Part I Background **1**

Chapter 1 Introduction to Databases **3**

1.1 Introduction 4

1.2 Traditional File-Based Systems 7
1.2.1 File-Based Approach 7
1.2.2 Limitations of the File-Based Approach 12

1.3 Database Approach 14
1.3.1 The Database 15
1.3.2 The Database Management System (DBMS) 16
1.3.3 (Database) Application Programs 17
1.3.4 Components of the DBMS Environment 18
1.3.5 Database Design: The Paradigm Shift 21

1.4 Roles in the Database Environment 21
1.4.1 Data and Database Administrators 21
1.4.2 Database Designers 22
1.4.3 Application Developers 23
1.4.4 End-Users 23

1.5 History of Database Management Systems 23

1.6 Advantages and Disadvantages of DBMSs 27

Chapter Summary *31*
Review Questions *32*
Exercises *32*

Chapter 2 Database Environment **35**

2.1 The Three-Level ANSI-SPARC Architecture 36
2.1.1 External Level 37
2.1.2 Conceptual Level 38
2.1.3 Internal Level 38
2.1.4 Schemas, Mappings, and Instances 39
2.1.5 Data Independence 40

2.2 Database Languages 41
2.2.1 The Data Definition Language (DDL) 42

2.2.2 The Data Manipulation Language (DML) 42
2.2.3 Fourth-Generation Languages (4GLs) 44

2.3 Data Models and Conceptual Modeling 45
2.3.1 Object-Based Data Models 46
2.3.2 Record-Based Data Models 46
2.3.3 Physical Data Models 49
2.3.4 Conceptual Modeling 49

2.4 Functions of a DBMS 49

Chapter Summary 54
Review Questions 55
Exercises 55

Chapter 3 Database Architectures and the Web **57**

3.1 Multi-user DBMS Architectures 58
3.1.1 Teleprocessing 58
3.1.2 File-Server Architecture 59
3.1.3 Traditional Two-Tier Client–Server Architecture 60
3.1.4 Three-Tier Client–Server Architecture 63
3.1.5 *N*-Tier Architectures 64
3.1.6 Middleware 65
3.1.7 Transaction Processing Monitors 67

3.2 Web Services and Service-Oriented Architectures 69
3.2.1 Web Services 69
3.2.2 Service-Oriented Architectures (SOA) 70

3.3 Distributed DBMSs 72

3.4 Data Warehousing 74

3.5 Components of a DBMS 77

3.6 Oracle Architecture 80
3.6.1 Oracle's Logical Database Structure 80
3.6.2 Oracle's Physical Database Structure 82

Chapter Summary 86
Review Questions 87
Exercises 87

Part 2 The Relational Model and Languages **89**

Chapter 4 The Relational Model **91**

4.1 Brief History of the Relational Model 92

4.2 Terminology 94
4.2.1 Relational Data Structure 94
4.2.2 Mathematical Relations 97

4.2.3 Database Relations 98
4.2.4 Properties of Relations 98
4.2.5 Relational Keys 100
4.2.6 Representing Relational Database Schemas 101

4.3 Integrity Constraints 103
4.3.1 Nulls 103
4.3.2 Entity Integrity 104
4.3.3 Referential Integrity 104
4.3.4 General Constraints 105

4.4 Views 105
4.4.1 Terminology 105
4.4.2 Purpose of Views 106
4.4.3 Updating Views 107

Chapter Summary *107*
Review Questions *108*
Exercises *108*

Chapter 5 Relational Algebra and Relational Calculus 109

5.1 The Relational Algebra 110
5.1.1 Unary Operations 110
5.1.2 Set Operations 113
5.1.3 Join Operations 116
5.1.4 Division Operation 119
5.1.5 Aggregation and Grouping Operations 120
5.1.6 Summary of the Relational Algebra Operations 122

5.2 The Relational Calculus 123
5.2.1 Tuple Relational Calculus 123
5.2.2 Domain Relational Calculus 126

5.3 Other Languages 128

Chapter Summary *129*
Review Questions *129*
Exercises *130*

Chapter 6 SQL: Data Manipulation 133

6.1 Introduction to SQL 134
6.1.1 Objectives of SQL 134
6.1.2 History of SQL 135
6.1.3 Importance of SQL 137
6.1.4 Terminology 137

6.2 Writing SQL Commands 137

6.3 Data Manipulation 138
6.3.1 Simple Queries 139
6.3.2 Sorting Results (ORDER BY Clause) 147

6.3.3 Using the SQL Aggregate Functions 149
6.3.4 Grouping Results (GROUP BY Clause) 151
6.3.5 Subqueries 154
6.3.6 ANY and ALL 156
6.3.7 Multi-table Queries 158
6.3.8 EXISTS and NOT EXISTS 164
6.3.9 Combining Result Tables (UNION, INTERSECT, EXCEPT) 165
6.3.10 Database Updates 167

