

THOMAS CONNOLLY & CAROLYN BEGG

DATABASE SYSTEMS

A Practical Approach to Design,
Implementation, and Management



www.booksites.net/connbegg


ADDISON
WESLEY

FOURTH EDITION



Contents

Preface

xxxiii

Part 1	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	22
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	24
1.6	Advantages and Disadvantages of DBMSs	26
	<i>Chapter Summary</i>	31
	<i>Review Questions</i>	32
	<i>Exercises</i>	32
Chapter 2	Database Environment	33
2.1	The Three-Level ANSI-SPARC Architecture	34
2.1.1	External Level	35

2.1.2 Conceptual Level	36
2.1.3 Internal Level	36
2.1.4 Schemas, Mappings, and Instances	37
2.1.5 Data Independence	38
2.2 Database Languages	39
2.2.1 The Data Definition Language (DDL)	40
2.2.2 The Data Manipulation Language (DML)	40
2.2.3 Fourth-Generation Languages (4GLs)	42
2.3 Data Models and Conceptual Modeling	43
2.3.1 Object-Based Data Models	44
2.3.2 Record-Based Data Models	45
2.3.3 Physical Data Models	47
2.3.4 Conceptual Modeling	47
2.4 Functions of a DBMS	48
2.5 Components of a DBMS	53
2.6 Multi-User DBMS Architectures	56
2.6.1 Teleprocessing	56
2.6.2 File-Server Architectures	56
2.6.3 Traditional Two-Tier Client-Server Architecture	57
2.6.4 Three-Tier Client-Server Architecture	60
2.6.5 Transaction Processing Monitors	62
<i>Chapter Summary</i>	64
<i>Review Questions</i>	65
<i>Exercises</i>	65
Part 2 The Relational Model and Languages	67
Chapter 3 The Relational Model	69
3.1 Brief History of the Relational Model	70
3.2 Terminology	71
3.2.1 Relational Data Structure	72
3.2.2 Mathematical Relations	75
3.2.3 Database Relations	76
3.2.4 Properties of Relations	77
3.2.5 Relational Keys	78
3.2.6 Representing Relational Database Schemas	79
3.3 Integrity Constraints	81
3.3.1 Nulls	81
3.3.2 Entity Integrity	82
3.3.3 Referential Integrity	83
3.3.4 General Constraints	83
3.4 Views	83
3.4.1 Terminology	84

3.4.2 Purpose of Views	84
3.4.3 Updating Views	85
<i>Chapter Summary</i>	86
<i>Review Questions</i>	87
<i>Exercises</i>	87
Chapter 4 Relational Algebra and Relational Calculus	88
4.1 The Relational Algebra	89
4.1.1 Unary Operations	89
4.1.2 Set Operations	92
4.1.3 Join Operations	95
4.1.4 Division Operation	99
4.1.5 Aggregation and Grouping Operations	100
4.1.6 Summary of the Relational Algebra Operations	102
4.2 The Relational Calculus	103
4.2.1 Tuple Relational Calculus	103
4.2.2 Domain Relational Calculus	107
4.3 Other Languages	109
<i>Chapter Summary</i>	110
<i>Review Questions</i>	110
<i>Exercises</i>	111
Chapter 5 SQL: Data Manipulation	112
5.1 Introduction to SQL	113
5.1.1 Objectives of SQL	113
5.1.2 History of SQL	114
5.1.3 Importance of SQL	116
5.1.4 Terminology	116
5.2 Writing SQL Commands	116
5.3 Data Manipulation	117
5.3.1 Simple Queries	118
5.3.2 Sorting Results (ORDER BY Clause)	127
5.3.3 Using the SQL Aggregate Functions	129
5.3.4 Grouping Results (GROUP BY Clause)	131
5.3.5 Subqueries	134
5.3.6 ANY and ALL	138
5.3.7 Multi-Table Queries	139
5.3.8 EXISTS and NOT EXISTS	146
5.3.9 Combining Result Tables (UNION, INTERSECT, EXCEPT)	147
5.3.10 Database Updates	149
<i>Chapter Summary</i>	154
<i>Review Questions</i>	155
<i>Exercises</i>	155

