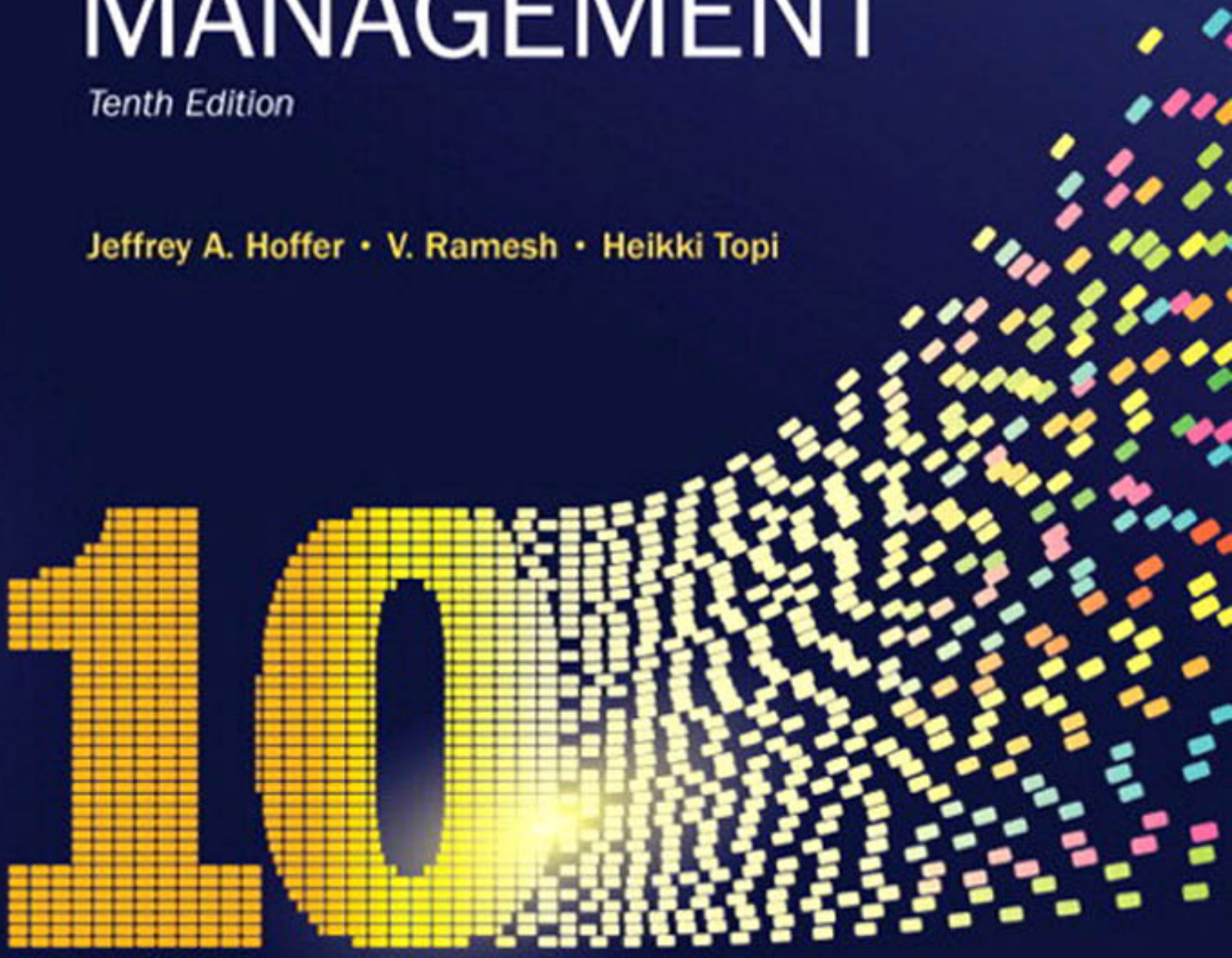
MODERN DATABASE MANAGEMENT

Tenth Edition

Jeffrey A. Hoffer • V. Ramesh • Heikki Topi



Editorial Director: Sally Yagan
Editor in Chief: Eric Svendsen
Executive Editor: Bob Horan
Editorial Project Manager: Kelly Loftus
Editorial Assistant: Jason Calcano
Director of Marketing: Patrice Lumumba Jones
Marketing Manager: Anne Fahlgren
Marketing Assistant: Melinda Jensen
Senior Managing Editor: Judy Leale
Project Manager: Becca Richter
Senior Operations Supervisor: Arnold Vila
Operations Specialist: Ilene Kahn
Senior Art Director: Jayne Conte
Cover Designer: Suzanne Behnke
Cover Art: Fotolia © vuifah
Manager, Visual Research: Karen Sanatar
Permissions Project Manager: Shannon Barbe
Media Project Manager, Editorial: Denise Vaughn
Media Project Manager, Production: Lisa Rinaldi
Supplements Editor: Kelly Loftus
Full-Service Project Management: PreMediaGlobal
Composition: PreMediaGlobal
Printer/Binder: Edwards Brothers
Cover Printer: Lehigh-Phoenix Color/Hagerstown
Text Font: Palatino

Credits and acknowledgments borrowed from other sources and reproduced, with permission, in this textbook appear on appropriate page within text.

Microsoft® and Windows® are registered trademarks of the Microsoft Corporation in the U.S.A. and other countries. Screen shots and icons reprinted with permission from the Microsoft Corporation. This book is not sponsored or endorsed by or affiliated with the Microsoft Corporation.

Copyright © 2011, 2009, 2007, 2005, 2002 Pearson Education, Inc., publishing as Prentice Hall, One Lake Street, Upper Saddle River, New Jersey 07458. All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use material from this work, please submit a written request to Pearson Education, Inc., Permissions Department, One Lake Street, Upper Saddle River, New Jersey 07458.

Many of the designations by manufacturers and seller to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed in initial caps or all caps.

Library of Congress Cataloging-in-Publication Data

Hoffer, Jeffrey A.
Modern database management / Jeffrey A. Hoffer, V. Ramesh, Heikki Topi. — 10th ed.
p. cm.
Includes index.
ISBN 0-13-608839-2 (alk. paper)
1. Database management. I. Ramesh, V. II. Topi, Heikki. III. Title.
QA76.9.D3M395 2011
005.74—dc22

2010017419

10 9 8 7 6 5 4 3 2 1

Prentice Hall
is an imprint of

PEARSON

www.pearsonhighered.com

ISBN 10: 0-13-608839-2
ISBN 13: 978-0-13-608839-4

Tenth Edition

MODERN DATABASE MANAGEMENT

Jeffrey A. Hoffer
University of Dayton

V. Ramesh
Indiana University

Heikki Topi
Bentley University

Prentice Hall
Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

To Patty, for her sacrifices, encouragement, and support for over 28 years of being a textbook author widow. To my students and colleagues, for being receptive and critical and for challenging me to be a better teacher.

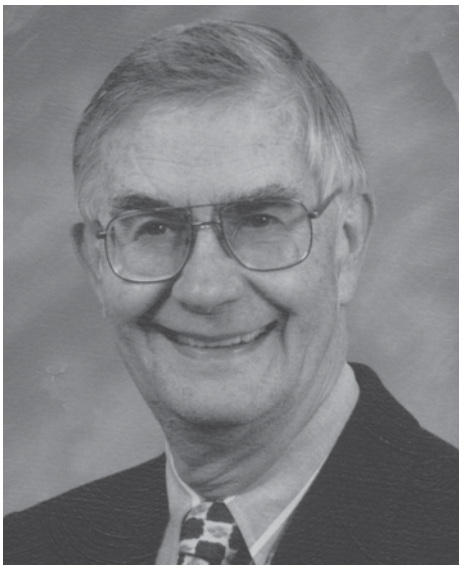
—J.A.H.

To Gayathri, for her sacrifices and patience these past 20 years. To my parents for letting me make the journey abroad, and to my cat, Raju, for being a part of our family for almost 20 years.

—V.R.

To Anne-Louise, for her loving support, encouragement, and patience. To Leila, whose laughter and joy of life continue to teach me about what is truly important. To my teachers, colleagues, and students, from whom I continue to learn every day.

—H.T.



Fred R. McFadden
1933–2009

Founding author of Modern Database Management, *Fred McFadden, passed away on August 9, 2009. Fred was a dedicated educator for 30 years in the College of Business at the University of Colorado, Colorado Springs. He received his bachelor's degree in Mechanical Engineering from Michigan State University, his MBA from the University of California, Los Angeles, and his PhD in Industrial Engineering from Stanford University. He began writing* Modern Database Management *in 1980 and was considered a leading information systems educator in database management, systems analysis, and decision support, all areas in which he was a scholarly author. Fred's work on the initial design of this textbook was pioneering, as few books existed then to present information technology to business students.*

Fred was an inspiration to his students and colleagues. An outstanding communicator with a strong sense of clarity and the needs of students, he was a mentor to his co-authors. Fred's first concern was always what was best for the students using the book, and he worked tirelessly to make passages succinct, readable, and motivating. He taught through examples and imaginatively told stories with graphics. He was skilled at blending the latest and best industry practices with leading research results into material accessible to all readers, whether undergraduate or graduate students. Fred was encouraging to his co-authors, always prepared to take on any writing assignment, yet never so prideful of his writing as to not accept comments with respect. Fred was actively involved in writing this text through the 8th edition, and he remained a confidant and guide after he ceased active writing.

