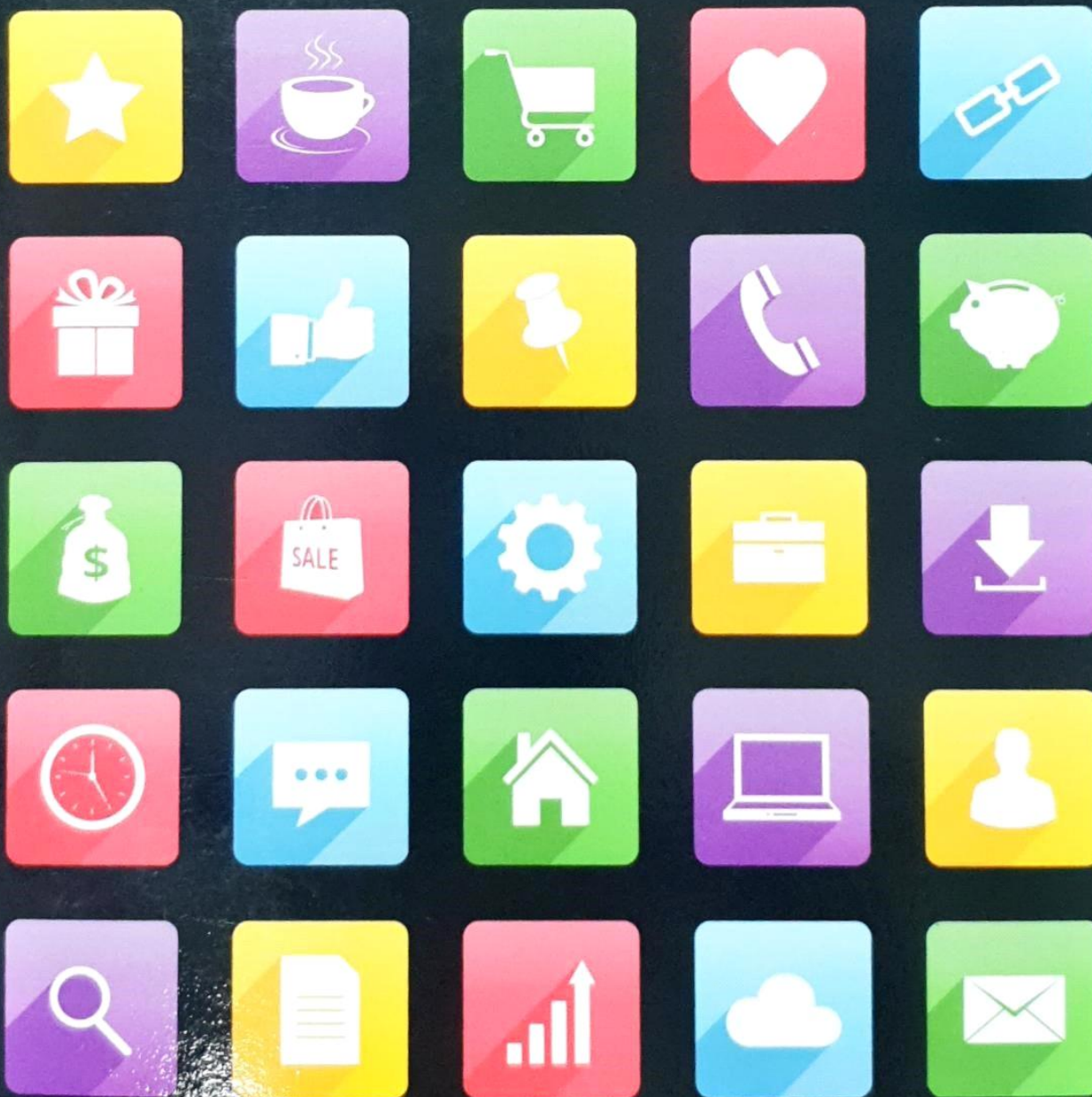


DATABASE SYSTEMS

A Practical Approach to
Design, Implementation, and Management

SIXTH EDITION



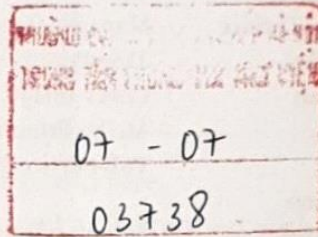
THOMAS CONNOLLY
CAROLYN BEGG

DATABASE SYSTEMS

A Practical Approach to Design, Implementation, and Management

SIXTH EDITION

JOHN F. KORTH, CAROLYN E. BEGG
UNIVERSITY OF THE WEST OF SCOTLAND



DATABASE SYSTEMS

A Practical Approach to Design, Implementation, and Management

SIXTH EDITION

THOMAS M. CONNOLLY | CAROLYN E. BEGG

UNIVERSITY OF THE WEST OF SCOTLAND

PEARSON

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

Brief Contents

Preface

xxxv

Part I Background

Chapter 1	Introduction to Databases	3
Chapter 2	Database Environment	35
Chapter 3	Database Architectures and the Web	57

Part 2 The Relational Model and Languages 99

Chapter 4	The Relational Model	101
Chapter 5	Relational Algebra and Relational Calculus	119
Chapter 6	SQL: Data Manipulation	143
Chapter 7	SQL: Data Definition	185
Chapter 8	Advanced SQL	223
Chapter 9	Object-Relational DBMSs	243