Chapter 6	SQL: Data Definition	157
6.1	The ISO SQL Data Types	158
6.1.1	SQL Identifiers	158
6.1.2	SQL Scalar Data Types	159
6.1.3	Exact Numeric Data	160
6.2	Integrity Enhancement Feature	164
6.2.1	Required Data	164
6.2.2	Domain Constraints	164
6.2.3	Entity Integrity	166
6.2.4	Referential Integrity	166
6.2.5	General Constraints	167
6.3	Data Definition	168
6.3.1	Creating a Database	168
6.3.2	Creating a Table (CREATE TABLE)	169
6.3.3	Changing a Table Definition (ALTER TABLE)	173
6.3.4	Removing a Table (DROP TABLE)	174
6.3.5	Creating an Index (CREATE INDEX)	175
6.3.6	Removing an Index (DROP INDEX)	176
6.4	Views	176
6.4.1	Creating a View (CREATE VIEW)	177
6.4.2	Removing a View (DROP VIEW)	179
6.4.3	View Resolution	180
6.4.4	Restrictions on Views	181
6.4.5	View Updatability	181
6.4.6	WITH CHECK OPTION	183
6.4.7	Advantages and Disadvantages of Views	184
6.4.8	View Materialization	186
6.5	Transactions	187
6.5.1	Immediate and Deferred Integrity Constraints	189
6.6	Discretionary Access Control	189
6.6.1	Granting Privileges to Other Users (GRANT)	191
6.6.2	Revoking Privileges from Users (REVOKE)	192
	<i>Chapter Summary</i>	194
	<i>Review Questions</i>	195
	<i>Exercises</i>	195
Chapter 7	Query-By-Example	198
7.1	Introduction to Microsoft Office Access Queries	199
7.2	Building Select Queries Using QBE	201
7.2.1	Specifying Criteria	202
7.2.2	Creating Multi-Table Queries	204
7.2.3	Calculating Totals	207
7.3	Using Advanced Queries	208

7.3.1	Parameter Query	208
7.3.2	Crosstab Query	209
7.3.3	Find Duplicates Query	212
7.3.4	Find Unmatched Query	214
7.3.5	Autolookup Query	215
7.4	Changing the Content of Tables Using Action Queries	215
7.4.1	Make-Table Action Query	215
7.4.2	Delete Action Query	217
7.4.3	Update Action Query	217
7.4.4	Append Action Query	221
	<i>Exercises</i>	224
Chapter 8	Commercial RDBMSs: Office Access and Oracle	225
8.1	Microsoft Office Access 2003	226
8.1.1	Objects	226
8.1.2	Microsoft Office Access Architecture	227
8.1.3	Table Definition	228
8.1.4	Relationships and Referential Integrity Definition	233
8.1.5	General Constraint Definition	234
8.1.6	Forms	236
8.1.7	Reports	238
8.1.8	Macros	239
8.1.9	Object Dependencies	242
8.2	Oracle9i	242
8.2.1	Objects	244
8.2.2	Oracle Architecture	245
8.2.3	Table Definition	252
8.2.4	General Constraint Definition	255
8.2.5	PL/SQL	255
8.2.6	Subprograms, Stored Procedures, Functions, and Packages	261
8.2.7	Triggers	263
8.2.8	Oracle Internet Developer Suite	267
8.2.9	Other Oracle Functionality	271
8.2.10	Oracle10g	271
	<i>Chapter Summary</i>	276
	<i>Review Questions</i>	277
Part 3	Database Analysis and Design Techniques	279
Chapter 9	Database Planning, Design, and Administration	281
9.1	The Information Systems Lifecycle	282
9.2	The Database System Development Lifecycle	283

9.3	Database Planning	285
9.4	System Definition	286
9.4.1	User Views	287
9.5	Requirements Collection and Analysis	288
9.5.1	Centralized Approach	289
9.5.2	View Integration Approach	289
9.6	Database Design	291
9.6.1	Approaches to Database Design	291
9.6.2	Data Modeling	292
9.6.3	Phases of Database Design	293
9.7	DBMS Selection	295
9.7.1	Selecting the DBMS	296
9.8	Application Design	299
9.8.1	Transaction Design	300
9.8.2	User Interface Design Guidelines	301
9.9	Prototyping	303
9.10	Implementation	304
9.11	Data Conversion and Loading	305
9.12	Testing	305
9.13	Operational Maintenance	306
9.14	CASE Tools	307
9.15	Data Administration and Database Administration	309
9.15.1	Data Administration	309
9.15.2	Database Administration	309
9.15.3	Comparison of Data and Database Administration	311
	<i>Chapter Summary</i>	311
	<i>Review Questions</i>	313
	<i>Exercises</i>	313
Chapter 10	Fact-Finding Techniques	314
10.1	When Are Fact-Finding Techniques Used?	315
10.2	What Facts Are Collected?	316
10.3	Fact-Finding Techniques	317
10.3.1	Examining Documentation	317
10.3.2	Interviewing	317
10.3.3	Observing the Enterprise in Operation	319
10.3.4	Research	319
10.3.5	Questionnaires	320
10.4	Using Fact-Finding Techniques – A Worked Example	321
10.4.1	The <i>DreamHome</i> Case Study – An Overview	321
10.4.2	The <i>DreamHome</i> Case Study – Database Planning	326

