

SECOND EDITION

DATABASE

DESIGN, APPLICATION DEVELOPMENT,
& ADMINISTRATION

MICHAEL V. MANNINO



CONTENTS

Preface vii

PART 1

FOUNDATIONS OF RELATIONAL DATABASES 1

1

INTRODUCTION TO DATABASE MANAGEMENT 3

Learning Objectives 3

Overview 3

1.1 Database Characteristics 4

1.2 Features of Database Management Systems 7

1.2.1 Database Definition 7

1.2.2 Nonprocedural Access 9

1.2.3 Application Development and Procedural
Language Interface 10

1.2.4 Other Features 11

1.3 Development of Database Technology and Market
Structure 12

1.3.1 Evolution of Database Technology 13

1.3.2 Current Market for Database Software 14

1.4 Architectures of Database Management Systems 15

1.4.1 Data Independence and the Three Schema
Architecture 15

1.4.2 Distributed Processing and the Client–Server
Architecture 17

1.5 Organizational Impacts of Database Technology 19

1.5.1 Interacting with Databases 19

1.5.2 Information Resource Management 20

Closing Thoughts 21

Review Concepts 22

Questions 22

Problems 24

References for Further Study 24

2

THE RELATIONAL DATA MODEL 25

Learning Objectives 25

Overview 25

2.1 Basic Elements 26

2.1.1 Tables 26

2.1.2 Connections among Tables 28

2.1.3 Alternative Terminology 29

2.2 Integrity Rules 29

2.2.1 Definition of the Integrity Rules 30

2.2.2 Applying the Integrity Rules 31

2.2.3 Graphical Representation of Referential
Integrity 34

2.3 Delete and Update Actions for Referenced Rows 35

2.4 Operators of Relational Algebra 37

2.4.1 Restrict (Select) and Project Operators 37

2.4.2 Extended Cross Product Operator 39

2.4.3 Join Operator 40

2.4.4 Outer Join Operator 43

2.4.5 Union, Intersection, and Difference
Operators 46

2.4.6 Summarize Operator 49

2.4.7 Divide Operator 50

2.4.8 Summary of Operators 51

Closing Thoughts 52

Review Concepts 53

Questions 53

Problems 55

References for Further Study 57

Appendix 2.A: CREATE TABLE Statements for the
University Database Tables 58

Appendix 2.B: SQL:1999 Syntax Summary 60

3**QUERY FORMULATION WITH SQL 63**

Learning Objectives 63

Overview 63

3.1 Background 64

3.1.1 Brief History of SQL 64

3.1.2 Scope of SQL 65

3.2 Getting Started with SELECT 66

3.2.1 Single Table Problems 69

3.2.2 Joining Tables 74

3.2.3 Summarizing Tables with GROUP BY and HAVING 76

3.2.4 Improving the Appearance of Results 80

3.3 Conceptual Evaluation Process for SELECT Statements 82

3.4 Critical Questions for Query Formulation 87

3.5 Refining Query Formulation Skills with Examples 88

3.5.1 Joining Multiple Tables with the Cross Product Style 89

3.5.2 Joining Multiple Tables with the Join Operator Style 92

3.5.3 Self-Joins and Multiple Joins between Two Tables 95

3.5.4 Combining Joins and Grouping 96

3.5.5 Traditional Set Operators in SQL 97

3.6 SQL Modification Statements 99

Closing Thoughts 101

Review Concepts 102

Questions 105

Problems 106

References for Further Study 115

Appendix 3.A: SQL:1999 Syntax Summary 116

Appendix 3.B: Syntax Differences among Major Database Products 120

P A R T 2**DATABASE DEVELOPMENT 121**

.....

