Preface xvii

Chapter 1 Introduction
1.1 Database-System Applications 1
1.2 Purpose of Database Systems 3
1.3 View of Data 5
1.4 Database Languages 9
1.5 Relational Databases 11
1.6 Database Design 14
1.7 Object-Based and Semistructured Databases 19
1.8 Data Storage and Querying 20
1.9 Transaction Management 22
1.10 Data Mining and Analysis 23
1.11 Database Architecture 24
1.12 Database Users and Administrators 26
1.13 History of Database Systems 28
1.14 Summary 30
Exercises 31
Bibliographical Notes 32

PART 1  ■  RELATIONAL DATABASES

Chapter 2 Relational Model
2.1 Structure of Relational Databases 37
2.2 Fundamental Relational-Algebra Operations 46
2.3 Additional Relational-Algebra Operations 55
2.4 Extended Relational-Algebra Operations 60
2.5 Null Values 66
2.6 Modification of the Database 68
2.7 Summary 70
Exercises 71
Bibliographical Notes 73
## Chapter 3  SQL

3.1 Background 75
3.2 Data Definition 77
3.3 Basic Structure of SQL Queries 80
3.4 Set Operations 87
3.5 Aggregate Functions 89
3.6 Null Values 91
3.7 Nested Subqueries 93
3.8 Complex Queries 97
3.9 Views 99
3.10 Modification of the Database 103
3.11 Joined Relations** 110
3.12 Summary 115

Exercises 116
Bibliographical Notes 120

## Chapter 4  Advanced SQL

4.1 SQL Data Types and Schemas 121
4.2 Integrity Constraints 126
4.3 Authorization 133
4.4 Embedded SQL 134
4.5 Dynamic SQL 137
4.6 Functions and Procedural Constructs** 145
4.7 Recursive Queries** 151
4.8 Advanced SQL Features** 155
4.9 Summary 158

Exercises 159
Bibliographical Notes 162

## Chapter 5  Other Relational Languages

5.1 The Tuple Relational Calculus 163
5.2 The Domain Relational Calculus 168
5.3 Query-by-Example 171
5.4 Datalog 180

5.5 Summary 194

Exercises 195
Bibliographical Notes 198

## PART 2  DATABASE DESIGN

## Chapter 6  Database Design and the E-R Model

6.1 Overview of the Design Process 201
6.2 The Entity-Relationship Model 204
6.3 Constraints 210
6.4 Entity-Relationship Diagrams 214
6.5 Entity-Relationship Design Issues 220
6.6 Weak Entity Sets 225
6.7 Extended E-R Features 227
6.8 Database Design for Banking Enterprise 236
6.9 Reduction to Relational Schemas 241
6.10 Other Aspects of Database Design 248
6.11 The Unified Modeling Language UML** 251

6.12 Summary 254

Exercises 256
Bibliographical Notes 261
Chapter 7  Relational Database Design

7.1 Features of Good Relational Designs 263
7.2 Atomic Domains and First Normal Form 268
7.3 Decomposition Using Functional Dependencies 270
7.4 Functional-Dependency Theory 278
7.5 Decomposition Using Functional Dependencies 288
7.6 Decomposition Using Multivalued Dependencies 293
7.7 More Normal Forms 298
7.8 Database-Design Process 299
7.9 Modeling Temporal Data 302
7.10 Summary 304

Exercises 306
Bibliographical Notes 310

Chapter 8  Application Design and Development

8.1 User Interfaces and Tools 311
8.2 Web Interfaces to Databases 314
8.3 Web Fundamentals 315
8.4 Servlets and JSP 321
8.5 Building Large Web Applications 326
8.6 Triggers 329
8.7 Authorization in SQL 335
8.8 Application Security 343
8.9 Summary 350

Exercises 352
Bibliographical Notes 357

PART 3  ■  OBJECT-BASED DATABASES AND XML

Chapter 9  Object-Based Databases

9.1 Overview 361
9.2 Complex Data Types 362
9.3 Structured Types and Inheritance in SQL 365
9.4 Table Inheritance 369
9.5 Array and Multiset Types in SQL 371
9.6 Object-Identity and Reference Types in SQL 376
9.7 Implementing O-R Features 378
9.8 Persistent Programming Languages 379
9.9 Object-Oriented versus Object-Relational 387

9.10 Summary 388

Exercises 389
Bibliographical Notes 393

Chapter 10  XML

10.1 Motivation 395
10.2 Structure of XML Data 399
10.3 XML Document Schema 402
10.4 Querying and Transformation 408
10.5 Application Program Interfaces to XML 420
10.6 Storage of XML Data 421
10.7 XML Applications 428
10.8 Summary 431