10.4.3 The <i>DreamHome</i> Case Study – System Definition	331
10.4.4 The <i>DreamHome</i> Case Study – Requirements Collection and Analysis	332
10.4.5 The <i>DreamHome</i> Case Study – Database Design	340
<i>Chapter Summary</i>	340
<i>Review Questions</i>	341
<i>Exercises</i>	341
Chapter 11 Entity–Relationship Modeling	342
11.1 Entity Types	343
11.2 Relationship Types	346
11.2.1 Degree of Relationship Type	347
11.2.2 Recursive Relationship	349
11.3 Attributes	350
11.3.1 Simple and Composite Attributes	351
11.3.2 Single-Valued and Multi-Valued Attributes	351
11.3.3 Derived Attributes	352
11.3.4 Keys	352
11.4 Strong and Weak Entity Types	354
11.5 Attributes on Relationships	355
11.6 Structural Constraints	356
11.6.1 One-to-One (1:1) Relationships	357
11.6.2 One-to-Many (1:*) Relationships	358
11.6.3 Many-to-Many (*:*) Relationships	359
11.6.4 Multiplicity for Complex Relationships	361
11.6.5 Cardinality and Participation Constraints	362
11.7 Problems with ER Models	364
11.7.1 Fan Traps	364
11.7.2 Chasm Traps	365
<i>Chapter Summary</i>	368
<i>Review Questions</i>	369
<i>Exercises</i>	369
Chapter 12 Enhanced Entity–Relationship Modeling	371
12.1 Specialization/Generalization	372
12.1.1 Superclasses and Subclasses	372
12.1.2 Superclass/Subclass Relationships	373
12.1.3 Attribute Inheritance	374
12.1.4 Specialization Process	374
12.1.5 Generalization Process	375
12.1.6 Constraints on Specialization/Generalization	378

12.1.7	Worked Example of using Specialization/Generalization to Model the Branch View of <i>DreamHome</i> Case Study	379
12.2	Aggregation	383
12.3	Composition	384
	<i>Chapter Summary</i>	385
	<i>Review Questions</i>	386
	<i>Exercises</i>	386
Chapter 13	Normalization	387
13.1	The Purpose of Normalization	388
13.2	How Normalization Supports Database Design	389
13.3	Data Redundancy and Update Anomalies	390
	13.3.1 Insertion Anomalies	391
	13.3.2 Deletion Anomalies	392
	13.3.3 Modification Anomalies	392
13.4	Functional Dependencies	392
	13.4.1 Characteristics of Functional Dependencies	393
	13.4.2 Identifying Functional Dependencies	397
	13.4.3 Identifying the Primary Key for a Relation using Functional Dependencies	399
13.5	The Process of Normalization	401
13.6	First Normal Form (1NF)	403
13.7	Second Normal Form (2NF)	407
13.8	Third Normal Form (3NF)	408
13.9	General Definitions of 2NF and 3NF	411
	<i>Chapter Summary</i>	412
	<i>Review Questions</i>	413
	<i>Exercises</i>	413
Chapter 14	Advanced Normalization	415
14.1	More on Functional Dependencies	416
	14.1.1 Inference Rules for Functional Dependencies	416
	14.1.2 Minimal Sets of Functional Dependencies	418
14.2	Boyce–Codd Normal Form (BCNF)	419
	14.2.1 Definition of Boyce–Codd Normal Form	419
14.3	Review of Normalization up to BCNF	422
14.4	Fourth Normal Form (4NF)	428
	14.4.1 Multi-Valued Dependency	428
	14.4.2 Definition of Fourth Normal Form	430
14.5	Fifth Normal Form (5NF)	430

14.5.1 Lossless-Join Dependency	430
14.5.2 Definition of Fifth Normal Form	431
<i>Chapter Summary</i>	433
<i>Review Questions</i>	433
<i>Exercises</i>	433

Part 4 Methodology

435

Chapter 15	Methodology – Conceptual Database Design	437
15.1	Introduction to the Database Design Methodology	438
15.1.1	What is a Design Methodology?	438
15.1.2	Conceptual, Logical, and Physical Database Design	439
15.1.3	Critical Success Factors in Database Design	440
15.2	Overview of the Database Design Methodology	440
15.3	Conceptual Database Design Methodology	442
	Step 1 Build Conceptual Data Model	442
	<i>Chapter Summary</i>	458
	<i>Review Questions</i>	459
	<i>Exercises</i>	460
Chapter 16	Methodology – Logical Database Design for the Relational Model	461
16.1	Logical Database Design Methodology for the Relational Model	462
	Step 2 Build and Validate Logical Data Model	462
	<i>Chapter Summary</i>	490
	<i>Review Questions</i>	491
	<i>Exercises</i>	492
Chapter 17	Methodology – Physical Database Design for Relational Databases	494
17.1	Comparison of Logical and Physical Database Design	495
17.2	Overview of Physical Database Design Methodology	496
17.3	The Physical Database Design Methodology for Relational Databases	497
	Step 3 Translate Logical Data Model for Target DBMS	497
	Step 4 Design File Organizations and Indexes	501
	Step 5 Design User Views	515
	Step 6 Design Security Mechanisms	516
	<i>Chapter Summary</i>	517
	<i>Review Questions</i>	517
	<i>Exercises</i>	518