4**INTRODUCTION TO DATABASE DEVELOPMENT 123**

Learning Objectives 123

Overview 123

4.1 Information Systems 124

4.1.1 Components of Information Systems 124

4.1.2 Information Systems Development Process 125

4.2 Goals of Database Development 127

4.2.1 Develop a Common Vocabulary 127

4.2.2 Define the Meaning of Data 127

4.2.3 Ensure Data Quality 128

4.2.4 Find an Efficient Implementation 129

4.3 Database Development Process 129

4.3.1 Phases of Database Development 129

4.3.2 Skills in Database Development 134

4.4 Tools of Database Development 136

4.4.1 Diagramming 136

4.4.2 Documentation 136

4.4.3 Analysis 137

4.4.4 Prototyping Tools 137

4.4.5 Commercial CASE Tools 137

Closing Thoughts 141

Review Concepts 141

Questions 142

Problems 143

References for Further Study 143

5**UNDERSTANDING ENTITY RELATIONSHIP DIAGRAMS 145**

Learning Objectives 145

Overview 145

5.1 Introduction to Entity Relationship Diagrams 146

5.1.1 Basic Symbols 146

5.1.2 Relationship Cardinality 147

5.1.3 Comparison to Relational Database Diagrams 151

5.2 Understanding Relationships 151

5.2.1 Identification Dependency (Weak Entities and Identifying Relationships) 152

5.2.2 Relationship Patterns 153

5.2.3 Equivalence between 1-M and M-N Relationships 158

5.3 Classification in the Entity Relationship Model 159

5.3.1 Generalization Hierarchies 159

5.3.2 Disjointness and Completeness Constraints 160

5.3.3 Multiple Levels of Generalization 161

5.4 Notation Summary and Diagram Rules 161

5.4.1 Notation Summary 162

5.4.2 Diagram Rules 162

Closing Thoughts 167

Review Concepts 168

Questions 168

Problems 169

References for Further Study 174

6**DEVELOPING DATA MODELS FOR BUSINESS DATABASES 175**

Learning Objectives 175

Overview	175
6.1 Developing an ERD for the Water Utility Database	176
6.1.1 Database Description	176
6.1.2 Initial ERD	177
6.1.3 Refinements to the Initial ERD	179
6.1.4 Documenting an ERD	184
6.1.5 Detecting Common Design Problems	186
6.2 Converting an ERD to Relational Tables	188
6.2.1 Basic Conversion Rules	188
6.2.2 Converting Optional 1-M Relationships	193
6.2.3 Converting Generalization Hierarchies	195
6.2.4 Converting 1-1 Relationships	198
6.2.5 Comprehensive Conversion Example	199
6.3 Comparison to Other Notations	202
6.3.1 ERD Variations	202
6.3.2 Class Diagram Notation of the Unified Modeling Language	203
Closing Thoughts	206
Review Concepts	206
Questions	207
Problems	208
References for Further Study	225
7	
NORMALIZATION OF RELATIONAL TABLES	227
Learning Objectives	227
Overview	227
7.1 Overview of Relational Database Design	228
7.1.1 Avoidance of Modification Anomalies	228
7.1.2 Functional Dependencies	229
7.2 Normal Forms	231
7.2.1 First Normal Form	232
7.2.2 Second and Third Normal Forms	233
7.2.3 Boyce-Codd Normal Form	236
7.2.4 Simple Synthesis Procedure	238
7.3 Refining M-Way Relationships	241
7.3.1 Relationship Independence	241
7.3.2 Multivalued Dependencies and Fourth Normal Form (4NF)	244
7.4 Higher-Level Normal Forms	245
7.4.1 Fifth Normal Form (5NF)	245
7.4.2 Domain Key Normal Form (DKNF)	247
7.5 Practical Concerns about Normalization	247
7.5.1 Role of Normalization in the Database Development Process	247
7.5.2 Analyzing the Normalization Objective	248
Closing Thoughts	248
Review Concepts	249
Questions	249

Problems	251
References for Further Study	259

8

PHYSICAL DATABASE DESIGN 261

Learning Objectives	261
Overview	261
8.1 Overview of Physical Database Design	262
8.1.1 Storage Level of Databases	262
8.1.2 Objectives and Constraints	263
8.1.3 Inputs, Outputs, and Environment	264
8.1.4 Difficulties	266
8.2 Inputs of Physical Database Design	266
8.2.1 Table Profiles	266
8.2.2 Application Profiles	268
8.3 File Structures	269
8.3.1 Sequential Files	269
8.3.2 Hash Files	270
8.3.3 Multiway Tree (Btree) Files	272
8.3.4 Bitmap Indexes	278
8.3.5 Summary of File Structures	281
8.4 Query Optimization	281
8.4.1 Translation Tasks	281
8.4.2 Optimization Tips	285
8.5 Index Selection	287
8.5.1 Problem Definition	287
8.5.2 Trade-offs and Difficulties	288
8.5.3 Selection Rules	290
8.6 Additional Choices in Physical Database Design	293
8.6.1 Denormalization	293
8.6.2 Record Formatting	295
8.6.3 Parallel Processing	296
8.6.4 Other Ways to Improve Performance	298
Closing Thoughts	299
Review Concepts	299
Questions	300
Problems	302
References for Further Study	308

PART 3

APPLICATION DEVELOPMENT WITH RELATIONAL DATABASES 309

.....

