoyoglu, D. B. Phatak, Juan Altmayer Pizzorno, Bruce Porter, Sunil Prabhakar, Jim Peterson, K. V. Raghavan, Nahid Rahman, Rajarshi Rakshit, Krithi Ramamritham, Mike Reiter, Greg Ric-

cardi, Odinaldo Rodriguez, Mark Roth, Marek Rusinkiewicz, Michael Rys, Sunita Sarawagi, N. L. Sarda, Patrick Schmid, Nikhil Sethi, S. Seshadri, Stewart Shen, Shashi Shekhar, Amit Sheth, Max Smolens, Nandit Soparkar, Greg Speegle, Jeff Storey, Dilys Thomas, Prem Thomas, Tim Wahls, Anita Whitehall, Christopher Wilson, Marianne Winslett, Weining Zhang, and Liu Zhenming.

Personal Notes

Sudarshan would like to acknowledge his wife, Sita, for her love, patience, and support, and children Madhur and Advaith for their love and joie de vivre. Hank would like to acknowledge his wife, Joan, and his children, Abby and Joe, for their love and understanding. Avi would like to acknowledge Valerie for her love, patience, and support during the revision of this book.

A. S.
H. F. K.
S. S.