- Chapter 18	Methodology – Monitoring and Tuning the Operational System	519
18.1	Denormalizing and Introducing Controlled Redundancy	519
	Step 7 Consider the Introduction of Controlled Redundancy	519
18.2	Monitoring the System to Improve Performance	532
	Step 8 Monitor and Tune the Operational System	532
	<i>Chapter Summary</i>	537
	<i>Review Questions</i>	537
	<i>Exercise</i>	537

Part 5 Selected Database Issues

539

Chapter 19	Security	541
19.1	Database Security	542
	19.1.1 Threats	543
19.2	Countermeasures – Computer-Based Controls	545
	19.2.1 Authorization	546
	19.2.2 Access Controls	547
	19.2.3 Views	550
	19.2.4 Backup and Recovery	550
	19.2.5 Integrity	551
	19.2.6 Encryption	551
	19.2.7 RAID (Redundant Array of Independent Disks)	552
19.3	Security in Microsoft Office Access DBMS	555
19.4	Security in Oracle DBMS	558
19.5	DBMSs and Web Security	562
	19.5.1 Proxy Servers	563
	19.5.2 Firewalls	563
	19.5.3 Message Digest Algorithms and Digital Signatures	564
	19.5.4 Digital Certificates	564
	19.5.5 Kerberos	565
	19.5.6 Secure Sockets Layer and Secure HTTP	565
	19.5.7 Secure Electronic Transactions and Secure Transaction Technology	566
	19.5.8 Java Security	566
	19.5.9 ActiveX Security	569
	<i>Chapter Summary</i>	570
	<i>Review Questions</i>	571
	<i>Exercises</i>	571
Chapter 20	Transaction Management	572
20.1	Transaction Support	573
	20.1.1 Properties of Transactions	575
	20.1.2 Database Architecture	576

20.2	Concurrency Control	577
20.2.1	The Need for Concurrency Control	577
20.2.2	Serializability and Recoverability	580
20.2.3	Locking Methods	587
20.2.4	Deadlock	594
20.2.5	Timestamping Methods	597
20.2.6	Multiversion Timestamp Ordering	600
20.2.7	Optimistic Techniques	601
20.2.8	Granularity of Data Items	602
20.3	Database Recovery	605
20.3.1	The Need for Recovery	606
20.3.2	Transactions and Recovery	607
20.3.3	Recovery Facilities	609
20.3.4	Recovery Techniques	612
20.3.5	Recovery in a Distributed DBMS	615
20.4	Advanced Transaction Models	615
20.4.1	Nested Transaction Model	616
20.4.2	Sagas	618
20.4.3	Multilevel Transaction Model	619
20.4.4	Dynamic Restructuring	620
20.4.5	Workflow Models	621
20.5	Concurrency Control and Recovery in Oracle	622
20.5.1	Oracle's Isolation Levels	623
20.5.2	Multiversion Read Consistency	623
20.5.3	Deadlock Detection	625
20.5.4	Backup and Recovery	625
	<i>Chapter Summary</i>	626
	<i>Review Questions</i>	627
	<i>Exercises</i>	628
Chapter 21	Query Processing	630
21.1	Overview of Query Processing	631
21.2	Query Decomposition	635
21.3	Heuristical Approach to Query Optimization	639
21.3.1	Transformation Rules for the Relational Algebra Operations	640
21.3.2	Heuristical Processing Strategies	645
21.4	Cost Estimation for the Relational Algebra Operations	646
21.4.1	Database Statistics	646
21.4.2	Selection Operation	647
21.4.3	Join Operation	654
21.4.4	Projection Operation	662
21.4.5	The Relational Algebra Set Operations	664
21.5	Enumeration of Alternative Execution Strategies	665

21.5.1	Pipelining	665
21.5.2	Linear Trees	666
21.5.3	Physical Operators and Execution Strategies	667
21.5.4	Reducing the Search Space	668
21.5.5	Enumerating Left-Deep Trees	669
21.5.6	Semantic Query Optimization	671
21.5.7	Alternative Approaches to Query Optimization	672
21.5.8	Distributed Query Optimization	672
21.6	Query Optimization in Oracle	673
21.6.1	Rule-Based and Cost-Based Optimization	673
21.6.2	Histograms	677
21.6.3	Viewing the Execution Plan	678
	<i>Chapter Summary</i>	680
	<i>Review Questions</i>	681
	<i>Exercises</i>	681