Besides his professional contributions, Fred more than anything else was a caring, gentle, passionate person. Growing up on a farm in Michigan taught him to love the outdoors and to have a strong sense of caring for his neighbor, whom Fred saw as everyone.

The co-authors of Modern Database Management, *10th edition, are humbled to dedicate this edition to Fred R. McFadden, our friend and colleague.*

This page intentionally left blank

BRIEF CONTENTS

PART I The Context of Database Management 1

Chapter 1 The Database Environment and Development Process 2

PART II Database Analysis 55

Chapter 2 Modeling Data in the Organization 57
Chapter 3 The Enhanced E-R Model 113

PART III Database Design 153

Chapter 4 Logical Database Design and the Relational Model 155
Chapter 5 Physical Database Design and Performance 207

PART IV Implementation 241

Chapter 6 Introduction to SQL 243
Chapter 7 Advanced SQL 289
Chapter 8 Database Application Development 335
Chapter 9 Data Warehousing 375

PART V Advanced Database Topics 431

Chapter 10 Data Quality and Integration 433
Chapter 11 Data and Database Administration 461
Chapter 12 Overview: Distributed Databases 512
Chapter 13 Overview: Object-Oriented Data Modeling 516
Chapter 14 Overview: Using Relational Databases to Provide Object Persistence 525

Appendices

Appendix A: Data Modeling Tools and Notation 535
Appendix B: Advanced Normal Forms 545
Appendix C: Data Structures 551
Glossary of Acronyms 563
Glossary of Terms 565
Index 573

Available Online at www.pearsonhighered.com/hoffer

Chapter 12 Distributed Databases 12-1
Chapter 13 Object-Oriented Data Modeling 13-1
Chapter 14 Using Relational Databases to Provide Object Persistence 14-1

This page intentionally left blank

CONTENTS

Preface xxv

Part I The Context of Database Management 1

An Overview of Part One 1

Chapter 1 The Database Environment and Development Process 2

Learning Objectives 2
Data Matter! 2
Introduction 3
Basic Concepts and Definitions 5
- Data 5
- Data Versus Information 6
- Metadata 7

Traditional File Processing Systems 8
- File Processing Systems at Pine Valley Furniture Company 8
- Disadvantages of File Processing Systems 9
  - Program-Data Dependence 9
  - Duplication of Data 9
  - Limited Data Sharing 9
  - Lengthy Development Times 9
  - Excessive Program Maintenance 9

The Database Approach 10
- Data Models 10
  - Entities 10
  - Relationships 10
- Relational Databases 10
- Database Management Systems 11
- Advantages of the Database Approach 12
  - Program-Data Independence 13
  - Planned Data Redundancy 13
  - Improved Data Consistency 13
  - Improved Data Sharing 13
  - Increased Productivity of Application Development 13
  - Enforcement of Standards 13
  - Improved Data Quality 14
  - Improved Data Accessibility and Responsiveness 14
  - Reduced Program Maintenance 14
  - Improved Decision Support 14
- Cautions About Database Benefits 15
- Costs and Risks of the Database Approach 15
  - New, Specialized Personnel 15
  - Installation and Management Cost and Complexity 15
  - Conversion Costs 15
  - Need for Explicit Backup and Recovery 15
  - Organizational Conflict 16

Components of the Database Environment 16

The Range of Database Applications 17
Personal Databases 18
Two-Tier Client/Server Databases 18
Multitier Client/Server Databases 19
Enterprise Applications 20
Evolution of Database Systems 21
1960s 21
1970s 23
1980s 23
1990s 23
2000 and Beyond 23
The Database Development Process 24
Systems Development Life Cycle 25
Planning—Enterprise Modeling 26
Planning—Conceptual Data Modeling 26
Analysis—Conceptual Data Modeling 26
Design—Logical Database Design 26
Design—Physical Database Design and Definition 27
Implementation—Database Implementation 27
Maintenance—Database Maintenance 27
Alternative IS Development Approaches 28
Three-Schema Architecture for Database Development 29
Managing the People Involved in Database Development 31
Developing a Database Application for Pine Valley Furniture Company 31
Simplified Project Data Model Example 33
A Current Pine Valley Furniture Company Project Request 35
Project Planning 36
Analyzing Database Requirements 37
Designing the Database 38
Using the Database 41
Administering the Database 42
Summary 42 • Key Terms 43 • Review Questions 44 • Problems and Exercises 45 • Field Exercises 46 • References 47 • Further Reading 47 • Web Resources 48
▶ **CASE: Mountain View Community Hospital 49**

Part II Database Analysis 55
An Overview of Part Two 55

Chapter 2 Modeling Data in the Organization 57
Learning Objectives 57
Introduction 57
The E-R Model: An Overview 59
Sample E-R Diagram 59
E-R Model Notation 61
Modeling the Rules of the Organization 62
Overview of Business Rules 63
The Business Rules Paradigm 63

Scope of Business Rules 64
Good Business Rules 64
Gathering Business Rules 64
Data Names and Definitions 65
Data Names 65
Data Definitions 66
Good Data Definitions 66
Modeling Entities and Attributes 68
Entities 68
Entity Type Versus Entity Instance 68
Entity Type Versus System Input, Output, or User 69
Strong Versus Weak Entity Types 69
Naming and Defining Entity Types 70
Attributes 72
Required Versus Optional Attributes 72
Simple Versus Composite Attributes 73
Single-Valued Versus Multivalued Attributes 73
Stored Versus Derived Attributes 74
Identifier Attribute 74
Naming and Defining Attributes 76
Modeling Relationships 77
Basic Concepts and Definitions in Relationships 78
Attributes on Relationships 79
Associative Entities 80
Degree of a Relationship 81
Unary Relationship 81
Binary Relationship 82
Ternary Relationship 83
Attributes or Entity? 84
Cardinality Constraints 86
Minimum Cardinality 87
Maximum Cardinality 87
Some Examples of Relationships and Their Cardinalities 87
A Ternary Relationship 88
Modeling Time-Dependent Data 89
Modeling Multiple Relationships Between Entity Types 92
Naming and Defining Relationships 93
E-R Modeling Example: Pine Valley Furniture Company 95
Database Processing at Pine Valley Furniture 97
Showing Product Information 97
Showing Product Line Information 98
Showing Customer Order Status 98
Showing Product Sales 100
Summary 100 • Key Terms 101 • Review Questions 101 • Problems and Exercises 102 • Field Exercises 108 • References 109 • Further Reading 109 • Web Resources 110
▶ **CASE: Mountain View Community Hospital 111**

Chapter 3 The Enhanced E-R Model 113
Learning Objectives 113
Introduction 113

Representing Supertypes and Subtypes 114
Basic Concepts and Notation 115
An Example of a Supertype/Subtype Relationship 116
Attribute Inheritance 117
When to Use Supertype/Subtype Relationships 117
Representing Specialization and Generalization 118
Generalization 118
Specialization 119
Combining Specialization and Generalization 120
Specifying Constraints in Supertype/Subtype Relationships 121
Specifying Completeness Constraints 121
Total Specialization Rule 121
Partial Specialization Rule 122
Specifying Disjointness Constraints 122
Disjoint Rule 122
Overlap Rule 122
Defining Subtype Discriminators 123
Disjoint Subtypes 123
Overlapping Subtypes 124
Defining Supertype/Subtype Hierarchies 125
An Example of a Supertype/Subtype Hierarchy 125
Summary of Supertype/Subtype Hierarchies 126
EER Modeling Example: Pine Valley Furniture Company 127
Entity Clustering 130
Packaged Data Models 133
A Revised Data Modeling Process with Packaged Data Models 135
Packaged Data Model Examples 137
Summary 142 • Key Terms 143 • Review Questions 143 • Problems and Exercises 144 • Field Exercises 147 • References 147 • Further Reading 147 • Web Resources 148
▶ **CASE: Case: Mountain View Community Hospital 149**

Part III Database Design 153

An Overview of Part Three 153

Chapter 4 Logical Database Design and the Relational Model 155

Learning Objectives 155
Introduction 155
The Relational Data Model 156
Basic Definitions 156
Relational Data Structure 157
Relational Keys 157
Properties of Relations 158
Removing Multivalued Attributes from Tables 158
Sample Database 158
Integrity Constraints 160
Domain Constraints 160
Entity Integrity 160
Referential Integrity 162