Chapter Summary *171*
Review Questions *172*
Exercises *172*

Chapter 7 SQL: Data Definition **175**

7.1 The ISO SQL Data Types 176
7.1.1 SQL Identifiers 176
7.1.2 SQL Scalar Data Types 177
7.1.3 Exact Numeric Data 178

7.2 Integrity Enhancement Feature 181
7.2.1 Required Data 182
7.2.2 Domain Constraints 182
7.2.3 Entity Integrity 183
7.2.4 Referential Integrity 184
7.2.5 General Constraints 185

7.3 Data Definition 185
7.3.1 Creating a Database 186
7.3.2 Creating a Table (CREATE TABLE) 187
7.3.3 Changing a Table Definition (ALTER TABLE) 190
7.3.4 Removing a Table (DROP TABLE) 191
7.3.5 Creating an Index (CREATE INDEX) 192
7.3.6 Removing an Index (DROP INDEX) 192

7.4 Views 193
7.4.1 Creating a View (CREATE VIEW) 193
7.4.2 Removing a View (DROP VIEW) 195
7.4.3 View Resolution 196
7.4.4 Restrictions on Views 197
7.4.5 View Updatability 197
7.4.6 WITH CHECK OPTION 198
7.4.7 Advantages and Disadvantages of Views 200
7.4.8 View Materialization 202

7.5 Transactions 203
7.5.1 Immediate and Deferred Integrity Constraints 204

7.6 Discretionary Access Control 204
7.6.1 Granting Privileges to Other Users (GRANT) 206
7.6.2 Revoking Privileges from Users (REVOKE) 207

Chapter Summary 209
Review Questions 210
Exercises 210

Chapter 8 Advanced SQL 213

8.1 The SQL Programming Language 214
8.1.1 Declarations 214
8.1.2 Assignments 215
8.1.3 Control Statements 216
8.1.4 Exceptions in PL/SQL 218
8.1.5 Cursors in PL/SQL 219

8.2 Subprograms, Stored Procedures, Functions, and Packages 222

8.3 Triggers 223

8.4 Recursion 229

Chapter Summary 230
Review Questions 231
Exercises 231

Chapter 9 Query-By-Example 233

9.1 Introduction to Microsoft Office Access Queries 234

9.2 Building Select Queries Using QBE 236
9.2.1 Specifying Criteria 237
9.2.2 Creating Multi-table Queries 239
9.2.3 Calculating Totals 242

9.3 Using Advanced Queries 242
9.3.1 Parameter Query 242
9.3.2 Crosstab Query 243
9.3.3 Find Duplicates Query 246
9.3.4 Find Unmatched Query 248
9.3.5 Autolookup Query 249

9.4 Changing the Content of Tables Using Action Queries 250
9.4.1 Make-Table Action Query 250
9.4.2 Delete Action Query 250
9.4.3 Update Action Query 253
9.4.4 Append Action Query 253

Exercises 258

Part 3 Database Analysis and Design 259

Chapter 10 Database System Development Lifecycle 261

10.1 The Information Systems Lifecycle 262

10.2 The Database System Development Lifecycle 263

10.3 Database Planning 263

10.4 System Definition 266
10.4.1 User Views 266

10.5 Requirements Collection and Analysis 266
10.5.1 Centralized Approach 268
10.5.2 View Integration Approach 268

10.6 Database Design 270
10.6.1 Approaches to Database Design 271
10.6.2 Data Modeling 271
10.6.3 Phases of Database Design 272

10.7 DBMS Selection 275
10.7.1 Selecting the DBMS 275

10.8 Application Design 279
10.8.1 Transaction Design 280
10.8.2 User Interface Design Guidelines 281

10.9 Prototyping 283

10.10 Implementation 283

10.11 Data Conversion and Loading 284

10.12 Testing 284

10.13 Operational Maintenance 285

10.14 CASE Tools 286

Chapter Summary *288*
Review Questions *289*
Exercises *290*

Chapter 11 Database Analysis and the *DreamHome* Case Study 291

11.1 When Are Fact-Finding Techniques Used? 292

11.2 What Facts Are Collected? 293

11.3 Fact-Finding Techniques 294
11.3.1 Examining Documentation 294
11.3.2 Interviewing 294
11.3.3 Observing the Enterprise in Operation 295

11.3.4 Research 296
11.3.5 Questionnaires 296

11.4 Using Fact-Finding Techniques: A Worked Example 297
11.4.1 The *DreamHome* Case Study—An Overview of the Current System 298
11.4.2 The *DreamHome* Case Study—Database Planning 302
11.4.3 The *DreamHome* Case Study—System Definition 308
11.4.4 The *DreamHome* Case Study—Requirements Collection and Analysis 309
11.4.5 The *DreamHome* Case Study—Database Design 317