Part 6 Distributed DBMSs and Replication 685

Chapter 22	Distributed DBMSs – Concepts and Design	687
22.1	Introduction	688
22.1.1	Concepts	689
22.1.2	Advantages and Disadvantages of DDBMSs	693
22.1.3	Homogeneous and Heterogeneous DDBMSs	697
22.2	Overview of Networking	699
22.3	Functions and Architectures of a DDBMS	703
22.3.1	Functions of a DDBMS	703
22.3.2	Reference Architecture for a DDBMS	704
22.3.3	Reference Architecture for a Federated MDBS	705
22.3.4	Component Architecture for a DDBMS	706
22.4	Distributed Relational Database Design	708
22.4.1	Data Allocation	709
22.4.2	Fragmentation	710
22.5	Transparencies in a DDBMS	719
22.5.1	Distribution Transparency	719
22.5.2	Transaction Transparency	722
22.5.3	Performance Transparency	725
22.5.4	DBMS Transparency	728
22.5.5	Summary of Transparencies in a DDBMS	728
22.6	Date's Twelve Rules for a DDBMS	729
	<i>Chapter Summary</i>	731
	<i>Review Questions</i>	732
	<i>Exercises</i>	732

Chapter 23	Distributed DBMSs – Advanced Concepts	734
23.1	Distributed Transaction Management	735
23.2	Distributed Concurrency Control	736
23.2.1	Objectives	736
23.2.2	Distributed Serializability	737
23.2.3	Locking Protocols	738
23.2.4	Timestamp Protocols	740
23.3	Distributed Deadlock Management	741
23.4	Distributed Database Recovery	744
23.4.1	Failures in a Distributed Environment	744
23.4.2	How Failures Affect Recovery	745
23.4.3	Two-Phase Commit (2PC)	746
23.4.4	Three-Phase Commit (3PC)	752
23.4.5	Network Partitioning	756
23.5	The X/Open Distributed Transaction Processing Model	758
23.6	Distributed Query Optimization	761
23.6.1	Data Localization	762
23.6.2	Distributed Joins	766
23.6.3	Global Optimization	767
23.7	Distribution in Oracle	772
23.7.1	Oracle's DDBMS Functionality	772
	<i>Chapter Summary</i>	777
	<i>Review Questions</i>	778
	<i>Exercises</i>	778
Chapter 24	Replication and Mobile Databases	780
24.1	Introduction to Database Replication	781
24.2	Benefits of Database Replication	781
24.3	Applications of Replication	783
24.4	Basic Components of Database Replication	783
24.5	Database Replication Environments	784
24.5.1	Synchronous Versus Asynchronous Replication	784
24.5.2	Data Ownership	784
24.6	Replication Servers	788
24.6.1	Replication Server Functionality	788
24.6.2	Implementation Issues	789
24.7	Introduction to Mobile Databases	792
24.7.1	Mobile DBMSs	794
24.8	Oracle Replication	794
24.8.1	Oracle's Replication Functionality	794

<i>Chapter Summary</i>	799
<i>Review Questions</i>	800
<i>Exercises</i>	800

Part 7 Object DBMSs 801

Chapter 25	Introduction to Object DBMSs	803
25.1	Advanced Database Applications	804
25.2	Weaknesses of RDBMSs	809
25.3	Object-Oriented Concepts	814
25.3.1	Abstraction, Encapsulation, and Information Hiding	814
25.3.2	Objects and Attributes	815
25.3.3	Object Identity	816
25.3.4	Methods and Messages	818
25.3.5	Classes	819
25.3.6	Subclasses, Superclasses, and Inheritance	820
25.3.7	Overriding and Overloading	822
25.3.8	Polymorphism and Dynamic Binding	823
25.3.9	Complex Objects	824
25.4	Storing Objects in a Relational Database	825
25.4.1	Mapping Classes to Relations	826
25.4.2	Accessing Objects in the Relational Database	827
25.5	Next-Generation Database Systems	828
25.6	Object-Oriented Database Design	830
25.6.1	Comparison of Object-Oriented Data Modeling and Conceptual Data Modeling	830
25.6.2	Relationships and Referential Integrity	831
25.6.3	Behavioral Design	834
25.7	Object-Oriented Analysis and Design with UML	836
25.7.1	UML Diagrams	837
25.7.2	Usage of UML in the Methodology for Database Design	842
	<i>Chapter Summary</i>	844
	<i>Review Questions</i>	845
	<i>Exercises</i>	846
Chapter 26	Object-Oriented DBMSs – Concepts	847
26.1	Introduction to Object-Oriented Data Models and OODBMSs	849
26.1.1	Definition of Object-Oriented DBMSs	849
26.1.2	Functional Data Models	850
26.1.3	Persistent Programming Languages	854
26.1.4	<i>The Object-Oriented Database System Manifesto</i>	857
26.1.5	Alternative Strategies for Developing an OODBMS	859