Part 3 Database Analysis and Design 295

Chapter 10	Database System Development Lifecycle	297
Chapter 11	Database Analysis and the <i>DreamHome</i> Case Study	327
Chapter 12	Entity-Relationship Modeling	357
Chapter 13	Enhanced Entity-Relationship Modeling	385
Chapter 14	Normalization	403
Chapter 15	Advanced Normalization	433

Part 4 Methodology 453

Chapter 16	Methodology—Conceptual Database Design	455
Chapter 17	Methodology—Logical Database Design for the Relational Model	479

Chapter 18	Methodology—Physical Database Design for Relational Databases	513
Chapter 19	Methodology—Monitoring and Tuning the Operational System	537
Part 5	Selected Database Issues	557
Chapter 20	Security and Administration	559
Chapter 21	Professional, Legal, and Ethical Issues in Data Management	593
Chapter 22	Transaction Management	619
Chapter 23	Query Processing	679
Part 6	Distributed DBMSs and Replication	735
Chapter 24	Distributed DBMSs—Concepts and Design	737
Chapter 25	Distributed DBMSs—Advanced Concepts	783
Chapter 26	Replication and Mobile Databases	827
Part 7	Object DBMSs	891
Chapter 27	Object-Oriented DBMSs—Concepts and Design	893
Chapter 28	Object-Oriented DBMSs—Standards and Systems	947
Part 8	The Web and DBMSs	997
Chapter 29	Web Technology and DBMSs	999
Chapter 30	Semistructured Data and XML	1081
Part 9	Business Intelligence	1173
Chapter 31	Data Warehousing Concepts	1175
Chapter 32	Data Warehousing Design	1209
Chapter 33	OLAP	1237
Chapter 34	Data Mining	1267

Appendices**I281**

A Users' Requirements Specification for <i>DreamHome</i> Case Study	A-1
B Other Case Studies	B-1
C Alternative ER Modeling Notations	C-1
D Summary of the Database Design Methodology for Relational Databases	D-1
E Introduction to Pyrrho: A Lightweight RDBMS	E-1
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-Oriented Concepts (Online)	K-1
L Example Web Scripts (Online)	L-1
M Query-By-Example (QBE) (Online)	M-1
N Third Generation Manifestos (Online)	N-1
O Postgres—An Early ORDBMS (Online)	O-1

References**R-1****Further Reading****FR-1****Index****IN-1**

Contents

Preface

xxxxv

Part I Background

Chapter 1 Introduction to Databases

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
	<i>Chapter Summary</i>	31
	<i>Review Questions</i>	32
	<i>Exercises</i>	32

Chapter 2 Database Environment

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	56
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)	71
3.3 Distributed DBMSs	72
3.4 Data Warehousing	75
3.5 Cloud Computing	77
3.5.1 Benefits and Risks of Cloud Computing	79
3.5.2 Cloud-based database solutions	82
3.6 Components of a DBMS	86
3.7 Oracle Architecture	89
3.7.1 Oracle's Logical Database Structure	89
3.7.2 Oracle's Physical Database Structure	92
Chapter Summary	96
Review Questions	97
Exercises	97
Part 2 The Relational Model and Languages	99
Chapter 4 The Relational Model	101
4.1 Brief History of the Relational Model	102
4.2 Terminology	104
4.2.1 Relational Data Structure	104

4.2.2 Mathematical Relations	107
4.2.3 Database Relations	108
4.2.4 Properties of Relations	108
4.2.5 Relational Keys	110
4.2.6 Representing Relational Database Schemas	111
4.3 Integrity Constraints	113
4.3.1 Nulls	113
4.3.2 Entity Integrity	114
4.3.3 Referential Integrity	114
4.3.4 General Constraints	115
4.4 Views	115
4.4.1 Terminology	115
4.4.2 Purpose of Views	116
4.4.3 Updating Views	117
Chapter Summary	117
Review Questions	118
Exercises	118
Chapter 5 Relational Algebra and Relational Calculus	119
5.1 The Relational Algebra	120
5.1.1 Unary Operations	120
5.1.2 Set Operations	123
5.1.3 Join Operations	126
5.1.4 Division Operation	129
5.1.5 Aggregation and Grouping Operations	130
5.1.6 Summary of the Relational Algebra Operations	132
5.2 The Relational Calculus	133
5.2.1 Tuple Relational Calculus	133
5.2.2 Domain Relational Calculus	136
5.3 Other Languages	138
Chapter Summary	139
Review Questions	139
Exercises	140
Chapter 6 SQL: Data Manipulation	143
6.1 Introduction to SQL	144
6.1.1 Objectives of SQL	144
6.1.2 History of SQL	145
6.1.3 Importance of SQL	147
6.1.4 Terminology	147
6.2 Writing SQL Commands	147