Creating Relational Tables 163
Well-Structured Relations 164
Transforming EER Diagrams into Relations 165
Step 1: Map Regular Entities 166
Composite Attributes 166
Multivalued Attributes 167
Step 2: Map Weak Entities 167
When to Create a Surrogate Key 169
Step 3: Map Binary Relationships 169
Map Binary One-to-Many Relationships 169
Map Binary Many-to-Many Relationships 170
Map Binary One-to-One Relationships 170
Step 4: Map Associative Entities 171
Identifier Not Assigned 171
Identifier Assigned 172
Step 5: Map Unary Relationships 173
Unary One-to-Many Relationships 173
Unary Many-to-Many Relationships 174
Step 6: Map Ternary (and n-ary) Relationships 175
Step 7: Map Supertype/Subtype Relationships 176
Summary of EER-to-Relational Transformations 178
Introduction to Normalization 178
Steps in Normalization 179
Functional Dependencies and Keys 179
Determinants 181
Candidate Keys 181
Normalization Example: Pine Valley Furniture Company 182
Step 0: Represent the View in Tabular Form 182
Step 1: Convert to First Normal Form 183
Remove Repeating Groups 183
Select the Primary Key 184
Anomalies in 1NF 184
Step 2: Convert to Second Normal Form 185
Step 3: Convert to Third Normal Form 186
Removing Transitive Dependencies 186
Determinants and Normalization 187
Step 4: Further Normalization 188
Merging Relations 188
An Example 188
View Integration Problems 189
Synonyms 189
Homonyms 189
Transitive Dependencies 190
Supertype/Subtype Relationships 190
A Final Step for Defining Relational Keys 190
Summary 192 • Key Terms 194 • Review Questions 194 • Problems and Exercises 195 • Field Exercises 202 • References 202 • Further Reading 202 • Web Resources 202
▶ **CASE: Case: Mountain View Community Hospital 203**

Chapter 5 Physical Database Design and Performance 207
Learning Objectives 207
Introduction 207
The Physical Database Design Process 208
Physical Database Design as a Basis for Regulatory Compliance 209
Data Volume and Usage Analysis 210
Designing Fields 211
Choosing Data Types 212
Coding Techniques 212
Handling Missing Data 214
Denormalizing and Partitioning Data 214
Denormalization 214
Opportunities for and Types of Denormalization 215
Denormalize with Caution 217
Partitioning 218
Designing Physical Database Files 220
File Organizations 221
Sequential File Organizations 222
Indexed File Organizations 222
Hashed File Organizations 225
Clustering Files 227
Designing Controls for Files 228
Using and Selecting Indexes 229
Creating a Unique Key Index 229
Creating a Secondary (Nonunique) Key Index 229
When to Use Indexes 230
Designing a Database for Optimal Query Performance 231
Parallel Query Processing 231
Overriding Automatic Query Optimization 232
Summary 233 • Key Terms 233 • Review Questions 234 • Problems and Exercises 234 • Field Exercises 237 • References 237 • Further Reading 237 • Web Resources 237

▶ **CASE: Mountain View Community Hospital 238**

Part IV Implementation 241
An Overview of Part Four 241

Chapter 6 Introduction to SQL 243
Learning Objectives 243
Introduction 243
Origins of the SQL Standard 245
The SQL Environment 246
Defining A Database in SQL 251
Generating SQL Database Definitions 252
Creating Tables 252
Creating Data Integrity Controls 255
Changing Table Definitions 256
Removing Tables 257

Inserting, Updating, and Deleting Data 257
Batch Input 258
Deleting Database Contents 259
Updating Database Contents 259
Internal Schema Definition in RDBMSs 260
Creating Indexes 260
Processing Single Tables 261
Clauses of the SELECT Statement 261
Using Expressions 263
Using Functions 264
Using Wildcards 267
Using Comparison Operators 267
Using Null Values 268
Using Boolean Operators 268
Using Ranges for Qualification 271
Using Distinct Values 271
Using IN and NOT IN with Lists 273
Sorting Results: The ORDER BY Clause 274
Categorizing Results: The GROUP BY Clause 275
Qualifying Results by Categories: The HAVING Clause 276
Using and Defining Views 278
MATERIALIZED VIEWS 281
Summary 281 • Key Terms 282 • Review Questions 282 • Problems and Exercises 283 • Field Exercises 286 • References 286 • Further Reading 287 • Web Resources 287
▶ **CASE: Mountain View Community Hospital 288**

Chapter 7 Advanced SQL 289
Learning Objectives 289
Introduction 289
Processing Multiple Tables 290
Equi-join 291
Natural Join 292
Outer Join 293
Union Join 295
Sample Join Involving Four Tables 295
Self-Join 297
Subqueries 298
Correlated Subqueries 303
Using Derived Tables 304
Combining Queries 305
Conditional Expressions 307
More Complicated SQL Queries 308
Tips for Developing Queries 310
Guidelines for Better Query Design 311
Ensuring Transaction Integrity 313

Data Dictionary Facilities 314
SQL:200n Enhancements and Extensions to SQL 317
Analytical and OLAP Functions 317
New Data Types 318
Other Enhancements 319
Programming Extensions 319
Triggers and Routines 320
Triggers 321
Routines 323
Embedded SQL and Dynamic SQL 326
Summary 328 • Key Terms 329 • Review Questions 329 • Problems and Exercises 330 • Field Exercises 333 • References 333 • Further Reading 333 • Web Resources 333

▶ **CASE: Mountain View Community Hospital 334**

Chapter 8 Database Application Development 335
Learning Objectives 335
Location, Location, Location! 335
Introduction 336
Client/Server Architectures 336
Partitioning an Application 337
Databases in a Two-Tier Architecture 339
A VB.NET Example 341
A Java Example 343
Three-Tier Architectures 344
Web Application Components 346
Languages for Creating Web Pages 348
Databases in Three-Tier Applications 348
A JSP Web Application 349
A PHP Example 353
An ASP.NET Example 353
Key Considerations in Three-Tier Applications 355
Stored Procedures 356
Transactions 357
Database Connections 359
Key Benefits of Three-Tier Applications 359
Extensible Markup Language (XML) 360
Storing XML Documents 362
Retrieving XML Documents 362
Displaying XML Data 365
XML and Web Services 365
Summary 369 • Key Terms 369 • Review Questions 370 • Problems and Exercises 370 • Field Exercises 371 • References 371 • Further Reading 371 • Web Resources 371

▶ **CASE: Mountain View Community Hospital 373**

Chapter 9 Data Warehousing 375
Learning Objectives 375
Introduction 375
Basic Concepts of Data Warehousing 377
A Brief History of Data Warehousing 378
The Need for Data Warehousing 378
Need for a Company-wide View 378
Need to Separate Operational and Informational Systems 380
Data Warehousing Success 381
Data Warehouse Architectures 382
Independent Data Mart Data Warehousing Environment 382
Dependent Data Mart and Operational Data Store Architecture: A Three-Level Approach 384
Logical Data Mart and Real-Time Data Warehouse Architecture 386
Three-Layer Data Architecture 389
Role of the Enterprise Data Model 390
Role of Metadata 390
Some Characteristics of Data Warehouse Data 390
Status Versus Event Data 390
Transient Versus Periodic Data 391
An Example of Transient and Periodic Data 391
Transient Data 391
Periodic Data 393
Other Data Warehouse Changes 393
The Derived Data Layer 394
Characteristics of Derived Data 394
The Star Schema 395
Fact Tables and Dimension Tables 395
Example Star Schema 396
Surrogate Key 398
Grain of the Fact Table 398
Duration of the Database 399
Size of the Fact Table 399
Modeling Date and Time 400
Variations of the Star Schema 401
Multiple Fact Tables 401
Factless Fact Tables 402
Normalizing Dimension Tables 403
Multivalued Dimensions 403
Hierarchies 404
Slowly Changing Dimensions 406
Determining Dimensions and Facts 408
Column Databases: A New Alternative for Data Warehouses 410
The User Interface 411
Role of Metadata 412
SQL OLAP Querying 412
Online Analytical Processing (OLAP) Tools 414
Slicing a Cube 415
Drill-Down 415
Summarizing More Than Three Dimensions 415

Data Visualization 415
Business Performance Management and Dashboards 417
Data-Mining Tools 418
Data-Mining Techniques 418
Data-Mining Applications 419
Summary 420 • Key Terms 420 • Review Questions 421 • Problems and Exercises 421 • Field Exercises 425 • References 426 • Further Reading 426 • Web Resources 426
▶ **CASE: Mountain View Community Hospital 428**

Part V Advanced Database Topics 431
An Overview of Part Five 431

Chapter 10 Data Quality and Integration 433
Learning Objectives 433
Introduction 433
Data Governance 434
Managing Data Quality 435
Characteristics of Quality Data 436
External Data Sources 437
Redundant Data Storage and Inconsistent Metadata 438
Data Entry Problems 438
Lack of Organizational Commitment 438
Data Quality Improvement 438
Get the Business Buy-in 438
Conduct a Data Quality Audit 439
Establish a Data Stewardship Program 440
Improve Data Capture Processes 441
Apply Modern Data Management Principles and Technology 441
Apply TQM Principles and Practices 441
Summary of Data Quality 442
Master Data Management 442
Data Integration: An Overview 443
General Approaches to Data Integration 444
Data Federation 444
Data Propagation 444
Data Integration for Data Warehousing: The Reconciled Data Layer 445
Characteristics of Data After ETL 446
The ETL Process 446
Mapping and Metadata Management 447
Extract 447
Cleanse 448
Load and Index 450
Data Transformation 452
Data Transformation Functions 452
Record-Level Functions 452
Field-Level Functions 453
Summary 455 • Key Terms 455 • Review Questions 456 • Problems and Exercises 456 • Field Exercises 457 • References 457 • Further Reading 458 • Web Resources 458
▶ **CASE: Mountain View Community Hospital 459**