Chapter Summary *318*
Review Questions *318*
Exercises *318*

Chapter 12 Entity–Relationship Modeling **321**

12.1 Entity Types 322

12.2 Relationship Types 324
12.2.1 Degree of Relationship Type 326
12.2.2 Recursive Relationship 328

12.3 Attributes 329
12.3.1 Simple and Composite Attributes 329
12.3.2 Single-valued and Multi-valued Attributes 330
12.3.3 Derived Attributes 330
12.3.4 Keys 331

12.4 Strong and Weak Entity Types 333

12.5 Attributes on Relationships 334

12.6 Structural Constraints 335
12.6.1 One-to-One (1:1) Relationships 336
12.6.2 One-to-Many (1 :*) Relationships 337
12.6.3 Many-to-Many (*:*) Relationships 338
12.6.4 Multiplicity for Complex Relationships 339
12.6.5 Cardinality and Participation Constraints 340

12.7 Problems with ER Models 342
12.7.1 Fan Traps 342
12.7.2 Chasm Traps 344

Chapter Summary *346*
Review Questions *346*
Exercises *347*

Chapter 13 Enhanced Entity–Relationship Modeling **349**

13.1 Specialization/Generalization 350
13.1.1 Superclasses and Subclasses 350

13.1.2 Superclass/Subclass Relationships 351
13.1.3 Attribute Inheritance 352
13.1.4 Specialization Process 352
13.1.5 Generalization Process 353
13.1.6 Constraints on Specialization/Generalization 356
13.1.7 Worked Example of using Specialization/Generalization to Model the Branch View of the *DreamHome* Case Study 357
13.2 Aggregation 361
13.3 Composition 362
Chapter Summary *363*
Review Questions *364*
Exercises *364*

Chapter 14 Normalization **365**
14.1 The Purpose of Normalization 366
14.2 How Normalization Supports Database Design 367
14.3 Data Redundancy and Update Anomalies 368
14.3.1 Insertion Anomalies 369
14.3.2 Deletion Anomalies 369
14.3.3 Modification Anomalies 370
14.4 Functional Dependencies 370
14.4.1 Characteristics of Functional Dependencies 370
14.4.2 Identifying Functional Dependencies 374
14.4.3 Identifying the Primary Key for a Relation Using Functional Dependencies 377
14.5 The Process of Normalization 378
14.6 First Normal Form (1NF) 380
14.7 Second Normal Form (2NF) 384
14.8 Third Normal Form (3NF) 385
14.9 General Definitions of 2NF and 3NF 387
Chapter Summary *389*
Review Questions *389*
Exercises *390*

Chapter 15 Advanced Normalization **393**
15.1 More on Functional Dependencies 394
15.1.1 Inference Rules for Functional Dependencies 394
15.1.2 Minimal Sets of Functional Dependencies 396

15.2 Boyce–Codd Normal Form (BCNF) 397
15.2.1 Definition of BCNF 397
15.3 Review of Normalization Up to BCNF 400
15.4 Fourth Normal Form (4NF) 405
15.4.1 Multi-Valued Dependency 406
15.4.2 Definition of Fourth Normal Form 407
15.5 Fifth Normal Form (5NF) 407
15.5.1 Lossless-Join Dependency 408
15.5.2 Definition of Fifth Normal Form 408
Chapter Summary *410*
Review Questions *410*
Exercises *411*

Part 4 Methodology **413**

Chapter 16 Methodology—Conceptual Database Design **415**
16.1 Introduction to the Database Design Methodology 416
16.1.1 What Is a Design Methodology? 416
16.1.2 Conceptual, Logical, and Physical Database Design 417
16.1.3 Critical Success Factors in Database Design 417
16.2 Overview of the Database Design Methodology 418
16.3 Conceptual Database Design Methodology 420
Step 1: Build Conceptual Data Model 420
Chapter Summary *436*
Review Questions *436*
Exercises *437*

Chapter 17 Methodology—Logical Database Design for the Relational Model **439**
17.1 Logical Database Design Methodology for the Relational Model 440
Step 2: Build Logical Data Model 440
Chapter Summary *468*
Review Questions *469*
Exercises *469*

Chapter 18 Methodology—Physical Database Design for Relational Databases **471**
18.1 Comparison of Logical and Physical Database Design 472
18.2 Overview of the Physical Database Design Methodology 473

18.3 The Physical Database Design Methodology for Relational Databases 474
Step 3: Translate Logical Data Model for Target DBMS 474
Step 4 479
Step 5: Design User Views 492
Step 6: Design Security Mechanisms 492
Chapter Summary *493*
Review Questions *494*
Exercises *494*