26.2	OODBMS Perspectives	860
26.2.1	Pointer Swizzling Techniques	862
26.2.2	Accessing an Object	865
26.3	Persistence	867
26.3.1	Persistence Schemes	868
26.3.2	Orthogonal Persistence	869
26.4	Issues in OODBMSs	871
26.4.1	Transactions	871
26.4.2	Versions	872
26.4.3	Schema Evolution	873
26.4.4	Architecture	876
26.4.5	Benchmarking	878
26.5	Advantages and Disadvantages of OODBMSs	881
26.5.1	Advantages	881
26.5.2	Disadvantages	883
	<i>Chapter Summary</i>	885
	<i>Review Questions</i>	886
	<i>Exercises</i>	887
Chapter 27	Object-Oriented DBMSs – Standards and Systems	888
27.1	Object Management Group	889
27.1.1	Background	889
27.1.2	The Common Object Request Broker Architecture	891
27.1.3	Other OMG Specifications	894
27.1.4	Model-Driven Architecture	897
27.2	Object Data Standard ODMG 3.0, 1999	897
27.2.1	Object Data Management Group	897
27.2.2	The Object Model	900
27.2.3	The Object Definition Language	908
27.2.4	The Object Query Language	911
27.2.5	Other Parts of the ODMG Standard	917
27.2.6	Mapping the Conceptual Design to a Logical (Object-Oriented) Design	920
27.3	ObjectStore	921
27.3.1	Architecture	921
27.3.2	Building an ObjectStore Application	924
27.3.3	Data Definition in ObjectStore	926
27.3.4	Data Manipulation in ObjectStore	929
	<i>Chapter Summary</i>	932
	<i>Review Questions</i>	934
	<i>Exercises</i>	934

Chapter 28	Object-Relational DBMSs	935
28.1	Introduction to Object-Relational Database Systems	936
28.2	The Third-Generation Database Manifestos	939
28.2.1	The <i>Third-Generation Database System Manifesto</i>	940
28.2.2	The <i>Third Manifesto</i>	940
28.3	Postgres – An Early ORDBMS	943
28.3.1	Objectives of Postgres	943
28.3.2	Abstract Data Types	943
28.3.3	Relations and Inheritance	944
28.3.4	Object Identity	946
28.4	SQL:1999 and SQL:2003	946
28.4.1	Row Types	947
28.4.2	User-Defined Types	948
28.4.3	Subtypes and Supertypes	951
28.4.4	User-Defined Routines	953
28.4.5	Polymorphism	955
28.4.6	Reference Types and Object Identity	956
28.4.7	Creating Tables	957
28.4.8	Querying Data	960
28.4.9	Collection Types	961
28.4.10	Typed Views	965
28.4.11	Persistent Stored Modules	966
28.4.12	Triggers	967
28.4.13	Large Objects	971
28.4.14	Recursion	972
28.5	Query Processing and Optimization	974
28.5.1	New Index Types	977
28.6	Object-Oriented Extensions in Oracle	978
28.6.1	User-Defined Data Types	978
28.6.2	Manipulating Object Tables	984
28.6.3	Object Views	985
28.6.4	Privileges	986
28.7	Comparison of ORDBMS and OODBMS	986
	<i>Chapter Summary</i>	988
	<i>Review Questions</i>	988
	<i>Exercises</i>	989

Part 8 Web and DBMSs 991

Chapter 29	Web Technology and DBMSs	993
29.1	Introduction to the Internet and Web	994

29.1.1 Intranets and Extranets	996
29.1.2 e-Commerce and e-Business	997
29.2 The Web	998
29.2.1 HyperText Transfer Protocol	999
29.2.2 HyperText Markup Language	1001
29.2.3 Uniform Resource Locators	1002
29.2.4 Static and Dynamic Web Pages	1004
29.2.5 Web Services	1004
29.2.6 Requirements for Web-DBMS Integration	1005
29.2.7 Advantages and Disadvantages of the Web-DBMS Approach	1006
29.2.8 Approaches to Integrating the Web and DBMSs	1011
29.3 Scripting Languages	1011
29.3.1 JavaScript and JScript	1012
29.3.2 VBScript	1012
29.3.3 Perl and PHP	1013
29.4 Common Gateway Interface	1014
29.4.1 Passing Information to a CGI Script	1016
29.4.2 Advantages and Disadvantages of CGI	1018
29.5 HTTP Cookies	1019
29.6 Extending the Web Server	1020
29.6.1 Comparison of CGI and API	1021
29.7 Java	1021
29.7.1 JDBC	1025
29.7.2 SQLJ	1030
29.7.3 Comparison of JDBC and SQLJ	1030
29.7.4 Container-Managed Persistence (CMP)	1031
29.7.5 Java Data Objects (JDO)	1035
29.7.6 Java Servlets	1040
29.7.7 JavaServer Pages	1041
29.7.8 Java Web Services	1042
29.8 Microsoft's Web Platform	1043
29.8.1 Universal Data Access	1045
29.8.2 Active Server Pages and ActiveX Data Objects	1046
29.8.3 Remote Data Services	1049
29.8.4 Comparison of ASP and JSP	1049
29.8.5 Microsoft .NET	1050
29.8.6 Microsoft Web Services	1054
29.8.7 Microsoft Office Access and Web Page Generation	1054
29.9 Oracle Internet Platform	1055
29.9.1 Oracle Application Server (OracleAS)	1056
Chapter Summary	1062
Review Questions	1063
Exercises	1064