6.3	Data Manipulation	148
6.3.1	Simple Queries	149
6.3.2	Sorting Results (ORDER BY Clause)	157
6.3.3	Using the SQL Aggregate Functions	159
6.3.4	Grouping Results (GROUP BY Clause)	161
6.3.5	Subqueries	164
6.3.6	ANY and ALL	166
6.3.7	Multi-table Queries	168
6.3.8	EXISTS and NOT EXISTS	174
6.3.9	Combining Result Tables (UNION, INTERSECT, EXCEPT)	175
6.3.10	Database Updates	177
	<i>Chapter Summary</i>	181
	<i>Review Questions</i>	182
	<i>Exercises</i>	182
Chapter 7	SQL: Data Definition	185
7.1	The ISO SQL Data Types	186
7.1.1	SQL Identifiers	186
7.1.2	SQL Scalar Data Types	187
7.2	Integrity Enhancement Feature	192
7.2.1	Required Data	192
7.2.2	Domain Constraints	192
7.2.3	Entity Integrity	193
7.2.4	Referential Integrity	194
7.2.5	General Constraints	195
7.3	Data Definition	196
7.3.1	Creating a Database	196
7.3.2	Creating a Table (CREATE TABLE)	197
7.3.3	Changing a Table Definition (ALTER TABLE)	200
7.3.4	Removing a Table (DROP TABLE)	201
7.3.5	Creating an Index (CREATE INDEX)	202
7.3.6	Removing an Index (DROP INDEX)	202
7.4	Views	203
7.4.1	Creating a View (CREATE VIEW)	203
7.4.2	Removing a View (DROP VIEW)	205
7.4.3	View Resolution	206
7.4.4	Restrictions on Views	207
7.4.5	View Updatability	207
7.4.6	WITH CHECK OPTION	208
7.4.7	Advantages and Disadvantages of Views	210
7.4.8	View Materialization	212

7.5	Transactions	213
7.5.1	Immediate and Deferred Integrity Constraints	214
7.6	Discretionary Access Control	214
7.6.1	Granting Privileges to Other Users (GRANT)	216
7.6.2	Revoking Privileges from Users (REVOKE)	217
	<i>Chapter Summary</i>	219
	<i>Review Questions</i>	220
	<i>Exercises</i>	220
Chapter 8	Advanced SQL	223
8.1	The SQL Programming Language	224
8.1.1	Declarations	224
8.1.2	Assignments	225
8.1.3	Control Statements	226
8.1.4	Exceptions in PL/SQL	228
8.1.5	Cursors in PL/SQL	229
8.2	Subprograms, Stored Procedures, Functions, and Packages	232
8.3	Triggers	233
8.4	Recursion	239
	<i>Chapter Summary</i>	240
	<i>Review Questions</i>	241
	<i>Exercises</i>	241
Chapter 9	Object-Relational DBMSs	243
9.1	Advanced Database Applications	244
9.2	Weaknesses of RDBMSs	249
9.3	Storing Objects in a Relational Database	254
9.3.1	Mapping Classes to Relations	255
9.3.2	Accessing Objects in the Relational Database	256
9.4	Introduction to Object-Relational Database Systems	257
9.5	SQL:2011	260
9.5.1	Row Types	261
9.5.2	User-Defined Types	262
9.5.3	Subtypes and Supertypes	265
9.5.4	User-Defined Routines	266
9.5.5	Polymorphism	269
9.5.6	Reference Types and Object Identity	270
9.5.7	Creating Tables	270
9.5.8	Querying Data	273

9.5.9 Collection Types	275
9.5.10 Typed Views	278
9.5.11 Persistent Stored Modules	279
9.5.12 Triggers	279
9.5.13 Large Objects	282
9.5.14 Recursion	283
9.6 Object-Oriented Extensions in Oracle	283
9.6.1 User-Defined Data Types	284
9.6.2 Manipulating Object Tables	289
9.6.3 Object Views	290
9.6.4 Privileges	291
Chapter Summary	292
Review Questions	292
Exercises	293

Part 3 Database Analysis and Design 295

Chapter 10 Database System Development Lifecycle	297
10.1 The Information Systems Lifecycle	298
10.2 The Database System Development Lifecycle	299
10.3 Database Planning	299
10.4 System Definition	302
10.4.1 User Views	302
10.5 Requirements Collection and Analysis	302
10.5.1 Centralized Approach	304
10.5.2 View Integration Approach	304
10.6 Database Design	306
10.6.1 Approaches to Database Design	307
10.6.2 Data Modeling	307
10.6.3 Phases of Database Design	308
10.7 DBMS Selection	311
10.7.1 Selecting the DBMS	311
10.8 Application Design	315
10.8.1 Transaction Design	316
10.8.2 User Interface Design Guidelines	317
10.9 Prototyping	319
10.10 Implementation	319
10.11 Data Conversion and Loading	320

10.12	Testing	320
10.13	Operational Maintenance	321
10.14	CASE Tools	322
	<i>Chapter Summary</i>	324
	<i>Review Questions</i>	325
	<i>Exercises</i>	326
Chapter 11	Database Analysis and the DreamHome Case Study	327
11.1	When Are Fact-Finding Techniques Used?	328
11.2	What Facts Are Collected?	329
11.3	Fact-Finding Techniques	330
11.3.1	Examining Documentation	330
11.3.2	Interviewing	330
11.3.3	Observing the Enterprise in Operation	331
11.3.4	Research	332
11.3.5	Questionnaires	332
11.4	Using Fact-Finding Techniques: A Worked -Example	333
11.4.1	The <i>DreamHome</i> Case Study—An Overview of the Current System	334
11.4.2	The <i>DreamHome</i> Case Study—Database Planning	338
11.4.3	The <i>DreamHome</i> Case Study—System Definition	344
11.4.4	The <i>DreamHome</i> Case Study—Requirements Collection and Analysis	345
11.4.5	The <i>DreamHome</i> Case Study—Database Design	353
	<i>Chapter Summary</i>	354
	<i>Review Questions</i>	354
	<i>Exercises</i>	354
Chapter 12	Entity–Relationship Modeling	357
12.1	Entity Types	358
12.2	Relationship Types	360
12.2.1	Degree of Relationship Type	362
12.2.2	Recursive Relationship	364
12.3	Attributes	365
12.3.1	Simple and Composite Attributes	365
12.3.2	Single-valued and Multi-valued Attributes	366
12.3.3	Derived Attributes	366
12.3.4	Keys	367
12.4	Strong and Weak Entity Types	369
12.5	Attributes on Relationships	370