Chapter 19 Methodology—Monitoring and Tuning the Operational System 495

19.1 Denormalizing and Introducing Controlled Redundancy 495
Step 7: Consider the Introduction of Controlled Redundancy 495
19.2 Monitoring the System to Improve Performance 508
Step 8: Monitor and Tune the Operational System 508
Chapter Summary *512*
Review Questions *513*
Exercises *513*

Part 5 Selected Database Issues 515

Chapter 20 Security and Administration 517

20.1 Database Security 518
20.1.1 Threats 519
20.2 Countermeasures—Computer-Based Controls 521
20.2.1 Authorization 522
20.2.2 Access Controls 523
20.2.3 Views 526
20.2.4 Backup and Recovery 526
20.2.5 Integrity 527
20.2.6 Encryption 527
20.2.7 RAID (Redundant Array of Independent Disks) 528
20.3 Security in Microsoft Office Access DBMS 531
20.4 Security in Oracle DBMS 533
20.5 DBMSs and Web Security 537
20.5.1 Proxy Servers 538
20.5.2 Firewalls 538

20.5.3 Message Digest Algorithms and Digital Signatures 539
20.5.4 Digital Certificates 539
20.5.5 Kerberos 540
20.5.6 Secure Sockets Layer and Secure HTTP 540
20.5.7 Secure Electronic Transactions and Secure Transaction Technology 541
20.5.8 Java Security 542
20.5.9 ActiveX Security 544

20.6 Data Administration and Database Administration 544
20.6.1 Data Administration 545
20.6.2 Databse Administration 546
20.6.3 Comparison of Data and Database Administration 546

Chapter Summary *547*
Review Questions *548*
Exercises *548*

Chapter 21 Professional, Legal, and Ethical Issues in Data Management 549

21.1 Defining Legal and Ethical Issues in IT 550
21.1.1 Defining Ethics in the Context of IT 550
21.1.2 The Difference Between Ethical and Legal Behavior 551
21.1.3 Ethical Behavior in IT 552

21.2 Legislation and Its Impact on the IT Function 553
21.2.1 Securities and Exchange Commission (SEC) Regulation National Market System (NMS) 553
21.2.2 The Sarbanes-Oxley Act, COBIT, and COSO 553
21.2.3 The Health Insurance Portability and Accountability Act 555
21.2.4 The European Union (EU) Directive on Data Protection of 1995 555
21.2.5 The United Kingdom's Data Protection Act of 1998 556
21.2.6 International Banking—Basel II Accords 557

21.3 Establishing a Culture of Legal and Ethical Data Stewardship 558
21.3.1 Developing an Organization-Wide Policy for Legal and Ethical Behavior 558
21.3.2 Professional Organizations and Codes of Ethics 559
21.3.3 Developing an Organization-Wide Policy for Legal and Ethical Behavior for *DreamHome* 561

21.4 Intellectual Property 563
21.4.1 Patent 563
21.4.2 Copyright 564
21.4.3 Trademark 564

21.4.4 Intellectual Property Rights Issues for Software 565
21.4.5 Intellectual Property Rights Issues for Data 566

Chapter Summary *566*
Review Questions *567*
Exercises *567*

Chapter 22 Transaction Management **569**

22.1 Transaction Support 570
22.1.1 Properties of Transactions 573
22.1.2 Database Architecture 573

22.2 Concurrency Control 574
22.2.1 The Need for Concurrency Control 574
22.2.2 Serializability and Recoverability 577
22.2.3 Locking Methods 585
22.2.4 Deadlock 591
22.2.5 Timestamping Methods 594
22.2.6 Multiversion Timestamp Ordering 597
22.2.7 Optimistic Techniques 598
22.2.8 Granularity of Data Items 599

22.3 Database Recovery 602
22.3.1 The Need for Recovery 602
22.3.2 Transactions and Recovery 603
22.3.3 Recovery Facilities 606
22.3.4 Recovery Techniques 609
22.3.5 Recovery in a Distributed DBMS 611

22.4 Advanced Transaction Models 611
22.4.1 Nested Transaction Model 613
22.4.2 Sagas 614
22.4.3 Multilevel Transaction Model 615
22.4.4 Dynamic Restructuring 616
22.4.5 Workflow Models 617

22.5 Concurrency Control and Recovery in Oracle 618
22.5.1 Oracle's Isolation Levels 619
22.5.2 Multiversion Read Consistency 619
22.5.3 Deadlock Detection 621
22.5.4 Backup and Recovery 621