9	
ADVANCED QUERY FORMULATION WITH SQL	311
Learning Objectives	311
Overview	311

- 9.1 Outer Join Problems 312
 - 9.1.1 SQL Support for Outer Join Problems 313
 - 9.1.2 Mixing Inner and Outer Joins 317
- 9.2 Understanding Nested Queries 318
 - 9.2.1 Type I Nested Queries 318
 - 9.2.2 Type II Nested Queries 321
 - 9.2.3 Solving Difference Problems with Type II Nested Queries 323
 - 9.2.4 Nested Queries in the FROM clause 328
- 9.3 Formulating Division Problems 330
 - 9.3.1 Review of the Divide Operator 330
 - 9.3.2 Simple Division Problems 331
 - 9.3.3 Advanced Division Problems 333
- 9.4 Null Value Considerations 336
 - 9.4.1 Effect on Simple Conditions 336
 - 9.4.2 Effect on Compound Conditions 337
 - 9.4.3 Effect on Aggregate Calculations and Grouping 339
- Closing Thoughts 340
- Review Concepts 341
- Questions 343
- Problems 344
- References for Further Study 348
- Appendix 9.A: Usage of Multiple Statements in Microsoft Access 349
- Appendix 9.B: SQL:1999 Syntax Summary 350

10

APPLICATION DEVELOPMENT WITH VIEWS 353

- Learning Objectives 353
- Overview 353
- 10.1 Background 354
 - 10.1.1 Motivation 354
 - 10.1.2 View Definition 355
- 10.2 Using Views for Retrieval 356
 - 10.2.1 Using Views in SELECT Statements 356
 - 10.2.2 Processing Queries with View References 358
- 10.3 Updating Using Views 360
 - 10.3.1 Single-Table Updatable Views 360
 - 10.3.2 Multiple-Table Updatable Views 364
- 10.4 Using Views in Hierarchical Forms 367
 - 10.4.1 What Is a Hierarchical Form? 367
 - 10.4.2 Relationship between Hierarchical Forms and Tables 369
 - 10.4.3 Query Formulation Skills for Hierarchical Forms 370
- 10.5 Using Views in Reports 374
 - 10.5.1 What Is a Hierarchical Report? 374
 - 10.5.2 Query Formulation Skills for Hierarchical Reports 375

- Closing Thoughts 377
- Review Concepts 377
- Questions 378
- Problems 379
- References for Further Study 386
- Appendix 10.A: SQL:1999 Syntax Summary 387
- Appendix 10.B: Rules for Updatable Join Views in Oracle 9i 388

11

STORED PROCEDURES AND TRIGGERS 389

- Learning Objectives 389
- Overview 389
 - 11.1 Database Programming Languages and PL/SQL 391
 - 11.1.1 Motivation for Database Programming Languages 391
 - 11.1.2 Design Issues 392
 - 11.1.3 PL/SQL Statements 394
 - 11.1.4 Executing PL/SQL Statements in Anonymous Blocks 401
 - 11.2 Stored Procedures 403
 - 11.2.1 PL/SQL Procedures 404
 - 11.2.2 PL/SQL Functions 407
 - 11.2.3 Using Cursors 410
 - 11.2.4 PL/SQL Packages 414
 - 11.3 Triggers 417
 - 11.3.1 Motivation and Classification of Triggers 418
 - 11.3.2 Oracle Triggers 418
 - 11.3.3 Understanding Trigger Execution 428
- Closing Thoughts 430
- Review Concepts 431
- Questions 432
- Problems 434
- References for Further Study 437
- Appendix 11.A: SQL:1999 Syntax Summary 438