Chapter 11 Data and Database Administration 461
Learning Objectives 461
Introduction 462
The Roles of Data and Database Administrators 463
Traditional Data Administration 463
Traditional Database Administration 465
Trends in Database Administration 466
Data Warehouse Administration 468
Summary of Evolving Data Administration Roles 469
The Open Source Movement and Database Management 469
Managing Data Security 471
Threats to Data Security 471
Establishing Client/Server Security 473
Server Security 473
Network Security 473
Application Security Issues in Three-Tier Client/Server Environments 473
Data Privacy 475
Database Software Data Security Features 476
Views 476
Integrity Controls 477
Authorization Rules 479
User-Defined Procedures 480
Encryption 480
Authentication Schemes 481
Passwords 481
Strong Authentication 482
Sarbanes-Oxley (SOX) and Databases 482
IT Change Management 483
Logical Access to Data 483
Personnel Controls 483
Physical Access Controls 483
IT Operations 484
Database Backup and Recovery 484
Basic Recovery Facilities 484
Backup Facilities 484
Journalizing Facilities 485
Checkpoint Facility 485
Recovery Manager 486
Recovery and Restart Procedures 486
Disk Mirroring 486
Restore/Rerun 487
Maintaining Transaction Integrity 487
Backward Recovery 488
Forward Recovery 489
Types of Database Failure 490
Aborted Transactions 490
Incorrect Data 490
System Failure 491
Database Destruction 491
Disaster Recovery 491

Controlling Concurrent Access 492
The Problem of Lost Updates 492
Serializability 492
Locking Mechanisms 493
Locking Level 493
Types of Locks 494
Deadlock 495
Managing Deadlock 495
Versioning 496
Data Dictionaries and Repositories 498
Data Dictionary 498
Repositories 498
Overview of Tuning the Database for Performance 500
Installation of the DBMS 500
Memory and Storage Space Usage 501
Input/Output (I/O) Contention 501
CPU Usage 502
Application Tuning 502
Data Availability 503
Costs of Downtime 503
Measures to Ensure Availability 504
Hardware Failures 504
Loss or Corruption of Data 504
Human Error 504
Maintenance Downtime 504
Network-Related Problems 505
Summary 505 • Key Terms 505 • Review Questions 506 • Problems and Exercises 507 • Field Exercises 509 • References 509 • Further Reading 510 • Web Resources 510

▶ **CASE: Mountain View Community Hospital 511**

Chapter 12 Overview: Distributed Databases 512
Learning Objectives 512
Overview 512
Objectives and Trade-offs 513
Options for Distributing a Database 513
Distributed DBMS 514
Query Optimization 514
Chapter Review 515 • References 515 • Further Reading 515 • Web Resources 515

Chapter 13 Overview: Object-Oriented Data Modeling 516
Learning Objectives 516
Overview 516
Unified Modeling Language 517
Object-Oriented Data Modeling 517
Representing Aggregation 523
Chapter Review 523 • References 523 • Further Reading 524 • Web Resources 524

Chapter 14 Overview: Using Relational Databases to Provide Object Persistence 525

Learning Objectives 525

Overview 525

Providing Persistence for Objects Using Relational Databases 526

Call-Level Application Programming Interfaces 527

SQL Query Mapping Frameworks 527

Object-Relational Mapping Frameworks 527

Proprietary Approaches 527

Selecting the Right Approach 528

Object-Relational Mapping Example 529

Mapping Files 529

Responsibilities of Object-Relational Mapping Frameworks 532

Summary 533 • Chapter Review 534 • References 534 • Further Reading 534 • Web Resources 534

Appendix A Data Modeling Tools and Notation 535

Comparing E-R Modeling Conventions 535

Visio Professional 2003 Notation 535

Entities 539

Relationships 539

CA ERwin Data Modeler r7.3 Notation 539

Entities 539

Relationships 539

Sybase PowerDesigner 15 Notation 541

Entities 542

Relationships 542

Oracle Designer Notation 542

Entities 542

Relationships 542

Comparison of Tool Interfaces and E-R Diagrams 542

Appendix B Advanced Normal Forms 545

Boyce-Codd Normal Form 545

Anomalies in Student Advisor 545

Definition of Boyce-Codd Normal Form (BCNF) 546

Converting a Relation to BCNF 546

Fourth Normal Form 547

Multivalued Dependencies 549

Higher Normal Forms 549

Key Terms 550 • References 550 • Web Resources 550

Appendix C Data Structures 551

Pointers 551

Data Structure Building Blocks 552

Linear Data Structures 554

Stacks 555

Queues 555

Sorted Lists 556
Multilists 558
Hazards of Chain Structures 558
Trees 559
Balanced Trees 559
Reference 562

Glossary of Acronyms 563
Glossary 565
Index 573

ONLINE CHAPTERS

Chapter 12 Distributed Databases 12-1
Learning Objectives 12-1
Introduction 12-1
Objectives and Trade-offs 12-4
Options for Distributing a Database 12-6
Data Replication 12-6
Snapshot Replication 12-7
Near-Real-Time Replication 12-8
Pull Replication 12-8
Database Integrity with Replication 12-8
When to Use Replication 12-8
Horizontal Partitioning 12-9
Vertical Partitioning 12-10
Combinations of Operations 12-11
Selecting the Right Data Distribution Strategy 12-12
Distributed DBMS 12-13
Location Transparency 12-15
Replication Transparency 12-16
Failure Transparency 12-17
Commit Protocol 12-17
Concurrency Transparency 12-18
Time-stamping 12-19
Query Optimization 12-19
Evolution of Distributed DBMSs 12-21
Remote Unit of Work 12-22
Distributed Unit of Work 12-22
Distributed Request 12-23
Distributed DBMS Products 12-23
Summary 12-24 • Key Terms 12-25 • Review Questions 12-25 • Problems and Exercises 12-26 • Field Exercises 12-27 • References 12-28 • Further Reading 12-28 • Web Resources 12-28

Chapter 13 Object-Oriented Data Modeling 13-1
Learning Objectives 13-1
Introduction 13-1
Unified Modeling Language 13-3
Object-Oriented Data Modeling 13-4
Representing Objects and Classes 13-4
Types of Operations 13-6
Representing Associations 13-7
Representing Association Classes 13-10
Representing Derived Attributes, Derived Associations, and Derived Roles 13-12
Representing Generalization 13-12
Interpreting Inheritance and Overriding 13-17

Representing Multiple Inheritance 13-18
Representing Aggregation 13-19
Business Rules 13-22
Object Modeling Example: Pine ValleyFurniture Company 13-23
Summary 13-25 • Key Terms 13-26 • Review Questions 13-26 • Problems and Exercises 13-29 • Field Exercises 13-35 • References 13-35 • Further Reading 13-36 • Web Resources 13-36

Chapter 14 Using Relational Databases to Provide Object Persistence 14-1
Learning Objectives 14-1
Introduction 14-1
Object-Relational Impedance Mismatch 14-3
Providing Persistence for Objects Using Relational Databases 14-6
Common Approaches 14-6
Call-Level Application Programming Interfaces 14-6
SQL Query Mapping Frameworks 14-7
Object-Relational Mapping Frameworks 14-7
Proprietary Approaches 14-7
Selecting the Right Approach 14-8
Call-Level APIs 14-8
SQL Query Mapping Frameworks 14-9
ORM Frameworks 14-9
Object-Relational Mapping Example Using Hibernate 14-10
Foundation 14-10
Mapping Files 14-11
Hibernate Configuration 14-15
Mapping Object-Oriented Structures to a Relational Database 14-16
Class 14-16
Inheritance: Superclass–Subclass 14-17
One-to-One Association 14-17
Many-to-One and One-to-Many Associations 14-17
Aggregation and Composition 14-19
Many-to-Many Associations 14-19
Responsibilities of Object-Relational Mapping Frameworks 14-20
HQL 14-21
Summary 14-25 • Key Terms 14-25 • Review Questions 14-26 • Problems and Exercises 14-26 • Field Exercises 14-27 • References 14-27 • Further Reading 14-27 • Web Resources 14-27

PREFACE

This text is designed to be used with an introductory course in database management. Such a course is usually required as part of an information systems curriculum in business schools, computer technology programs, and applied computer science departments. The Association for Information Systems (AIS), the Association for Computing Machinery (ACM), and the International Federation of Information Processing Societies (IFIPS) curriculum guidelines (e.g., IS 2010) all outline this type of database management course. Previous editions of this text have been used successfully for more than 27 years at both the undergraduate and graduate levels, as well as in management and professional development programs.

WHAT'S NEW IN THIS EDITION?