Chapter Summary *624*
Review Questions *625*
Exercises *625*

Chapter 23 Query Processing **627**

23.1 Overview of Query Processing 629

23.2 Query Decomposition 632

23.3 Heuristical Approach to Query Optimization 636
23.3.1 Transformation Rules for the Relational Algebra Operations 636
23.3.2 Heuristical Processing Strategies 641
23.4 Cost Estimation for the Relational Algebra Operations 642
23.4.1 Database Statistics 642
23.4.2 Selection Operation ($S = \sigma_p(R)$) 643
23.4.3 Join Operation ($T = (R \bowtie_F S)$) 650
23.4.4 Projection Operation ($S = \Pi_{A_1, A_2, \ldots, A_m}(R)$) 657
23.4.5 The Relational Algebra Set Operations ($T = R \cup S$, $T = R \cap S$, $T = R - S$) 659
23.5 Enumeration of Alternative Execution Strategies 660
23.5.1 Pipelining 661
23.5.2 Linear Trees 661
23.5.3 Physical Operators and Execution Strategies 662
23.5.4 Reducing the Search Space 664
23.5.5 Enumerating Left-Deep Trees 665
23.5.6 Semantic Query Optimization 666
23.5.7 Alternative Approaches to Query Optimization 667
23.5.8 Distributed Query Optimization 668
23.6 Query Optimization in Oracle 668
23.6.1 Rule-Based and Cost-Based Optimization 668
23.6.2 Histograms 672
23.6.3 Viewing the Execution Plan 674
Chapter Summary *675*
Review Questions *676*
Exercises *676*

Part 6 Distributed DBMSs and Replication 679

Chapter 24 Distributed DBMSs—Concepts and Design 681

24.1 Introduction 682
24.1.1 Concepts 683
24.1.2 Advantages and Disadvantages of DDBMSs 687
24.1.3 Homogeneous and Heterogeneous DDBMSs 690
24.2 Overview of Networking 693
24.3 Functions and Architectures of a DDBMS 697
24.3.1 Functions of a DDBMS 697
24.3.2 Reference Architecture for a DDBMS 697
24.3.3 Reference Architecture for a Federated MDBS 699
24.3.4 Component Architecture for a DDBMS 700
24.4 Distributed Relational Database Design 701
24.4.1 Data Allocation 702
24.4.2 Fragmentation 703

24.5 Transparencies in a DDBMS 712
24.5.1 Distribution Transparency 712
24.5.2 Transaction Transparency 715
24.5.3 Performance Transparency 718
24.5.4 DBMS Transparency 720
24.5.5 Summary of Transparencies in a DDBMS 720

24.6 Date's Twelve Rules for a DDBMS 721

Chapter Summary *723*
Review Questions *724*
Exercises *724*

Chapter 25 Distributed DBMSs—Advanced Concepts 727

25.1 Distributed Transaction Management 728

25.2 Distributed Concurrency Control 729
25.2.1 Objectives 729
25.2.2 Distributed Serializability 730
25.2.3 Locking Protocols 730
25.2.4 Timestamp Protocols 733

25.3 Distributed Deadlock Management 733

25.4 Distributed Database Recovery 737
25.4.1 Failures in a Distributed Environment 737
25.4.2 How Failures Affect Recovery 738
25.4.3 Two-Phase Commit (2PC) 739
25.4.4 Three-Phase Commit (3PC) 745
25.4.5 Network Partitioning 749

25.5 The X/Open Distributed Transaction Processing Model 750

25.6 Distributed Query Optimization 753
25.6.1 Data Localization 754
25.6.2 Distributed Joins 758
25.6.3 Global Optimization 759

25.7 Distribution in Oracle 763
25.7.1 Oracle's DDBMS Functionality 763

Chapter Summary *768*
Review Questions *769*
Exercises *770*

Chapter 26 Replication and Mobile Databases 771

26.1 Introduction to Database Replication 772
26.1.1 Synchronous Versus Asynchronous Replication 773
26.1.2 Applications of Replication 774

26.2 Replication Servers 774
26.2.1 Replication Server Functionality 775
26.2.2 Data Ownership 775
26.2.3 Implmentation Issues 779

26.3. Introduction to Mobile Databases 782
26.3.1 Mobile DBMSs 784
26.3.2 Issues with Mobile DBMSs 784

26.4 Oracle Replication 790
26.4.1 Oracle's Replication Functionality 790

Chapter Summary 796
Review Questions 797
Exercises 797

Part 7 Object DBMSs 799

Chapter 27 Object-Oriented DBMSs—Concepts and Design 801

27.1 Advanced Database Applications 803

27.2 Weaknesses of RDBMSs 807

27.3 Storing Objects in a Relational Database 812
27.3.1 Mapping Classes to Relations 813
27.3.2 Accessing Objects in the Relational Database 814

27.4 Next-Generation Database Systems 816

27.5 Introduction to OODBMSs 817
27.5.1 Definition of Object-Oriented DBMSs 818
27.5.2 Functional Data Models 819
27.5.3 Persistent Programming Languages 824
27.5.4 The *Object-Oriented Database System Manifesto* 825
27.5.5 Alternative Strategies for Developing an OODBMS 828

27.6 Persistence in OODBMSs 829
27.6.1 Pointer Swizzling Techniques 831
27.6.2 Accessing an Object 834
27.6.3 Persistence Schemes 836
27.6.4 Orthogonal Persistence 837