12

VIEW DESIGN AND INTEGRATION 439

- Learning Objectives 439
- Overview 439
 - 12.1 Motivation for View Design and Integration 440
 - 12.2 View Design with Forms 441
 - 12.2.1 Form Analysis 441
 - 12.2.2 Analysis of M-Way Relationships Using Forms 450
 - 12.3 View Integration 452
 - 12.3.1 Incremental and Parallel Integration Approaches 452
 - 12.3.2 View Integration Examples 455

Closing Thoughts 458
 Review Concepts 458
 Questions 459
 Problems 459
 References for Further Study 462

13

DATABASE DEVELOPMENT FOR STUDENT LOAN, LIMITED 463

Learning Objectives 463

Overview 463

- 13.1 Case Description 464
 - 13.1.1 Overview 464
 - 13.1.2 Flow of Work 465
- 13.2 Conceptual Data Modeling 470
 - 13.2.1 ERD for the Loan Origination Form 470
 - 13.2.2 Incremental Integration after Adding the Disclosure Letter 470
 - 13.2.3 Incremental Integration after Adding the Statement of Account 473
 - 13.2.4 Incremental Integration after Adding the Loan Activity Report 475
- 13.3 Refining the Conceptual Schema 476
 - 13.3.1 Schema Conversion 476
 - 13.3.2 Normalization 478
- 13.4 Physical Database Design and Application Development 479
 - 13.4.1 Application and Table Profiles 480
 - 13.4.2 Index Selection 482
 - 13.4.3 Derived Data and Denormalization Decisions 482
 - 13.4.4 Other Implementation Decisions 483
 - 13.4.5 Application Development 483

Closing Thoughts 486

Review Concepts 486

Questions 487

Problems 488

Appendix 13.A: Glossary of Form and Report Fields 489

Appendix 13.B: CREATE TABLE Statements 492

P A R T 4

MANAGING DATABASE ENVIRONMENTS 497

.....

14

DATA AND DATABASE ADMINISTRATION 499

Learning Objectives 499

Overview 499

14.1 Organizational Context for Managing Databases 500

14.1.1 Database Support for Management Decision Making 500

14.1.2 Information Resource Management to Knowledge Management 501

14.1.3 Responsibilities of Data Administrators and Database Administrators 503

14.2 Tools of Database Administration 504

14.2.1 Security 504

14.2.2 Integrity Constraints 509

14.2.3 Management of Triggers and Stored Procedures 513

14.2.4 Data Dictionary Manipulation 515

14.3 Processes for Database Specialists 517

14.3.1 Data Planning 517

14.3.2 Selection and Evaluation of Database Management Systems 519

14.4 Managing Database Environments 523

14.4.1 Transaction Processing 523

14.4.2 Data Warehouse Processing 524

14.4.3 Distributed Environments 524

14.4.4 Object Database Management 525

Closing Thoughts 526

Review Concepts 527

Questions 528

Problems 531

References for Further Study 531

Appendix 14.A: SQL:1999 Syntax Summary 532

15

TRANSACTION MANAGEMENT 535

Learning Objectives 535

Overview 535

15.1 Basics of Database Transactions 536

15.1.1 Transaction Examples 536

15.1.2 Transaction Properties 539

15.2 Concurrency Control 540

15.2.1 Objective of Concurrency Control 540

15.2.2 Interference Problems 541

15.2.3 Concurrency Control Tools 543

15.3 Recovery Management 547

15.3.1 Data Storage Devices and Failure Types 547

15.3.2 Recovery Tools 548

15.3.3 Recovery Processes 550

15.4 Transaction Design Issues 554

15.4.1 Transaction Boundary and Hot Spots 555

15.4.2 Isolation Levels 557

15.4.3 Timing of Integrity Constraint Enforcement 559

15.4.4 Save Points 561

15.5 Workflow Management 561

15.5.1 Characterizing Workflows 561

15.5.2 Enabling Technologies 563

Closing Thoughts 564
 Review Concepts 565
 Questions 566
 Problems 568
 References for Further Study 572
 Appendix 15.A: SQL:1999 Syntax Summary 573