This 10th edition of *Modern Database Management* updates and expands materials in areas undergoing rapid change due to improved managerial practices, database design tools and methodologies, and database technology. Later we detail changes to each chapter. The themes of this 10th edition reflect the major trends in the information systems field and the skills required of modern information systems graduates:

- Data quality and database processing accuracy, which are extremely important with the national and international regulations such as the Sarbanes-Oxley Act, Basel II, COSI, and HIPAA that now require organizations to comply with standards for reporting accurate financial data and ensuring data privacy. Material on data quality and master data management has been updated with a stronger coverage of the people, process, and technology aspects and internationally accepted best practices for information systems development and management (specifically, ITIL).
- Integration of data from multiple internal and external databases and data sources, which is now common for building data warehouses and other types of enterprise systems, and dealing with the rapid organizational changes in information systems brought on by corporate reorganizations, mergers, and acquisitions. These first two bullets are implemented with the revised Chapter 10 on data quality and integration, which updates and improves the focus of the material and introduces the latest principles in these areas.
- Demonstrating knowledge of how to use databases in the context of developing database applications in two and three-tier client/server environments. In this 10th edition (in Chapters 8 and 14), we provide examples of how to connect to databases from popular programming languages such as Java and VB.NET as well Web development languages such as Java Server Pages (JSP), ASP.NET, and PHP. Coverage of XML has also been revised to emphasize the role of XML in data storage and retrieval.
- Linking object-oriented information systems development environments (such as Java Technology and Microsoft .NET) with mainstream technology for maintaining organizational data—relational databases—and in the process dealing with significant paradigm differences between object-oriented and relational frameworks. This major change that was introduced for the ninth edition and has been updated for the 10th edition reflects what is a rapidly changing environment for database processing.

Also, we are very excited to now provide on the student Companion Web site several new, custom-developed short videos that address key concepts and skills from different sections of the book. These videos, produced using Camtasia by the textbook authors, help students to learn difficult material by using both the printed text and a mini lecture or tutorial. Videos have been developed to support Chapters 1 (introduction to database), 2 and 3 (conceptual data modeling), 4 (normalization), and 6 and 7 (SQL).

More will be produced with future editions. Look for special icons on the opening page of these chapters to call attention to these videos, and go to **www.pearsonhighered.com/hoffer** to find these videos.

Specific improvements to the textbook have been made in the following areas:

- Arranged the Problems and Exercises into roughly increasing order of difficulty to make it easier for instructors and students to select problems and exercises for practice and assignments.
- Applied standard data naming conventions throughout the book to make it easier for students to distinguish data elements from conceptual to physical forms.
- Clarified system requirements through systems modeling and design and outlined a process to use the increasingly popular industry and business function commercial data models to speed up the systems development process. The new material focuses on changes to the database development process when an organization uses packaged data models. Students are now better prepared to understand why these data models are important and how to read and work with (tailor) them.
- Expanded coverage of SQL, with a few more frequently used components of the language. We have also created new figures to graphically depict the set processing logic of SQL queries, which gives students, especially visual learners, new tools to use when writing queries.
- Included new screen captures to reflect the latest database technologies and an updated Web Resources section in each chapter that lists Web sites that can provide the student with information on the latest database trends and expanded background details on important topics covered in the text.
- Reduced the length of the printed book, which we began doing with the eighth edition. The reduced length is more consistent with what our reviewers say can be covered in a database course today, given the need for depth of coverage in the most important topics. Specifically, for the 10th edition, we combined the first two chapters from the ninth edition into one, so that students can more quickly cover/review background topics and then dig into the material central to database management. We have also combined the two chapters from the ninth edition on client/server and Internet databases into one chapter addressing database issues in a multitier computing environment. We continue to update the chapters on distributed databases, the object-oriented data model, and using relational databases to provide object persistence, including an overview in the printed textbook and full versions on the textbook's Web site. Care has been given to the layout of figures and tables to also reduce the length of the book, while adding some new figures and figure elements to better link the text narrative with the figures. The reduced length should encourage more students to purchase and read the text, without any loss of coverage and learning. The book is also now available through CourseSmart, an innovative e-book delivery system.

MODERN DATABASE MANAGEMENT: A RETRO AND FUTURE PERSPECTIVE

This 10th edition is a humbling milestone. We are extremely grateful for the support of adopters, reviewers, students, colleagues, editors, and publisher staff who have been with us for some or, in a few cases, all of the past 27 years. Database technology has "grown up" over these years, from a resource for only the most sophisticated organizations to being a mainstay of almost any computing environment. Some topics, such as relational databases, have been a central part of the text from the beginning; other topics, such as data warehousing, business intelligence, object-oriented databases, and databases on the Internet, are newer topics. Database management used to be able to be explained in 531 pages that were about 80 percent the size of current pages, and now it takes 624 larger pages (really, we aren't just wordier). One of the original authors of this text is still a co-author, while a newer generation of database academic

experts now contributes to these pages with zest and creativity. The original book authors were educated in fields other than business information systems, whereas today our newer authors are experienced and educated in this rich field central to the success of modern organizations.

As a book that we believe has succeeded in leading the database management textbook market, this book is positioned to continue (in some printed or electronic form) for at least another 27 years. Writing this book has been and remains an awesome responsibility. We authors realize that the course that this text supports will be the foundation for student careers with databases. Over the years, we've seen students reading our book on airplanes while traveling on business, and, believe it or not, reading it on a Florida beach during spring break. The authors remain committed to presenting material with sound pedagogy, including topics (both easy and difficult, traditional and emerging) that are critical for the practical success of database professionals, and being informed by research that reveals what will be the "next big thing" in database management. It is in this spirit that we celebrate our milestone edition, and lay the foundation for many more editions to come.

FOR THOSE NEW TO MODERN DATABASE MANAGEMENT

Modern Database Management has been a leading text since its first edition in 1983. In spite of this market leadership position, some instructors have used other good database management texts. Why might you want to switch at this time? There are several good reasons to switch to *Modern Database Management*, including:

- One of our goals, in every edition, has been to lead other books in coverage of the latest principles, concepts, and technologies. See what we have added for the 10th edition in "What's New in This Edition." In the past, we have led in coverage of object-oriented data modeling and UML, Internet databases, data warehousing, and the use of CASE tools in support of data modeling. For the 10th edition, we are taking the lead on database development for Internet-based applications, data quality and integration, the linking of object-oriented development environments with relational databases, and the increasingly important role of packaged database model as a component of agile, rapid development of information systems. We also have for the first time Camtasia-produced tutorial videos to accompany the book, with more to come for future editions.
- While remaining current, this text focuses on what leading practitioners say is most important for database developers. We work with many practitioners, including the professionals of the Data Management Association (DAMA) and The Data Warehousing Institute (TDWI), leading consultants, technology leaders, and authors of articles in the most widely read professional publications. We draw on these experts to ensure that what the book includes is important and covers not only important entry-level knowledge and skills, but also those fundamentals and mindsets that lead to long-term career success.
- In this highly successful book in its 10th edition, material is presented in a way that has been viewed as very accessible to students. Our methods have been refined through continuous market feedback for over 27 years, as well as through our own teaching. Overall, the pedagogy of the book is sound. We use many illustrations that help to make important concepts and techniques clear. We use the most modern notations. The organization of the book is flexible, so you can use chapters in whatever sequence makes sense for your students. We supplement the book with data sets to facilitate hands-on, practical learning, and with new media resources to make some of the more challenging topics more engaging.
- You may have particular interest in introducing SQL early in your course. Our text can accommodate this. First, we cover SQL in depth, devoting two full chapters to this core technology of the database field. Second, we include many SQL examples in early chapters. Third, many instructors have successfully used the two SQL chapters early in their course. Although logically appearing in the life cycle of systems development as Chapters 6 and 7, part of the implementation section of the text, many instructors have used these chapters immediately after

Chapter 1 or in parallel with other early chapters. Finally, we use SQL throughout the book, for example, to illustrate Web application connections to relational databases in Chapter 8, online analytical processing in Chapter 9, and accessing relational databases from object-oriented development environments in Chapter 14.

- We have the latest in supplements and Web site support for the text. See the supplement package for details on all the resources available to you and your students.
- This text is written to be part of a modern information systems curriculum with a strong business systems development focus. Topics are included and addressed so as to reinforce principles from other typical courses, such as systems analysis and design, networking, Web site design and development, MIS principles, and computer programming. Emphasis is on the development of the database component of modern information systems and on the management of the data resource. Thus, the text is practical, supports projects and other hands-on class activities, and encourages linking database concepts to concepts being learned throughout the curriculum the student is taking.

SUMMARY OF ENHANCEMENTS TO EACH CHAPTER

The following sections present a chapter-by-chapter description of the major changes in this edition. Each chapter description presents a statement of the purpose of that chapter, followed by a description of the changes and revisions that have been made for the 10th edition. Each paragraph concludes with a description of the strengths that have been retained from prior editions.