27.7 Issues in OODBMSs 839
27.7.1 Transactions 839
27.7.2 Versions 840
27.7.3 Schema Evolution 841
27.7.4 Architecture 844
27.7.5 Benchmarking 846

27.8 Advantages and Disadvantages of OODBMSs 849
27.8.1 Advantages 849
27.8.2 Disadvantages 851

27.9 Object-Oriented Database Design 853
27.9.1 Comparison of Object-Oriented Data Modeling and Conceptual Data Modeling 853
27.9.2 Relationships and Referential Integrity 854
27.9.3 Behavioral Design 856

27.10 Object-Oriented Analysis and Design with UML 858
27.10.1 UML Diagrams 859
27.10.2 Usage of UML in the Methodology for Database Design 864

Chapter Summary *866*
Review Questions *867*
Exercises *868*

Chapter 28 Object-Oriented DBMSs—Standards and Systems 871

28.1 Object Management Group 872
28.1.1 Background 872
28.1.2 The Common Object Request Broker Architecture 875
28.1.3 Other OMG Specifications 880
28.1.4 Model-Driven Architecture 883

28.2 Object Data Standard ODMG 3.0, 1999 883
28.2.1 Object Data Management Group 885
28.2.2 The Object Model 886
28.2.3 The Object Definition Language 894
28.2.4 The Object Query Language 897
28.2.5 Other Parts of the ODMG Standard 903
28.2.6 Mapping the Conceptual Design to a Logical (Object-Oriented) Design 906

28.3 ObjectStore 907
28.3.1 Architecture 907
28.3.2 Building an ObjectStore Application 910
28.3.3 Data Definition in ObjectStore 911
28.3.4 Data Manipulation in ObjectStore 915

Chapter Summary *918*
Review Questions *919*
Exercises *919*

Chapter 29 Object-Relational DBMSs 921

29.1 Introduction to Object-Relational Database Systems 922

29.2 The Third-Generation Database Manifestos 925
29.2.1 The *Third-Generation Database System Manifesto* 926
29.2.2 *The Third Manifesto* 926

29.3 Postgres—An Early ORDBMS 929
29.3.1 Objectives of Postgres 929
29.3.2 Abstract Data Types 929
29.3.3 Relations and Inheritance 930
29.3.4 Object Identity 931

29.4 SQL:2008 932
29.4.1 Row Types 933
29.4.2 User-Defined Types 934
29.4.3 Subtypes and Supertypes 936
29.4.4 User-Defined Routines 939
29.4.5 Polymorphism 940
29.4.6 Reference Types and Object Identity 941
29.4.7 Creating Tables 942
29.4.8 Querying Data 945
29.4.9 Collection Types 946
29.4.10 Typed Views 950
29.4.11 Persistent Stored Modules 950
29.4.12 Triggers 951
29.4.13 Large Objects 954
29.4.14 Recursion 955

29.5 Query Processing and Optimization 955
29.5.1 New Index Types 959

29.6 Object-Oriented Extensions in Oracle 959
29.6.1 User-Defined Data Types 960
29.6.2 Manipulating Object Tables 965
29.6.3 Object Views 966
29.6.4 Privileges 967

29.7 Comparison of ORDBMS and OODBMS 968

Chapter Summary *969*
Review Questions *969*
Exercises *970*

Part 8 The Web and DBMSs **971**

Chapter 30 Web Technology and DBMSs **973**

30.1 Introduction to the Internet and the Web 974
30.1.1 Intranets and Extranets 976
30.1.2 e-Commerce and e-Business 977

30.2 The Web 978
30.2.1 HyperText Transfer Protocol 979
30.2.2 HyperText Markup Language 981
30.2.3 Uniform Resource Locators 982
30.2.4 Static and Dynamic Web Pages 982

30.2.5 Web Services 984
30.2.6 Requirements for Web–DBMS Integration 985
30.2.7 Advantages and Disadvantages of the Web–DBMS Approach 986
30.2.8 Approaches to Integrating the Web and DBMSs 990

30.3 Scripting Languages 991
30.3.1 JavaScript and JScript 991
30.3.2 VBScript 992
30.3.3 Perl and PHP 993

30.4 Common Gateway Interface (CGI) 993
30.4.1 Passing Information to a CGI Script 995
30.4.2 Advantages and Disadvantages of CGI 997

30.5 HTTP Cookies 998

30.6 Extending the Web Server 999
30.6.1 Comparison of CGI and API 1000

30.7 Java 1000
30.7.1 JDBC 1004
30.7.2 SQLJ 1010
30.7.3 Comparison of JDBC and SQLJ 1010
30.7.4 Container-Managed Persistence (CMP) 1011
30.7.5 Java Data Objects (JDO) 1015
30.7.6 JPA (Java Persistence API) 1022
30.7.7 Java Servlets 1030
30.7.8 JavaServer Pages 1030
30.7.9 Java Web Services 1031