12.6	Structural Constraints	371
12.6.1	One-to-One (1:1) Relationships	372
12.6.2	One-to-Many (1:*) Relationships	373
12.6.3	Many-to-Many (*:*) Relationships	374
12.6.4	Multiplicity for Complex Relationships	375
12.6.5	Cardinality and Participation Constraints	376
12.7	Problems with ER Models	378
12.7.1	Fan Traps	378
12.7.2	Chasm Traps	380
	<i>Chapter Summary</i>	382
	<i>Review Questions</i>	382
	<i>Exercises</i>	383
Chapter 13	Enhanced Entity–Relationship Modeling	385
13.1	Specialization/Generalization	386
13.1.1	Superclasses and Subclasses	386
13.1.2	Superclass/Subclass Relationships	387
13.1.3	Attribute Inheritance	388
13.1.4	Specialization Process	388
13.1.5	Generalization Process	389
13.1.6	Constraints on Specialization/Generalization	392
13.1.7	Worked Example of using Specialization/ Generalization to Model the Branch View of the <i>DreamHome</i> Case Study	393
13.2	Aggregation	397
13.3	Composition	398
	<i>Chapter Summary</i>	399
	<i>Review Questions</i>	400
	<i>Exercises</i>	400
Chapter 14	Normalization	403
14.1	The Purpose of Normalization	404
14.2	How Normalization Supports Database Design	405
14.3	Data Redundancy and Update Anomalies	406
14.3.1	Insertion Anomalies	407
14.3.2	Deletion Anomalies	407
14.3.3	Modification Anomalies	408
14.4	Functional Dependencies	408
14.4.1	Characteristics of Functional Dependencies	408
14.4.2	Identifying Functional Dependencies	412
14.4.3	Identifying the Primary Key for a Relation Using Functional Dependencies	415
14.5	The Process of Normalization	416

14.6	First Normal Form (1NF)	418
14.7	Second Normal Form (2NF)	422
14.8	Third Normal Form (3NF)	423
14.9	General Definitions of 2NF and 3NF	425
	<i>Chapter Summary</i>	427
	<i>Review Questions</i>	427
	<i>Exercises</i>	428
Chapter 15	Advanced Normalization	433
15.1	More on Functional Dependencies	434
15.1.1	Inference Rules for Functional Dependencies	434
15.1.2	Minimal Sets of Functional Dependencies	436
15.2	Boyce–Codd Normal Form (BCNF)	437
15.2.1	Definition of BCNF	437
15.3	Review of Normalization Up to BCNF	440
15.4	Fourth Normal Form (4NF)	445
15.4.1	Multi-Valued Dependency	446
15.4.2	Definition of Fourth Normal Form	447
15.5	Fifth Normal Form (5NF)	447
15.5.1	Lossless-Join Dependency	448
15.5.2	Definition of Fifth Normal Form	448
	<i>Chapter Summary</i>	450
	<i>Review Questions</i>	450
	<i>Exercises</i>	451
Part 4	Methodology	453
Chapter 16	Methodology—Conceptual Database Design	455
16.1	Introduction to the Database Design Methodology	456
16.1.1	What Is a Design Methodology?	456
16.1.2	Conceptual, Logical, and Physical Database Design	457
16.1.3	Critical Success Factors in Database Design	457
16.2	Overview of the Database Design Methodology	458
16.3	Conceptual Database Design Methodology	460
	Step 1: Build Conceptual Data Model	460
	<i>Chapter Summary</i>	476
	<i>Review Questions</i>	476
	<i>Exercises</i>	477

Chapter 17	Methodology—Logical Database Design for the Relational Model	479
17.1	Logical Database Design Methodology for the Relational Model	480
	Step 2: Build Logical Data Model	480
	<i>Chapter Summary</i>	508
	<i>Review Questions</i>	509
	<i>Exercises</i>	509
Chapter 18	Methodology—Physical Database Design for Relational Databases	513
18.1	Comparison of Logical and Physical Database Design	514
18.2	Overview of the Physical Database Design Methodology	515
18.3	The Physical Database Design Methodology for Relational Databases	516
	Step 3: Translate Logical Data Model for Target DBMS	516
	Step 4: Design File Organizations and Indexes	521
	Step 5: Design User Views	534
	Step 6: Design Security Mechanisms	534
	<i>Chapter Summary</i>	535
	<i>Review Questions</i>	536
	<i>Exercises</i>	536
Chapter 19	Methodology—Monitoring and Tuning the Operational System	537
19.1	Denormalizing and Introducing Controlled Redundancy	537
	Step 7: Consider the Introduction of Controlled Redundancy	537
19.2	Monitoring the System to Improve Performance	550
	Step 8: Monitor and Tune the Operational System	550
	<i>Chapter Summary</i>	554
	<i>Review Questions</i>	555
	<i>Exercises</i>	555
Part 5	Selected Database Issues	557
Chapter 20	Security and Administration	559
20.1	Database Security	560
	20.1.1 Threats	561