Part I: The Context of Database Management

CHAPTER 1: THE DATABASE ENVIRONMENT AND DEVELOPMENT PROCESS This chapter discusses the role of databases in organizations and previews the major topics in the remainder of the text. This chapter has undergone extensive reorganization for the 10th edition because it is a consolidation of two previous chapters, allowing students to more quickly cover material that previews the rest of the book. After presenting a brief introduction to the basic terminology associated with storing and retrieving data, the chapter presents a well organized comparison of traditional file-processing systems and modern database technology. The chapter then introduces the core components of a database environment and the range of database applications that are currently in use within organizations—personal, two-tier, multitier, and enterprise applications. The explanation of enterprise databases includes databases that are part of enterprise resource planning systems and data warehouses. A brief history of the evolution of database technology, from pre-database files to modern object-relational technologies, is also presented. The chapter then goes on to explain the process of database development in the context of structured life cycle, prototyping, and agile methodologies. The presentation remains consistent with the companion systems analysis text by Hoffer, George, and Valacich. The chapter also discusses important issues in database development, including management of the diverse group of people involved in database development and frameworks for understanding database architectures and technologies (e.g., the three-schema architecture). Reviewers frequently note the compatibility of this chapter with what students learn in systems analysis and design classes.

Part II: Database Analysis

CHAPTER 2: MODELING DATA IN THE ORGANIZATION This chapter presents a thorough introduction to conceptual data modeling with the entity-relationship (E-R) model. The chapter title emphasizes the reason for the entity-relationship model: to unambiguously document the rules of the business that influence database design. Specific subsections explain in detail how to name and define elements of a data model, which are essential

in developing an unambiguous E-R diagram. In the 10th edition, we have provided some new problems and exercises, improved an example from Pine Valley Furniture to show the relationship of conceptual database design with implementation, and provided more annotations in figures to better highlight key elements and better link text to figures. The chapter continues to proceed from simple to more complex examples, and it concludes with a comprehensive E-R diagram for the Pine Valley Furniture Company.

CHAPTER 3: THE ENHANCED E-R MODEL This chapter presents a discussion of several advanced E-R data model constructs, primarily supertype/subtype relationships. A major change in this chapter is the elimination of the section on business rules, which many adopters and reviewers said they did not have time to cover in class. The most significant addition is a more thorough description of how to conduct a data modeling project when using a packaged data model; this new material better prepares students for working with commercial off-the-shelf (COTS) software and purchased data models, which support major efforts to implement patterns and reusability for application deployment in organizations. As in Chapter 2, figures have been improved with more annotations to clarify important data modeling structures. The chapter continues to present a thorough coverage of supertype/subtype relationships and includes a comprehensive example of an extended E-R data model for the Pine Valley Furniture Company.

Part III: Database Design

CHAPTER 4: LOGICAL DATABASE DESIGN AND THE RELATIONAL MODEL This chapter describes the process of converting a conceptual data model to the relational data model, as well as how to merge new relations into an existing normalized database. It provides a conceptually sound and practically relevant introduction to normalization, emphasizing the importance of the use of functional dependencies and determinants as the basis for normalization. Concepts of normalization and normal forms are extended in Appendix B. The chapter features a discussion of the characteristics of foreign keys and introduces the important concept of a nonintelligent enterprise key. Enterprise keys (also called surrogate keys for data warehouses) are being emphasized as some concepts of object-orientation migrate into the relational technology world. A number of new review questions and problems and exercises are included, and revision also has further clarified the presentation of some of the key concepts. The chapter continues to emphasize the basic concepts of the relational data model and the role of the database designer in the logical design process.

CHAPTER 5: PHYSICAL DATABASE DESIGN AND PERFORMANCE This chapter describes the steps that are essential in achieving an efficient database design, with a strong focus on those aspects of database design and implementation that are typically within the control of a database professional in a modern database environment. The revised chapter is significantly shorter than the previous one, but we believe the reduction in length has been achieved without loss of significant content. The cuts are in areas that are either not relevant anymore because of changes in technology or not directly related to database design (e.g., storage technologies). Consequently, the chapter has a stronger and clearer focus on the core concepts. Several new review questions and problems and exercises are included. The chapter contains an emphasis on ways to improve database performance, with references to specific techniques available in Oracle and other DBMSs to improve database processing performance. The discussion of indexes includes descriptions of the types of indexes (primary and secondary indexes, join index, hash index table) that are widely available in database technologies as techniques to improve query processing speed. Appendix C provides excellent background on fundamental data structures for programs of study that need coverage of this topic. The chapter continues to emphasize the physical design process and the goals of that process.

Part IV: Implementation

CHAPTER 6: INTRODUCTION TO SQL This chapter presents a thorough introduction to the SQL used by most DBMSs (SQL:1999) and introduces the changes that are included in the latest proposed standard (SQL:200n). The coverage of SQL is extensive and divided into this and the next chapter. This chapter includes examples of SQL code, using mostly SQL:1999 and SQL:200n syntax, as well as some Oracle 11g and Microsoft SQL Server syntax. Some unique features of MySQL are mentioned. Views, both dynamic and materialized, are also covered. Chapter 6 explains the SQL commands needed to create and maintain a database and to program single-table queries. The history and SQL technology environment sections have been streamlined for the 10th edition. Coverage of dual-table, IS NULL/IS NOT NULL, more built-in functions, derived tables, and rules for aggregate functions and the GROUP BY clause are included or improved. New problems and exercises have been added to the chapter. It continues to use the Pine Valley Furniture Company case to illustrate a wide variety of practical queries and query results.

CHAPTER 7: ADVANCED SQL This chapter continues the description of SQL, with a careful explanation of multiple-table queries, transaction integrity, data dictionaries, triggers and stored procedures (the differences between which are now more clearly explained), and embedded SQL in other programming language programs. All forms of the OUTER JOIN command are covered. Standard SQL is also used in Chapter 7. This chapter illustrates how to store the results of a query in a derived table, the CAST command to convert data between different data types, and the CASE command for doing conditional processing in SQL. The chapter reduces its coverage of online analytical processing (OLAP) features of SQL:200n, which are also covered in Chapter 9. A new section on self-joins and an explanation of when to use EXISTS (NOT EXISTS) versus IN (NOT IN) are now included. The explanation of cursors with embedded SQL is enhanced. Emphasis continues on the set-processing style of SQL compared with the record-processing of programming languages with which the student may be familiar. New problems and exercises have been added to the chapter. The chapter continues to contain a clear explanation of subqueries and correlated subqueries, two of the most complex and powerful constructs in SQL.

CHAPTER 8: DATABASE APPLICATION DEVELOPMENT This chapter provides a modern discussion of the concepts of client/server architecture and applications, middleware, and database access in contemporary database environments. Technologies that are commonly used to create two- and three-tier applications are presented. Many figures are included to show the options in multitiered networks, including application and database servers, database processing distribution alternatives among network tiers, and browser (thin) clients. New to this edition is the presentation of sample application programs that demonstrate how to access databases from popular programming languages such as, Java, VB.NET, ASP.NET, JSP, and PHP. This chapter lays the technology groundwork for the Internet topics presented in the remainder of the text and highlights some of the key considerations in creating three-tier Internet-based applications. The chapter also presents expanded coverage of the role of Extensible Markup Language (XML) and related technologies in data storage and retrieval. Topics covered include basics of XML schemas, XQuery, and XSLT. The chapter concludes with an overview of Web services, associated standards and technologies, and their role in seamless, secure movement of data in Web-based applications. A brief introduction to service-oriented architecture (SOA) is also presented. Security topics, including Web security, are covered in Chapter 11.

CHAPTER 9: DATA WAREHOUSING This chapter describes the basic concepts of data warehousing, the reasons data warehousing is regarded as critical to competitive advantage in many organizations, and the database design activities and structures unique to data warehousing. An updated section reviews best practices for determining requirements for a dimensional model. A short new section introduces the emerging column databases technology, which has been developed especially for data warehousing applications.

New exercises provide hands-on practice with a data mart, using SQL and a BI tool called MicroStrategy that is supported on Teradata University Network. Topics include alternative data warehouse architectures and the dimensional data model (or star schema) for data warehouses. Coverage of architectures has been streamlined consistent with trends in data warehousing, and a deep explanation of how to handle slowly changing dimensional data is provided. Operational data store; independent, dependent, and logical data marts; and various forms of online analytical processing (OLAP) are defined (including the SAMPLE SQL command, which is useful for analyzing data from market research activities). User interfaces, including OLAP, data visualization, business performance management and dashboards, and data mining are also described.

Part V: Advanced Database Topics

CHAPTER 10: DATA QUALITY AND INTEGRATION This chapter, first introduced in the ninth edition, has been reorganized to better reflect the nature of enterprise data management (EDM) activities in organizations. The principles of data governance, which are at the core of EDM activities, are introduced first. This is followed by coverage of data quality. This chapter describes the need for an active program to manage data quality in organizations and outlines the steps that are considered today to be best practices for data quality management. Quality data are defined, and reasons for poor-quality data are identified. Methods for data quality improvement, such as data auditing, improving data capturing (a key part of database design), data stewardship and governance, TQM principles, modern data management technologies, and high-quality data models are all discussed. The current hot topic of master data management, one approach to integrating key business data, is motivated and explained. Different approaches to data integration are overviewed, and the reasons for each are outlined. The ETL process for data warehousing is discussed in detail. The authors believe that the material covered in this chapter continues to represent a major step forward in database management textbooks.