30.8 Microsoft's Web Platform 1032
30.8.1 Universal Data Access 1034
30.8.2 Active Server Pages and ActiveX Data Objects 1035
30.8.3 Remote Data Services 1036
30.8.4 Comparison of ASP and JSP 1039
30.8.5 Microsoft .NET 1039
30.8.6 Microsoft Web Services 1044

30.9 Oracle Internet Platform 1044
30.9.1 Oracle Application Server (OracleAS) 1045

Chapter Summary *1051*
Review Questions *1052*
Exercises *1053*

Chapter 31 Semistructured Data and XML **1055**

31.1 Semistructured Data 1056
31.1.1 Object Exchange Model (OEM) 1058
31.1.2 Lore and Lorel 1059

31.2 Introduction to XML 1063
31.2.1 Overview of XML 1066
31.2.2 Document Type Definitions (DTDs) 1068
31.3 XML-Related Technologies 1071
31.3.1 DOM and SAX Interfaces 1072
31.3.2 Namespaces 1073
31.3.3 XSL and XSLT 1073
31.3.4 XPath (XML Path Language) 1074
31.3.5 XPointer (XML Pointer Language) 1075
31.3.6 XLink (XML Linking Language) 1076
31.3.7 XHTML 1076
31.3.8 Simple Object Access Protocol (SOAP) 1077
31.3.9 Web Services Description Language (WSDL) 1077
31.3.10 Universal Discovery, Description and Integration (UDDI) 1078
31.4 XML Schema 1081
31.4.1 Resource Description Framework (RDF) 1087
31.5 XML Query Languages 1091
31.5.1 Extending Lore and Lorel to Handle XML 1092
31.5.2 XML Query Working Group 1093
31.5.3 XQuery—A Query Language for XML 1094
31.5.4 XML Information Set 1104
31.5.5 XQuery 1.0 and XPath 2.0 Data Model (XDM) 1105
31.5.6 XQuery Update Facility 1.0 1111
31.5.7 Formal Semantics 1113
31.6 XML and Databases 1121
31.6.1 Storing XML in Databases 1121
31.6.2 XML and SQL 1124
31.6.3 Native XML Databases 1135
31.7 XML in Oracle 1136
Chapter Summary *1139*
Review Questions *1141*
Exercises *1142*

Part 9 Business Intelligence 1143

Chapter 32 Data Warehousing Concepts 1145

32.1 Introduction to Data Warehousing 1146
32.1.1 The Evolution of Data Warehousing 1146
32.1.2 Data Warehousing Concepts 1147
32.1.3 Benefits of Data Warehousing 1148
32.1.4 Comparison of OLTP Systems and Data Warehousing 1148
32.1.5 Problems of Data Warehousing 1150
32.1.6 Real-Time Data Warehouse 1152

32.2 Data Warehouse Architecture 1153
32.2.1 Operational Data 1153
32.2.2 Operational Data Store 1153
32.2.3 ETL Manager 1154
32.2.4 Warehouse Manager 1154
32.2.5 Query Manager 1155
32.2.6 Detailed Data 1155
32.2.7 Lightly and Highly Summarized Data 1155
32.2.8 Archive/Backup Data 1155
32.2.9 Metadata 1156
32.2.10 End-User Access Tools 1156

32.3 Data Warehousing Tools and Technologies 1157
32.3.1 Extraction, Transformation, and Loading (ETL) 1158
32.3.2 Data Warehouse DBMS 1159
32.3.3 Data Warehouse Metadata 1162
32.3.4 Administration and Management Tools 1164

32.4 Data Mart 1164
32.4.1 Reasons for Creating a Data Mart 1165

32.5 Data Warehousing Using Oracle 1165
32.5.1 New Warehouse Features in Oracle 10*g*/11*g* 1168

Chapter Summary *1169*
Review Questions *1170*
Exercise *1171*

Chapter 33 Data Warehousing Design 1173

33.1 Designing a Data Warehouse Database 1174

33.2 Data Warehouse Development Methodologies 1174

33.3 Kimball's Business Dimensional Lifecycle 1176

33.4 Dimensionality Modeling 1177
33.4.1 Comparison of DM and ER models 1180

33.5 The Dimensional Modeling Stage of Kimball's Business Dimensional Lifecycle 1181
33.5.1 Create a High-Level Dimensional Model (Phase I) 1181
Step 1: Select Business Process 1181
Step 2: Declare Grain 1183
Step 3: Choose Dimensions 1183
Step 4: Identify Facts 1185
33.5.2 Identify All Dimension Attributes for the Dimensional Model (Phase II) 1186