20.2	Countermeasures—Computer-Based Controls	563
20.2.1	Authorization	564
20.2.2	Access Controls	565
20.2.3	Views	568
20.2.4	Backup and Recovery	568
20.2.5	Integrity	569
20.2.6	Encryption	569
20.2.7	RAID (Redundant Array of Independent Disks)	570
20.3	Security in Microsoft Office Access DBMS	573
20.4	Security in Oracle DBMS	575
20.5	DBMSs and Web Security	579
20.5.1	Proxy Servers	580
20.5.2	Firewalls	580
20.5.3	Message Digest Algorithms and Digital Signatures	581
20.5.4	Digital Certificates	581
20.5.5	Kerberos	582
20.5.6	Secure Sockets Layer and Secure HTTP	582
20.5.7	Secure Electronic Transactions and Secure Transaction Technology	583
20.5.8	Java Security	584
20.5.9	ActiveX Security	586
20.6	Data Administration and Database Administration	586
20.6.1	Data Administration	587
20.6.2	Database Administration	588
20.6.3	Comparison of Data and Database Administration	588
	<i>Chapter Summary</i>	589
	<i>Review Questions</i>	590
	<i>Exercises</i>	590
Chapter 21	Professional, Legal, and Ethical Issues in Data Management	593
21.1	Defining Legal and Ethical Issues in IT	594
21.1.1	Defining Ethics in the Context of IT	594
21.1.2	The Difference Between Ethical and Legal Behavior	595
21.1.3	Ethical Behavior in IT	596
21.2	Legislation and Its Impact on the IT Function	597
21.2.1	Securities and Exchange Commission (SEC) Regulation National Market System (NMS)	597
21.2.2	The Sarbanes-Oxley Act, COBIT, and COSO	598
21.2.3	The Health Insurance Portability and Accountability Act	601
21.2.4	The European Union (EU) Directive on Data Protection of 1995	602
21.2.5	The United Kingdom's Data Protection Act of 1998	603

21.2.6 Access to Information Laws	604
21.2.7 International Banking—Basel II Accords	606
21.3 Establishing a Culture of Legal and Ethical Data Stewardship	607
21.3.1 Developing an Organization-Wide Policy for Legal and Ethical Behavior	607
21.3.2 Professional Organizations and Codes of Ethics	608
21.3.3 Developing an Organization-Wide Policy for Legal and Ethical Behavior for <i>DreamHome</i>	611
21.4 Intellectual Property	612
21.4.1 Patent	613
21.4.2 Copyright	613
21.4.3 Trademark	614
21.4.4 Intellectual Property Rights Issues for Software	614
21.4.5 Intellectual Property Rights Issues for Data	616
Chapter Summary	616
Review Questions	617
Exercises	618
Chapter 22 Transaction Management	619
22.1 Transaction Support	620
22.1.1 Properties of Transactions	623
22.1.2 Database Architecture	623
22.2 Concurrency Control	624
22.2.1 The Need for Concurrency Control	624
22.2.2 Serializability and Recoverability	627
22.2.3 Locking Methods	635
22.2.4 Deadlock	641
22.2.5 Timestamping Methods	644
22.2.6 Multiversion Timestamp Ordering	647
22.2.7 Optimistic Techniques	648
22.2.8 Granularity of Data Items	649
22.3 Database Recovery	652
22.3.1 The Need for Recovery	652
22.3.2 Transactions and Recovery	653
22.3.3 Recovery Facilities	656
22.3.4 Recovery Techniques	659
22.3.5 Recovery in a Distributed DBMS	661
22.4 Advanced Transaction Models	661
22.4.1 Nested Transaction Model	663
22.4.2 Sagas	664
22.4.3 Multilevel Transaction Model	665
22.4.4 Dynamic Restructuring	666
22.4.5 Workflow Models	667

22.5	Concurrency Control and Recovery in Oracle	668
22.5.1	Oracle's Isolation Levels	669
22.5.2	Multiversion Read Consistency	669
22.5.3	Deadlock Detection	671
22.5.4	Backup and Recovery	671
	<i>Chapter Summary</i>	674
	<i>Review Questions</i>	675
	<i>Exercises</i>	676
Chapter 23	Query Processing	679
23.1	Overview of Query Processing	681
23.2	Query Decomposition	684
23.3	Heuristical Approach to Query Optimization	688
23.3.1	Transformation Rules for the Relational Algebra Operations	688
23.3.2	Heuristical Processing Strategies	693
23.4	Cost Estimation for the Relational Algebra Operations	694
23.4.1	Database Statistics	694
23.4.2	Selection Operation ($S = \sigma_p(R)$)	695
23.4.3	Join Operation ($T = (R \bowtie_f S)$)	702
23.4.4	Projection Operation ($S = \Pi_{A_1, A_2, \dots, A_m}(R)$)	709
23.4.5	The Relational Algebra Set Operations ($T = R \cup S, T = R \cap S, T = R - S$)	711
23.5	Enumeration of Alternative Execution Strategies	712
23.5.1	Pipelining	713
23.5.2	Linear Trees	713
23.5.3	Physical Operators and Execution Strategies	714
23.5.4	Reducing the Search Space	716
23.5.5	Enumerating Left-Deep Trees	717
23.5.6	Semantic Query Optimization	718
23.5.7	Alternative Approaches to Query Optimization	719
23.5.8	Distributed Query Optimization	720
23.6	Query Processing and Optimization	720
23.6.1	New Index Types	723
23.7	Query Optimization in Oracle	724
23.7.1	Rule-Based and Cost-Based Optimization	724
23.7.2	Histograms	728
23.7.3	Viewing the Execution Plan	730
	<i>Chapter Summary</i>	731
	<i>Review Questions</i>	732
	<i>Exercises</i>	733