16

DATA WAREHOUSE TECHNOLOGY AND MANAGEMENT 575

Learning Objectives 575

Overview 575

- 16.1 Basic Concepts 576
 - 16.1.1 Transaction Processing versus Decision Support 576
 - 16.1.2 Characteristics of Data Warehouses 577
 - 16.1.3 Architectures for Data Warehouses 578
 - 16.1.4 Data Mining 580
 - 16.1.5 Applications of Data Warehouses 582
- 16.2 Multidimensional Representation of Data 582
 - 16.2.1 Example of a Multidimensional Data Cube 583
 - 16.2.2 Multidimensional Terminology 585
 - 16.2.3 Time-Series Data 586
 - 16.2.4 Data Cube Operations 587
- 16.3 Relational DBMS Support for Data Warehouses 590
 - 16.3.1 Relational Data Modeling for Multidimensional Data 591
 - 16.3.2 Dimension Representation 594
 - 16.3.3 Extensions to the GROUP BY Clause for Multidimensional Data 596
 - 16.3.4 Materialized Views and Query Rewriting 605
 - 16.3.5 Storage and Optimization Technologies 609
- 16.4 Populating a Data Warehouse 611
 - 16.4.1 Sources of Data 611
 - 16.4.2 Workflow for Maintaining a Data Warehouse 612
 - 16.4.3 Optimizing the Refresh Process 614

Closing Thoughts 616

Review Concepts 617

Questions 617

Problems 619

References for Further Study 624

17

CLIENT-SERVER PROCESSING AND DISTRIBUTED DATABASES 625

Learning Objectives 625

Overview 625

- 17.1 Overview of Distributed Processing and Distributed Data 626

- 17.1.1 Evolution of Distributed Processing and Distributed Data 626

- 17.1.2 Motivation for Distributed Processing 628

- 17.1.3 Motivation for Distributed Data 629

- 17.1.4 Summary of Advantages and Disadvantages 630

- 17.2 Client-Server Database Architectures 630

- 17.2.1 Design Issues 631

- 17.2.2 Description of Architectures 633

- 17.3 Web Database Connectivity 636

- 17.3.1 Internet Basics 637

- 17.3.2 Common Gateway Interface 640

- 17.3.3 Server-Side Connectivity 641

- 17.3.4 Client-Side Connectivity 642

- 17.3.5 Summary of Web Database Connectivity Approaches 643

- 17.4 Architectures for Distributed Database Management Systems 643

- 17.4.1 Component Architecture 644

- 17.4.2 Schema Architectures 646

- 17.5 Transparency for Distributed Database Processing 648

- 17.5.1 Motivating Example 648

- 17.5.2 Fragmentation Transparency 650

- 17.5.3 Location Transparency 651

- 17.5.4 Local Mapping Transparency 652

- 17.6 Distributed Database Processing 654

- 17.6.1 Distributed Query Processing 654

- 17.6.2 Distributed Transaction Processing 656

Closing Thoughts 659

Review Concepts 659

Questions 660

Problems 662

References for Further Study 664

18

OBJECT DATABASE MANAGEMENT SYSTEMS 665

Learning Objectives 665

Overview 665

- 18.1 Motivation for Object Database Management 666

- 18.1.1 Complex Data 666

- 18.1.2 Type System Mismatch 667

- 18.1.3 Application Examples 667

- 18.2 Object-Oriented Principles 668

- 18.2.1 Encapsulation 668

- 18.2.2 Inheritance 670

- 18.2.3 Polymorphism 672

- 18.2.4 Programming Languages versus DBMSs 673

- 18.3 Architectures for Object Database Management 674

- 18.3.1 Large Objects and External Software 674

- 18.3.2 Specialized Media Servers 675

18.3.3 Object Database Middleware	676
18.3.4 Object Relational Database Management Systems for User-Defined Types	677
18.3.5 Object-Oriented Database Management Systems	679
18.3.6 Summary of Object Database Architectures	680
18.4 Object Database Features in SQL:1999	680
18.4.1 User-Defined Types	681
18.4.2 Table Definitions and Subtable Families	683
18.4.3 Manipulating Complex Objects and Subtable Families	687
18.5 Object Database Features in Oracle 9i	689
18.5.1 Defining User-Defined Types and Typed Tables in Oracle 9i	689

18.5.2 Using Typed Tables in Oracle 9i	693
18.5.3 Other Object Features in Oracle 9i	695
Closing Thoughts	696
Review Concepts	697
Questions	697
Problems	699
References for Further Study	702

Glossary 703

Bibliography 720

Index 723