33.6 Data Warehouse Development Issues 1189

33.7 Data Warehousing Design Using Oracle 1190
33.7.1 Oracle Warehouse Builder Components 1190

33.7.2 Using Oracle Warehouse Builder 1191
33.7.3 New Warehouse Builder Features in Oracle 10*g*/11*g* 1195

Chapter Summary *1196*
Review Questions *1197*
Exercises *1198*

Chapter 34 OLAP **1199**

34.1 Online Analytical Processing 1200
34.1.1 OLAP Benchmarks 1201

34.2 OLAP Applications 1201

34.3 Multidimensional Data Model 1203
34.3.1 Alternative Multidimensional Data Representations 1203
34.3.2 Dimensional Hierarchy 1205
34.3.3 Multidimensional Operations 1207
34.3.4 Multidimensional Schemas 1207

34.4 OLAP Tools 1207
34.4.1 Codd's Rules for OLAP Tools 1208
34.4.2 OLAP Server—Implementation Issues 1209
34.4.3 Categories of OLAP Server 1210

34.5 OLAP Extensions to the SQL Standard 1214
34.5.1 Extended Grouping Capabilities 1214
34.5.2 Elememtary OLAP Operators 1219

34.6 Oracle OLAP 1221
34.6.1 Oracle OLAP Environment 1221
34.6.2 Platform for Business Intelligence Applications 1222
34.6.3 Oracle Database 1222
34.6.4 Oracle OLAP 1224
34.6.5 Performance 1225
34.6.6 System Management 1226
34.6.7 System Requirements 1226
34.6.8 New OLAP Features in Oracle 11*g* 1226

Chapter Summary *1226*
Review Questions *1227*
Exercises *1227*

Chapter 35 Data Mining **1229**

35.1 Data Mining 1230

35.2 Data Mining Techniques 1230
35.2.1 Predictive Modeling 1232
35.2.2 Database Segmentation 1233

35.2.3 Link Analysis 1234
35.2.4 Deviation Detection 1235
35.3 The Data Mining Process 1236
35.3.1 The CRISP-DM Model 1236
35.4 Data Mining Tools 1237
35.5 Data Mining and Data Warehousing 1238
35.6 Oracle Data Mining (ODM) 1239
35.6.1 Data Mining Capabilities 1239
35.6.2 Enabling Data Mining Applications 1239
35.6.3 Predictions and Insights 1240
35.6.4 Oracle Data Mining Environment 1240
35.6.5 New Data Mining Features in Oracle 11*g* 1241
Chapter Summary *1241*
Review Questions *1242*
Exercises *1242*

Appendices **1243**

A Users' Requirements Specification for *DreamHome* Case Study **A-1**

A.1 Branch User Views of *DreamHome* A-1
A.1.1 Data Requirements A-1
A.1.2 Transaction Requirements (Sample) A-3
A.2 Staff User Views of *DreamHome* A-4
A.2.1 Data Requirements A-4
A.2.2 Transaction Requirements (Sample) A-5

B Other Case Studies **B-1**

B.1 The *University Accommodation Office* Case Study B-1
B.1.1 Data Requirements B-1
B.1.2 Query Transactions (Sample) B-3
B.2 The *EasyDrive School of Motoring* Case Study B-4
B.2.1 Data Requirements B-4
B.2.2 Query Transactions (Sample) B-5
B.3 The *Wellmeadows Hospital* Case Study B-5
B.3.1 Data Requirements B-5
B.3.2 Transaction Requirements (Sample) B-12

C Alternative ER Modeling Notations **C-1**

C.1 ER Modeling Using the Chen Notation C-1
C.2 ER Modeling Using the Crow's Feet Notation C-1

D Summary of the Database Design Methodology for Relational Databases D-1

Step 1: Build Conceptual Data Model D-1
Step 2: Build Logical Data Model D-2
Step 3: Translate Logical Data Model for Target DBMS D-5
Step 4: Design File Organizations and Indexes D-5
Step 5: Design User Views D-5
Step 6: Design Security Mechanisms D-5
Step 7: Consider the Introduction of Controlled Redundancy D-6
Step 8: Monitor and Tune the Operational System D-6

E Introduction to Pyrrho: A Lightweight RDBMS E-1

E.1 Pyrrho Features E-2
E.2 Download and Install Pyrrho E-2
E.3 Getting Started E-3
E.4 The Connection String E-3
E.5 Pyrrho's Security Model E-4
E.6 Pyrrho SQL Syntax E-4

F File Organizations and Indexes (Online) F-1

G When Is a DBMS Relational? (Online) G-1

H Commercial DBMSs: Access and Oracle (Online) H-1

I Programmatic SQL (Online) I-1

J Estimating Disk Space Requirements (Online) J-1

K Introduction to Object-Orientation (Online) K-1

L Example Web Scripts (Online) L-1

References R-1
Further Reading FR-1
Index In-1