Part 6	Distributed DBMSs and Replication	735
Chapter 24	Distributed DBMSs—Concepts and Design	737
24.1	Introduction	738
24.1.1	Concepts	739
24.1.2	Advantages and Disadvantages of DDBMSs	743
24.1.3	Homogeneous and Heterogeneous DDBMSs	746
24.2	Overview of Networking	749
24.3	Functions and Architectures of a DDBMS	753
24.3.1	Functions of a DDBMS	753
24.3.2	Reference Architecture for a DDBMS	753
24.3.3	Reference Architecture for a Federated MDBS	755
24.3.4	Component Architecture for a DDBMS	756
24.4	Distributed Relational Database Design	757
24.4.1	Data Allocation	758
24.4.2	Fragmentation	759
24.5	Transparencies in a DDBMS	768
24.5.1	Distribution Transparency	768
24.5.2	Transaction Transparency	771
24.5.3	Performance Transparency	774
24.5.4	DBMS Transparency	776
24.5.5	Summary of Transparencies in a DDBMS	776
24.6	Date's Twelve Rules for a DDBMS	777
	<i>Chapter Summary</i>	779
	<i>Review Questions</i>	780
	<i>Exercises</i>	780
Chapter 25	Distributed DBMSs—Advanced Concepts	783
25.1	Distributed Transaction Management	784
25.2	Distributed Concurrency Control	785
25.2.1	Objectives	785
25.2.2	Distributed Serializability	786
25.2.3	Locking Protocols	786
25.3	Distributed Deadlock Management	789
25.4	Distributed Database Recovery	792
25.4.1	Failures in a Distributed Environment	793
25.4.2	How Failures Affect Recovery	794
25.4.3	Two-Phase Commit (2PC)	794
25.4.4	Three-Phase Commit (3PC)	801
25.4.5	Network Partitioning	804
25.5	The X/Open Distributed Transaction Processing Model	806
25.6	Distributed Query Optimization	808

25.6.1 Data Localization	810
25.6.2 Distributed Joins	813
25.6.3 Global Optimization	814
25.7 Distribution in Oracle	818
25.7.1 Oracle's DDBMS Functionality	818
Chapter Summary	824
Review Questions	824
Exercises	825
Chapter 26 Replication and Mobile Databases	827
26.1 Introduction to Data Replication	828
26.1.1 Applications of Replication	829
26.1.2 Replication Model	830
26.1.3 Functional Model of Replication Protocols	831
26.1.4 Consistency	832
26.2 Replication Architecture	832
26.2.1 Kernel-Based Replication	832
26.2.2 Middleware-Based Replication	833
26.2.3 Processing of Updates	834
26.2.4 Propagation of Updates	836
26.2.5 Update Location (Data Ownership)	836
26.2.6 Termination Protocols	840
26.3 Replication Schemes	840
26.3.1 Eager Primary Copy	841
26.3.2 Lazy Primary Copy	846
26.3.3 Eager Update Anywhere	850
26.3.4 Lazy Update Anywhere	851
26.3.5 Update Anywhere with Uniform Total Order Broadcast	855
26.3.6 SI and Uniform Total Order Broadcast Replication	859
26.4 Introduction to Mobile Databases	865
26.4.1 Mobile DBMSs	867
26.4.2 Issues with Mobile DBMSs	868
26.5 Oracle Replication	881
26.5.1 Oracle's Replication Functionality	881
Chapter Summary	888
Review Questions	889
Exercises	889
Part 7 Object DBMSs	891
Chapter 27 Object-Oriented DBMSs—Concepts and Design	893
27.1 Next-Generation Database Systems	895