Chapter 30	Semistructured Data and XML	1065
30.1	Semistructured Data	1066
30.1.1	Object Exchange Model (OEM)	1068
30.1.2	Lore and Lorel	1069
30.2	Introduction to XML	1073
30.2.1	Overview of XML	1076
30.2.2	Document Type Definitions (DTDs)	1078
30.3	XML-Related Technologies	1082
30.3.1	DOM and SAX Interfaces	1082
30.3.2	Namespaces	1083
30.3.3	XSL and XSLT	1084
30.3.4	XPath (XML Path Language)	1085
30.3.5	XPointer (XML Pointer Language)	1085
30.3.6	XLink (XML Linking Language)	1086
30.3.7	XHTML	1087
30.3.8	Simple Object Access Protocol (SOAP)	1087
30.3.9	Web Services Description Language (WSDL)	1088
30.3.10	Universal Discovery, Description and Integration (UDDI)	1088
30.4	XML Schema	1091
30.4.1	Resource Description Framework (RDF)	1098
30.5	XML Query Languages	1100
30.5.1	Extending Lore and Lorel to Handle XML	1100
30.5.2	XML Query Working Group	1101
30.5.3	XQuery – A Query Language for XML	1103
30.5.4	XML Information Set	1114
30.5.5	XQuery 1.0 and XPath 2.0 Data Model	1115
30.5.6	Formal Semantics	1121
30.6	XML and Databases	1128
30.6.1	Storing XML in Databases	1129
30.6.2	XML and SQL	1132
30.6.3	Native XML Databases	1137
30.7	XML in Oracle	1139
	<i>Chapter Summary</i>	1142
	<i>Review Questions</i>	1144
	<i>Exercises</i>	1145
Part 9	Business Intelligence	1147
Chapter 31	Data Warehousing Concepts	1149
31.1	Introduction to Data Warehousing	1150
31.1.1	The Evolution of Data Warehousing	1150
31.1.2	Data Warehousing Concepts	1151

31.1.3	Benefits of Data Warehousing	1152
31.1.4	Comparison of OLTP Systems and Data Warehousing	1153
31.1.5	Problems of Data Warehousing	1154
31.2	Data Warehouse Architecture	1156
31.2.1	Operational Data	1156
31.2.2	Operational Data Store	1157
31.2.3	Load Manager	1158
31.2.4	Warehouse Manager	1158
31.2.5	Query Manager	1158
31.2.6	Detailed Data	1159
31.2.7	Lightly and Highly Summarized Data	1159
31.2.8	Archive/Backup Data	1159
31.2.9	Metadata	1159
31.2.10	End-User Access Tools	1160
31.3	Data Warehouse Data Flows	1161
31.3.1	Inflow	1162
31.3.2	Upflow	1163
31.3.3	Downflow	1164
31.3.4	Outflow	1164
31.3.5	Metaflow	1165
31.4	Data Warehousing Tools and Technologies	1165
31.4.1	Extraction, Cleansing, and Transformation Tools	1165
31.4.2	Data Warehouse DBMS	1166
31.4.3	Data Warehouse Metadata	1169
31.4.4	Administration and Management Tools	1171
31.5	Data Marts	1171
31.5.1	Reasons for Creating a Data Mart	1173
31.5.2	Data Marts Issues	1173
31.6	Data Warehousing Using Oracle	1175
31.6.1	Oracle9i	1175
	<i>Chapter Summary</i>	1178
	<i>Review Questions</i>	1180
	<i>Exercise</i>	1180

Chapter 32	Data Warehousing Design	1181
32.1	Designing a Data Warehouse Database	1182
32.2	Dimensionality Modeling	1183
32.2.1	Comparison of DM and ER models	1186
32.3	Database Design Methodology for Data Warehouses	1187
32.4	Criteria for Assessing the Dimensionality of a Data Warehouse	1195
32.5	Data Warehousing Design Using Oracle	1196
32.5.1	Oracle Warehouse Builder Components	1197
32.5.2	Using Oracle Warehouse Builder	1198