CHAPTER 11: DATA AND DATABASE ADMINISTRATION This chapter presents a thorough discussion of the importance and roles of data and database administration and describes a number of the key issues that arise when these functions are performed. This chapter emphasizes the changing roles and approaches of data and database administration, with emphasis on data quality and high performance. It contains a thorough discussion of database backup procedures, as well as extensively expanded and consolidated coverage of data security threats and responses, and data availability. The data security topics include database security policies, procedures, and technologies (including encryption and smart cards). New to this edition is expanded coverage of the role of databases in Sarbanes-Oxley compliance. We have again added to our discussion of open source DBMS to cover more on the benefits and hazards of this technology and how to choose an open source DBMS. In addition, the topic of heartbeat queries is included in the coverage of database performance improvements. The chapter continues to emphasize the critical importance of data and database management in managing data as a corporate asset.

CHAPTER 12: DISTRIBUTED DATABASES This chapter reviews the role, technologies, and unique database design opportunities of distributed databases. The objectives and trade-offs for distributed databases, data replication alternatives, factors in selecting a data distribution strategy, and distributed database vendors and products are covered. This chapter provides thorough coverage of database concurrency access controls. The revision of the chapter introduces several technical updates that are related to the significant advancements in both data management and networking technologies, which form the context for distributed database. An overview of this chapter is included in the printed textbook, and the full version of this chapter has been moved to the textbook's Web site. Many reviewers indicated that they seldom are able to cover this chapter in an introductory course, but having the material available is critical for advanced students or special topics. Having an overview in the printed text with the full chapter available to students provides the greatest flexibility and economy.

CHAPTER 13: OBJECT-ORIENTED DATA MODELING This chapter presents an introduction to object-oriented modeling using Object Management Group's Unified Modeling Language (UML). This chapter has been carefully reviewed to ensure consistency with the latest UML notation and best industry practices. UML provides an industry-standard notation for representing classes and objects. The chapter continues to emphasize basic object-oriented concepts, such as inheritance, encapsulation, composition, and polymorphism. The revised version of the chapter also includes several brand-new modeling exercises. As with Chapters 12 and 14, the full version of this chapter is available on the textbook's Web site, with a brief overview included in the printed text.

CHAPTER 14: USING RELATIONAL DATABASES TO PROVIDE OBJECT PERSISTENCE This chapter presents an up-to-date approach to how relational databases are used with object-oriented development environments, such as Java EE and Microsoft .NET. Object-oriented and relational approaches have critical design mismatches, which are outlined in the chapter, along with ways database and application developers can deal with these issues. The chapter reviews call-level application program interfaces, SQL query mapping frameworks, and object-relational mapping frameworks as approaches to providing object persistence, which is an essential need in modern development environments that integrate object-oriented development and relational databases. The chapter has been revised to take into account the changing landscape of object-relational mapping (ORM) technologies and the strengthening of the Java Persistence API (JPA) standard. Object-relational mapping is illustrated using the XML mapping files of Hibernate, the most popular ORM framework and the most widely used implementation of the JPA standard. As with Chapters 12 and 13, the full version of this chapter is available on the textbook's Web site, with a brief overview included in the printed text.

Appendices

The 10th edition contains three appendices intended for those who wish to explore certain topics in greater depth.

APPENDIX A: DATA MODELING TOOLS AND NOTATION This appendix addresses a need raised by many readers—how to translate the E-R notation in the text into the form used by the CASE tool or DBMS used in class. Specifically, this appendix compares the notations of CA ERwin Data Modeler r7.3, Oracle Designer 10g, Sybase PowerDesigner 15, and Microsoft Visio Pro 2003. Tables and illustrations show the notations used for the same constructs in each of these popular software packages.

APPENDIX B: ADVANCED NORMAL FORMS This appendix presents a description (with examples) of Boyce-Codd and fourth normal forms, including an example of BCNF to show how to handle overlapping candidate keys. Other normal forms are briefly introduced. The Web Resources section includes a reference for information on many advanced normal form topics.

APPENDIX C: DATA STRUCTURES This appendix describes several data structures that often underlie database implementations. Topics include the use of pointers, stacks, queues, sorted lists, inverted lists, and trees.

PEDAGOGY

A number of additions and improvements have been made to chapter-end materials to provide a wider and richer range of choices for the user. The most important of these improvements are the following:

1. ***Review Questions*** Questions have been updated to support new and enhanced chapter material.
2. ***Problems and Exercises*** This section has been reviewed in every chapter, and many chapters contain new problems and exercises to support updated chapter material. Of special interest are questions in many chapters that give students

opportunities to use the data sets provided for the text. Also, Problems and Exercises have be re-sequenced into roughly increasing order of difficulty, which should help instructors and students to find exercises appropriate for what they want to accomplish.

3. ***Field Exercises*** This section provides a set of "hands-on" minicases that can be assigned to individual students or to small teams of students. Field exercises range from directed field trips to Internet searches and other types of research exercises.
4. ***Case*** The Mountain View Community Hospital (MVCH) case was updated for the 10th edition only to the extent that chapters have been combined from the ninth edition. In each chapter, the case begins with a description of a realistic, modern hospital situation as it relates to that chapter. The case then presents a series of case questions and exercises that focus on specific aspects of the case. The final section includes project assignments, which tie together some issues and activities across chapters. These project assignments can be completed by individual students or by small project teams. This case provides an excellent means for students to gain hands-on experience with the concepts and tools they have studied.

5. ***Web Resources*** Each chapter contains a list of updated and validated URLs for Web sites that contain information that supplements the chapter. These Web sites cover online publication archives, vendors, electronic publications, industry standards organizations, and many other sources. These sites allow students and instructors to find updated product information, innovations that have appeared since the printing of the book, background information to explore topics in greater depth, and resources for writing research papers.

We have also updated the pedagogical features that helped make the 10th edition widely accessible to instructors and students. These features include the following:

1. ***Learning objectives*** appear at the beginning of each chapter, as a preview of the major concepts and skills students will learn from that chapter. The learning objectives also provide a great study review aid for students as they prepare for assignments and examinations.
2. ***Chapter introductions and summaries*** both encapsulate the main concepts of each chapter and link material to related chapters, providing students with a comprehensive conceptual framework for the course.
3. ***The chapter review*** includes the Review Questions, Problems and Exercises, and Field Exercises discussed earlier, also contains a Key Terms list to test the student's grasp of important concepts, basic facts, and significant issues.
4. ***A running glossary*** defines key terms in the page margins as they are discussed in the text. These terms are also defined at the end of the text, in the Glossary of Terms. Also included is the end-of-book Glossary of Acronyms for abbreviations commonly used in database management.

ORGANIZATION

We encourage instructors to customize their use of this book to meet the needs of both their curriculum and student career paths. The modular nature of the text, its broad coverage, extensive illustrations, and its inclusion of advanced topics and emerging issues make customization easy. The many references to current publications and Web sites can help instructors develop supplemental reading lists or expand classroom discussion beyond material presented in the text. The use of appendices for several advanced topics allows instructors to easily include or omit these topics.

The modular nature of the text allows the instructor to omit certain chapters or to cover chapters in a different sequence. For example, an instructor who wishes to emphasize data modeling may cover Chapter 13 on object-oriented data modeling along with or instead of Chapters 2 and 3. An instructor who wishes to cover only basic entity-relationship concepts (but not the enhanced E-R model) may skip Chapter 3 or cover it after Chapter 4 on the relational model. Three of the advanced topic

chapters—Chapters 12 through 14—are provided in overview form in the printed text and in full version on the book's Companion Web site; this gives the instructor added flexibility to cover these advanced topics at different levels.

We have contacted many adopters of *Modern Database Management* and asked them to share with us their syllabi. Most adopters cover the chapters in sequence, but several alternative sequences have also been successful. These alternatives include:

- Some instructors cover Chapter 11 on data and database administration immediately after Chapter 5 on physical database design and the relational model.
- To cover SQL as early as possible, instructors have effectively covered Chapters 6 and 7 immediately after Chapter 4; some have even covered Chapter 6 immediately after Chapter 1.
- Many instructors have students read appendices along with chapters, such as reading Appendix A on data modeling notations with Chapters 2 or 3 on E-R modeling, Appendix B on advanced normal forms with Chapter 4 on the relational model, and Appendix C on data structures with Chapter 5.

CASE TOOLS

Modern Database Management, 10th edition, offers adopters the option of acquiring outstanding CASE tools software packages from Microsoft and Oracle. Students can purchase this book packaged with the full editions of Microsoft Visio Pro and Oracle 11g at a greatly reduced fee. We are proud to offer such highly valued, powerful software packages to students at such a low cost. These packages can be used to draw data models, generate normalized relations from conceptual data models, and generate database definition code, among other tasks. These tools are also useful in other courses on information systems development.