27.2	Introduction to OODBMSs	897
27.2.1	Definition of Object-Oriented DBMSs	897
27.2.2	Functional Data Models	898
27.2.3	Persistent Programming Languages	903
27.2.4	Alternative Strategies for Developing an OODBMS	905
27.3	Persistence in OODBMSs	906
27.3.1	Pointer Swizzling Techniques	908
27.3.2	Accessing an Object	911
27.3.3	Persistence Schemes	913
27.3.4	Orthogonal Persistence	914
27.4	Issues in OODBMSs	916
27.4.1	Transactions	916
27.4.2	Versions	917
27.4.3	Schema Evolution	918
27.4.4	Architecture	921
27.4.5	Benchmarking	923
27.5	Advantages and Disadvantages of OODBMSs	926
27.5.1	Advantages	926
27.5.2	Disadvantages	928
27.6	Comparison of ORDBMS and OODBMS	930
27.7	Object-Oriented Database Design	931
27.7.1	Comparison of Object-Oriented Data Modeling and Conceptual Data Modeling	931
27.7.2	Relationships and Referential Integrity	932
27.7.3	Behavioral Design	934
27.8	Object-Oriented Analysis and Design with UML	936
27.8.1	UML Diagrams	937
27.8.2	Usage of UML in the Methodology for Database Design	942
	<i>Chapter Summary</i>	944
	<i>Review Questions</i>	945
	<i>Exercises</i>	945
Chapter 28	Object-Oriented DBMSs—Standards and Systems	947
28.1	Object Management Group	948
28.1.1	Background	948
28.1.2	The Common Object Request Broker Architecture	951
28.1.3	Other OMG Specifications	956
28.1.4	Model-Driven Architecture	959
28.2	Object Data Standard ODMG 3.0, 1999	959
28.2.1	Object Data Management Group	961
28.2.2	The Object Model	962

28.2.3 The Object Definition Language	970
28.2.4 The Object Query Language	973
28.2.5 Other Parts of the ODMG Standard	979
28.2.6 Mapping the Conceptual Design to a Logical (Object-Oriented) Design	982
28.3 ObjectStore	983
28.3.1 Architecture	983
28.3.2 Building an ObjectStore Application	986
28.3.3 Data Definition in ObjectStore	987
28.3.4 Data Manipulation in ObjectStore	991
Chapter Summary	994
Review Questions	995
Exercises	995

Part 8 The Web and DBMSs 997

Chapter 29 Web Technology and DBMSs 999

29.1 Introduction to the Internet and the Web	1000
29.1.1 Intranets and Extranets	1002
29.1.2 e-Commerce and e-Business	1003
29.2 The Web	1004
29.2.1 HyperText Transfer Protocol	1005
29.2.2 HyperText Markup Language	1007
29.2.3 Uniform Resource Locators	1009
29.2.4 Static and Dynamic Web Pages	1010
29.2.5 Web Services	1010
29.2.6 Requirements for Web-DBMS Integration	1011
29.2.7 Advantages and Disadvantages of the Web-DBMS Approach	1012
29.2.8 Approaches to Integrating the Web and DBMSs	1016
29.3 Scripting Languages	1017
29.3.1 JavaScript and JScript	1017
29.3.2 VBScript	1018
29.3.3 Perl and PHP	1019
29.4 Common Gateway Interface (CGI)	1019
29.4.1 Passing Information to a CGI Script	1021
29.4.2 Advantages and Disadvantages of CGI	1023
29.5 HTTP Cookies	1024
29.6 Extending the Web Server	1025
29.6.1 Comparison of CGI and API	1026
29.7 Java	1026
29.7.1 JDBC	1030
29.7.2 SQLJ	1036

29.7.3 Comparison of JDBC and SQLJ	1036
29.7.4 Container-Managed Persistence (CMP)	1037
29.7.5 Java Data Objects (JDO)	1041
29.7.6 JPA (Java Persistence API)	1048
29.7.7 Java Servlets	1056
29.7.8 JavaServer Pages	1056
29.7.9 Java Web Services	1057
29.8 Microsoft's Web Platform	1059
29.8.1 Universal Data Access	1060
29.8.2 Active Server Pages and ActiveX Data Objects	1061
29.8.3 Remote Data Services	1062
29.8.4 Comparison of ASP and JSP	1065
29.8.5 Microsoft .NET	1065
29.8.6 Microsoft Web Services	1070
29.9 Oracle Internet Platform	1071
29.9.1 Oracle WebLogic Server	1072
29.9.2 Oracle Metadata Repository	1073
29.9.3 Oracle Identity Management	1073
29.9.4 Oracle Portal	1074
29.9.5 Oracle WebCenter	1074
29.9.6 Oracle Business Intelligence (BI) Discoverer	1074
29.9.7 Oracle SOA (Service-Oriented Architecture) Suite	1075
<i>Chapter Summary</i>	1078
<i>Review Questions</i>	1079
<i>Exercises</i>	1079
 Chapter 30 Semistructured Data and XML	 1081
30.1 Semistructured Data	1082
30.1.1 Object Exchange Model (OEM)	1084
30.1.2 Lore and Lorel	1085
30.2 Introduction to XML	1089
30.2.1 Overview of XML	1092
30.2.2 Document Type Definitions (DTDs)	1094
30.3 XML-Related Technologies	1097
30.3.1 DOM and SAX Interfaces	1098
30.3.2 Namespaces	1099
30.3.3 XSL and XSLT	1099
30.3.4 XPath (XML Path Language)	1100
30.3.5 XPointer (XML Pointer Language)	1101
30.3.6 XLink (XML Linking Language)	1102
30.3.7 XHTML	1102
30.3.8 Simple Object Access Protocol (SOAP)	1103
30.3.9 Web Services Description Language (WSDL)	1104