	<i>Chapter Summary</i>	1202
	<i>Review Questions</i>	1203
	<i>Exercises</i>	1203
Chapter 33	OLAP	1204
33.1	Online Analytical Processing	1205
33.1.1	OLAP Benchmarks	1206
33.2	OLAP Applications	1206
33.2.1	OLAP Benefits	1208
33.3	Representation of Multi-Dimensional Data	1209
33.4	OLAP Tools	1211
33.4.1	Codd's Rules for OLAP Tools	1211
33.4.2	Categories of OLAP Tools	1214
33.5	OLAP Extensions to the SQL Standard	1217
33.5.1	Extended Grouping Capabilities	1218
33.5.2	Elementary OLAP Operators	1222
33.6	Oracle OLAP	1224
33.6.1	Oracle OLAP Environment	1225
33.6.2	Platform for Business Intelligence Applications	1225
33.6.3	Oracle9i Database	1226
33.6.4	Oracle OLAP	1228
33.6.5	Performance	1229
33.6.6	System Management	1229
33.6.7	System Requirements	1230
	<i>Chapter Summary</i>	1230
	<i>Review Questions</i>	1231
	<i>Exercises</i>	1231
Chapter 34	Data Mining	1232
34.1	Data Mining	1233
34.2	Data Mining Techniques	1233
34.2.1	Predictive Modeling	1235
34.2.2	Database Segmentation	1236
34.2.3	Link Analysis	1237
34.2.4	Deviation Detection	1238
34.3	The Data Mining Process	1239
34.3.1	The CRISP-DM Model	1239
34.4	Data Mining Tools	1241
34.5	Data Mining and Data Warehousing	1242
34.6	Oracle Data Mining (ODM)	1242
34.6.1	Data Mining Capabilities	1242
34.6.2	Enabling Data Mining Applications	1243

34.6.3 Predictions and Insights	1243
34.6.4 Oracle Data Mining Environment	1243
<i>Chapter Summary</i>	1245
<i>Review Questions</i>	1246
<i>Exercises</i>	1246

Appendices **1247**

A	Users' Requirements Specification for <i>DreamHome</i> Case Study	1249
A.1	Branch User Views of <i>DreamHome</i>	1249
A.1.1	Data Requirements	1249
A.1.2	Transaction Requirements (Sample)	1251
A.2	Staff User Views of <i>DreamHome</i>	1252
A.2.1	Data Requirements	1252
A.2.2	Transaction Requirements (Sample)	1253
B	Other Case Studies	1255
B.1	The <i>University Accommodation Office</i> Case Study	1255
B.1.1	Data Requirements	1255
B.1.2	Query Transactions (Sample)	1257
B.2	The <i>EasyDrive School of Motoring</i> Case Study	1258
B.2.1	Data Requirements	1258
B.2.2	Query Transactions (Sample)	1259
B.3	The <i>Wellmeadows Hospital</i> Case Study	1260
B.3.1	Data Requirements	1260
B.3.2	Transaction Requirements (Sample)	1266
C	File Organizations and Indexes	1268
C.1	Basic Concepts	1269
C.2	Unordered Files	1270
C.3	Ordered Files	1271
C.4	Hash Files	1272
C.4.1	Dynamic Hashing	1275
C.4.2	Limitations of Hashing	1276
C.5	Indexes	1277
C.5.1	Types of Index	1277
C.5.2	Indexed Sequential Files	1278
C.5.3	Secondary Indexes	1279
C.5.4	Multilevel Indexes	1280
C.5.5	B ⁺ -trees	1280

C.5.6 Bitmap Indexes	1283
C.5.7 Join Indexes	1284
C.6 Clustered and Non-Clustered Tables	1286
C.6.1 Indexed Clusters	1286
C.6.2 Hash Clusters	1287
C.7 Guidelines for Selecting File Organizations	1288
<i>Appendix Summary</i>	1291
D When is a DBMS Relational?	1293
E Programmatic SQL	1298
E.1 Embedded SQL	1299
E.1.1 Simple Embedded SQL Statements	1299
E.1.2 SQL Communications Area	1301
E.1.3 Host Language Variables	1303
E.1.4 Retrieving Data Using Embedded SQL and Cursors	1304
E.1.5 Using Cursors to Modify Data	1310
E.1.6 ISO Standard for Embedded SQL	1311
E.2 Dynamic SQL	1312
E.3 The Open Database Connectivity (ODBC) Standard	1313
E.3.1 The ODBC Architecture	1314
E.3.2 ODBC Conformance Levels	1315
<i>Appendix Summary</i>	1318
<i>Review Questions</i>	1319
<i>Exercises</i>	1319
F Alternative ER Modeling Notations	1320
F.1 ER Modeling Using the Chen Notation	1320
F.2 ER Modeling Using the Crow's Feet Notation	1320
G Summary of the Database Design Methodology for Relational Databases	1326
References	1332
Further Reading	1345
Index	1356