THE SUPPLEMENT PACKAGE: WWW.PEARSONHIGHERED.COM/HOFFER

A comprehensive and flexible technology support package is available to enhance the teaching and learning experience. All instructor and student supplements are available on the text Web site: www.pearsonhighered.com/hoffer.

FOR STUDENTS The following online resources are available to students:

- The *Web Resources* module includes the Web links referenced at the end of each chapter in the text to help students further explore database management topics on the Web.
- A full *glossary* is available, along with a glossary of acronyms.
- *Links to sites where students can use our data sets* are provided. Although our data sets are provided in formats that are easily loaded on computers at your university or on student PCs, some instructors will not want the responsibility of supporting local data sets. The application service providers with whom we have developed arrangements (e.g., **www.teradatastudentnetwork.com**) provide thin-client interfaces to SQL coding environments. See the text's Web site and the inside front cover for more details.
- *Complete chapters on distributed databases, object-oriented data modeling, and object-oriented development with relational databases* allow you to learn in depth about topics that are overviewed in Chapters 12 through 14 of the textbook.
- *Accompanying databases* are also provided. Two versions of the Pine Valley Furniture Company case have been created and populated for the 10th edition. One version is scoped to match the textbook examples. A second version is fleshed out with more data and tables, as well as sample forms, reports, and modules coded in Visual Basic. This version is not complete, however, so that students can create missing tables and additional forms, reports, and modules. Databases are provided in several formats (ASCII tables, Oracle script, and Microsoft Access), but formats vary for the two versions. Some documentation of the databases is also provided. Both versions of the PVFC database are also provided on Teradata University Network.

- *Several new, custom-developed short videos that address key concepts and skills from different sections of the book* help students to learn material that may be more difficult to understand by using both the printed text and a mini lecture.

FOR INSTRUCTORS The following online resources are available to instructors:

- The *Instructor's Resource Manual* by Chelley Vician, University of St. Thomas, provides chapter-by-chapter instructor objectives, classroom ideas, and answers to Review Questions, Problems and Exercises, Field Exercises, and Project Case Questions. The Instructor's Resource Manual is available for download on the instructor area of the text's Web site.
- The *Test Item File and TestGen* by John P. Russo, Wentworth Institute of Technology, includes a comprehensive set of test questions in multiple-choice, true/false, and short-answer format, ranked according to level of difficulty and referenced with page numbers and topic headings from the text. The Test Item File is available in Microsoft Word and as the computerized TestGen. TestGen is a comprehensive suite of tools for testing and assessment. It allows instructors to easily create and distribute tests for their courses, either by printing and distributing through traditional methods or by online delivery via a local area network (LAN) server. Test Manager features Screen Wizards to assist you as you move through the program, and the software is backed with full technical support.
- *PowerPoint presentation slides* by Michel Mitri, James Madison University, feature lecture notes that highlight key terms and concepts. Instructors can customize the presentation by adding their own slides or editing existing ones.
- The *Image Library* is a collection of the text art organized by chapter. It includes all figures, tables, and screenshots (as permission allows) and can be used to enhance class lectures and PowerPoint slides.
- *Accompanying databases* are also provided. Two versions of the Pine Valley Furniture Company case have been created and populated for the 10th edition. One version is scoped to match the textbook examples. A second version is fleshed out with more data and tables, and sample forms, reports, and modules coded in Visual Basic. This version is not complete, however, so that students can create missing tables and additional forms, reports, and modules. Databases are provided in several formats (ASCII tables, Oracle script, and Microsoft Access), but formats vary for the two versions. Some documentation of the databases is also provided. Both versions of the PVFC database are also provided on Teradata University Network.
- A white paper by Willard Baird, Progress Telecom, *Achieving Optimal Database Performance*, provides supplemental reading for students interested in tuning an Oracle database. This paper provides a real-world perspective from a very experienced database administrator. It offers students and instructors an opportunity to consider the differences between material taught in classrooms and the hands-on experience gained through professional database administration.

MATERIALS FOR YOUR ONLINE COURSE

Pearson Prentice Hall supports our adopters using online courses by providing files ready for upload into Blackboard course management systems for our testing, quizzing, and other supplements. Please contact your local Pearson Prentice Hall representative for further information on your particular course.

COURSESMART eTEXTBOOK

CourseSmart is an exciting new choice for students looking to save money. As an alternative to purchasing the print textbook, students can purchase an electronic version of the same content and save up to 50% off the suggested list price of the print text. With a CourseSmart eTextbook, students can search the text, make notes online, print out reading assignments that incorporate lecture notes, and bookmark important passages for later review. For more information, or to purchase access to the CourseSmart eTextbook, visit **www.coursesmart.com**.

ACKNOWLEDGMENTS

We are grateful to numerous individuals who contributed to the preparation of *Modern Database Management*, 10th edition. First, we wish to thank our reviewers for their detailed suggestions and insights, characteristic of their thoughtful teaching style. As always, analysis of topics and depth of coverage provided by the reviewers was crucial. Our reviewers and others who gave us many useful comments to improve the text include Tamara Babaian, Bentley University; Gary Baram, Temple University; Timothy Bridges, University of Central Oklahoma; Bijoy Bordoloi, Southern Illinois University, Edwardsville; Traci Carte, University of Oklahoma; Wingyan Chung, Santa Clara University; Jon Gant, Syracuse University; Jinzhu Gao, University of the Pacific; Monica Garfield, Bentley University; Rick Gibson, American University; William H. Hochstettler III, Franklin University; Weiling Ke, Clarkson University; Dongwon Lee, Pennsylvania State University; Ingyu Lee, Troy University; Brian Mennecke, Iowa State University; Dat-Dao Nguyen, California State University, Northridge;Lara Preiser-Houy, California State Polytechnic University, Pomona; John Russo, Wentworth Institute of Technology; Ioulia Rytikova, George Mason University; Richard Segall, Arkansas State University; Chelley Vician, University of St. Thomas; and Daniel S. Weaver, Messiah College.

We received excellent input from people in industry, including Todd Walter, Carrie Ballinger, Rob Armstrong, and Dave Schoeff (all of Teradata Corp); Chad Gronbach and Philip DesAutels (Microsoft Corp.); Peter Gauvin (Ball Aerospace); Mridvika Raisinghani (Ernst & Young); Nisha Subramanian (Cook Medical, Inc.); Paul Longhurst (Overstock.com); Derek Strauss (Gavroshe International); Richard Hackathorn (Bolder Technology); and Michael Alexander (Open Access Technology, International).

We also thank Klara Nelson at the University of Tampa, who authored the Mountain View Community Hospital case study. This extensive real-world situation is a notable addition to the text. Linda Jayne, formerly operations manager of Suncoast Hospital, Largo, Florida, provided many relevant stories and validation of the situations faced by Mountain View Community Hospital, introduced in the eighth edition. We appreciate her taking the time to meet with us and to review the existing case.

Thanks are also due Willard Baird, Progress Telecom, a highly experienced Oracle database administrator, who has co-authored the Web site piece *Achieving Optimal Database Performance* with us. Willard is able to convey trends in database administration to us very clearly and in a timely fashion, and his experience with the trend toward more holistic approaches to database optimization that have occurred over the past few years have led to our decision to put up a piece on the Web site for students and instructors who desire a deeper understanding of this area.

We have special admiration for and gratitude to Chelley Vician of University of St. Thomas, author of the *Instructor's Resource Manual*. Chelley has been extremely careful in preparing the *Instructor's Resource Manual* and in the process has helped us clarify and fix various parts of the text. Chelley has added great value to this book. Christina Cooper, Indiana University, has been of considerable assistance to Chelley Vician, to ensure notational changes in the *Instructor's Resource Manual*. We also thank Sven Aelterman, Troy University, for his many excellent suggestions for improvements and clarifications throughout the text. And we thank Pragya Seth, an MS in Information Systems (MSIS) student at Indiana University, for help in reviewing material in several chapters.

We are also grateful to the staff and associates of Pearson Prentice Hall for their support and guidance throughout this project. In particular, we wish to thank Executive Editor Bob Horan, who coordinated the planning for the text, Editorial Project Manager Kelly Loftus, who kept us on track and made sure everything was complete, Production Project Manager Becca Richter, Senior Marketing Manager Anne Fahlgren, Media Project Manager Denise Vaughn, and Marketing Assistant Melinda Jensen. We extend special thanks to Jen Carley at PMG, whose supervision of the production process was excellent.

Finally, we give immeasurable thanks to our spouses, who endured many evenings and weekends of solitude for the thrill of seeing a book cover hang on a den

wall. In particular, we marvel at the commitment of Patty Hoffer, who has lived the lonely life of a textbook author's spouse through 10 editions over more than 28 years of late-night and weekend writing. Anne-Louise Klaus knew this time what to expect, and we want to sincerely thank her for being willing to continue her wholehearted support for Heikki's involvement in the project. For Gayathri Mani, this was the first time, but her unwavering support and understanding are equally appreciated. Much of the value of this text is due to their patience, encouragement, and love, but we alone bear the responsibility for any errors or omissions between the covers.

Jeffrey A. Hoffer

V. Ramesh

Heikki Topi