30.3.10 Universal Discovery, Description, and Integration (UDDI)	1104
30.3.11 JSON (JavaScript Object Notation)	1106
30.4 XML Schema	1108
30.4.1 Resource Description Framework (RDF)	1114
30.5 XML Query Languages	1118
30.5.1 Extending Lore and Lorel to Handle XML	1119
30.5.2 XML Query Working Group	1120
30.5.3 XQuery—A Query Language for XML	1121
30.5.4 XML Information Set	1131
30.5.5 XQuery 1.0 and XPath 2.0 Data Model (XDM)	1132
30.5.6 XQuery Update Facility 1.0	1138
30.5.7 Formal Semantics	1140
30.6 XML and Databases	1148
30.6.1 Storing XML in Databases	1148
30.6.2 XML and SQL	1151
30.6.3 Native XML Databases	1165
30.7 XML in Oracle	1166
Chapter Summary	1169
Review Questions	1171
Exercises	1172

Part 9 Business Intelligence

1173

Chapter 31 Data Warehousing Concepts

1175

31.1 Introduction to Data Warehousing	1176
31.1.1 The Evolution of Data Warehousing	1176
31.1.2 Data Warehousing Concepts	1177
31.1.3 Benefits of Data Warehousing	1178
31.1.4 Comparison of OLTP Systems and Data Warehousing	1178
31.1.5 Problems of Data Warehousing	1180
31.1.6 Real-Time Data Warehouse	1182
31.2 Data Warehouse Architecture	1183
31.2.1 Operational Data	1183
31.2.2 Operational Data Store	1183
31.2.3 ETL Manager	1184
31.2.4 Warehouse Manager	1184
31.2.5 Query Manager	1185
31.2.6 Detailed Data	1185
31.2.7 Lightly and Highly Summarized Data	1185
31.2.8 Archive/Backup Data	1185
31.2.9 Metadata	1186
31.2.10 End-User Access Tools	1186

31.3	Data Warehousing Tools and Technologies	1187
31.3.1	Extraction, Transformation, and Loading (ETL)	1188
31.3.2	Data Warehouse DBMS	1189
31.3.3	Data Warehouse Metadata	1192
31.3.4	Administration and Management Tools	1194
31.4	Data Mart	1194
31.4.1	Reasons for Creating a Data Mart	1195
31.5	Data Warehousing and Temporal Databases	1195
31.5.1	Temporal Extensions to the SQL Standard	1198
31.6	Data Warehousing Using Oracle	1200
31.6.1	Warehouse Features in Oracle 11g	1203
31.6.2	Oracle Support for Temporal Data	1204
	<i>Chapter Summary</i>	1205
	<i>Review Questions</i>	1206
	<i>Exercises</i>	1207

Chapter 32 Data Warehousing Design 1209

32.1	Designing a Data Warehouse Database	1210
32.2	Data Warehouse Development Methodologies	1210
32.3	Kimball's Business Dimensional Lifecycle	1212
32.4	Dimensionality Modeling	1213
32.4.1	Comparison of DM and ER models	1216
32.5	The Dimensional Modeling Stage of Kimball's Business Dimensional Lifecycle	1217
32.5.1	Create a High-Level Dimensional Model (Phase I)	1217
32.5.2	Identify All Dimension Attributes for the Dimensional Model (Phase II)	1222
32.6	Data Warehouse Development Issues	1225
32.7	Data Warehousing Design Using Oracle	1226
32.7.1	Oracle Warehouse Builder Components	1226
32.7.2	Using Oracle Warehouse Builder	1227
32.7.3	Warehouse Builder Features in Oracle 11g	1231
	<i>Chapter Summary</i>	1232
	<i>Review Questions</i>	1233
	<i>Exercises</i>	1234

Chapter 33 OLAP 1237

33.1	Online Analytical Processing	1238
33.1.1	OLAP Benchmarks	1239

33.2	OLAP Applications	1239
33.3	Multidimensional Data Model	1241
33.3.1	Alternative Multidimensional Data Representations	1241
33.3.2	Dimensional Hierarchy	1243
33.3.3	Multidimensional Operations	1245
33.3.4	Multidimensional Schemas	1245
33.4	OLAP Tools	1245
33.4.1	Codd's Rules for OLAP Tools	1246
33.4.2	OLAP Server—Implementation Issues	1247
33.4.3	Categories of OLAP Servers	1248
33.5	OLAP Extensions to the SQL Standard	1252
33.5.1	Extended Grouping Capabilities	1252
33.5.2	Elementary OLAP Operators	1257
33.6	Oracle OLAP	1259
33.6.1	Oracle OLAP Environment	1259
33.6.2	Platform for Business Intelligence Applications	1260
33.6.3	Oracle Database	1260
33.6.4	Oracle OLAP	1262
33.6.5	Performance	1263
33.6.6	System Management	1264
33.6.7	System Requirements	1264
33.6.8	OLAP Features in Oracle 11g	1264
	<i>Chapter Summary</i>	1265
	<i>Review Questions</i>	1265
	<i>Exercises</i>	1265
Chapter 34	Data Mining	1267
34.1	Data Mining	1268
34.2	Data Mining Techniques	1268
34.2.1	Predictive Modeling	1270
34.2.2	Database Segmentation	1271
34.2.3	Link Analysis	1272
34.2.4	Deviation Detection	1273
34.3	The Data Mining Process	1274
34.3.1	The CRISP-DM Model	1274
34.4	Data Mining Tools	1275
34.5	Data Mining and Data Warehousing	1276
34.6	Oracle Data Mining (ODM)	1277
34.6.1	Data Mining Capabilities	1277