Exercises 433
Bibliographical Notes 436
PART 4  DATA STORAGE AND QUERYING

Chapter 11  Storage and File Structure

11.1 Overview of Physical Storage Media  
11.2 Magnetic Disks  
11.3 RAID  
11.4 Tertiary Storage  
11.5 Storage Access  
11.6 File Organization  
11.7 Organization of Records in Files  
11.8 Data-Dictionary Storage  
11.9 Summary  
Exercises  
Bibliographical Notes

Chapter 12  Indexing and Hashing

12.1 Basic Concepts  
12.2 Ordered Indices  
12.3 B+-Tree Index Files  
12.4 B-Tree Index Files  
12.5 Multiple-Key Access  
12.6 Static Hashing  
12.7 Dynamic Hashing  
12.8 Comparison of Ordered Indexing and Hashing  
12.9 Bitmap Indices  
12.10 Index Definition in SQL  
12.11 Summary  
Exercises  
Bibliographical Notes

Chapter 13  Query Processing

13.1 Overview  
13.2 Measures of Query Cost  
13.3 Selection Operation  
13.4 Sorting  
13.5 Join Operation  
13.6 Other Operations  
13.7 Evaluation of Expressions  
13.8 Summary  
Exercises  
Bibliographical Notes

Chapter 14  Query Optimization

14.1 Overview  
14.2 Transformation of Relational Expressions  
14.3 Estimating Statistics of Expression Results  
14.4 Choice of Evaluation Plans  
14.5 Materialized Views**  
14.6 Summary  
Exercises  
Bibliographical Notes
## PART 5  TRANSACTION MANAGEMENT

### Chapter 15 Transactions
- 15.1 Transaction Concept 609
- 15.2 Transaction State 612
- 15.3 Implementation of Atomicity and Durability 615
- 15.4 Concurrent Executions 617
- 15.5 Serializability 620
- 15.6 Recoverability 626
- 15.7 Implementation of Isolation 627
- 15.8 Testing for Serializability 628
- 15.9 Summary 630

### Exercises 632

### Bibliographical Notes 633

### Chapter 16 Concurrency Control
- 16.1 Lock-Based Protocols 635
- 16.2 Timestamp-Based Protocols 648
- 16.3 Validation-Based Protocols 651
- 16.4 Multiple Granularity 653
- 16.5 Multiversion Schemes 656
- 16.6 Deadlock Handling 659
- 16.7 Insert and Delete Operations 664
- 16.8 Weak Levels of Consistency 667
- 16.9 Concurrency in Index Structures∗∗ 669
- 16.10 Summary 673

### Exercises 676

### Bibliographical Notes 680

### Chapter 17 Recovery System
- 17.1 Failure Classification 683
- 17.2 Storage Structure 684
- 17.3 Recovery and Atomicity 688
- 17.4 Log-Based Recovery 689
- 17.5 Recovery with Concurrent Transactions 697
- 17.6 Buffer Management 699
- 17.7 Failure with Loss of Nonvolatile Storage 702
- 17.8 Advanced Recovery Techniques∗∗ 703
- 17.9 Remote Backup Systems 711
- 17.10 Summary 713

### Exercises 716

### Bibliographical Notes 718

## PART 6  DATA MINING AND INFORMATION RETRIEVAL

### Chapter 18 Data Analysis and Mining
- 18.1 Decision-Support Systems 723
- 18.2 Data Analysis and OLAP 725
- 18.3 Data Warehousing 736
- 18.4 Data Mining 739
- 18.5 Summary 752

### Exercises 754

### Bibliographical Notes 756
### Chapter 19  Information Retrieval

19.1 Overview 759
19.2 Relevance Ranking Using Terms 761
19.3 Relevance Using Hyperlinks 763
19.4 Synonyms, Homonyms and Ontologies 768
19.5 Indexing of Documents 769
19.6 Measuring Retrieval Effectiveness 770
19.7 Web Search Engines 771
19.8 Information Retrieval and Structured Data 772
19.9 Directories 773
19.10 Summary 776
Exercises 777
Bibliographical Notes 779

### PART 7  ■  SYSTEM ARCHITECTURE

### Chapter 20  Database-System Architectures

20.1 Centralized and Client–Server Architectures 783
20.2 Server System Architectures 786
20.3 Parallel Systems 790
20.4 Distributed Systems 797
20.5 Network Types 801
20.6 Summary 803
Exercises 805
Bibliographical Notes 807

### Chapter 21  Parallel Databases

21.1 Introduction 809
21.2 I/O Parallelism 810
21.3 Interquery Parallelism 814
21.4 Intraquery Parallelism 815
21.5 Intraoperation Parallelism 816
21.6 Interoperation Parallelism 824
21.7 Design of Parallel Systems 826
21.8 Summary 827
Exercises 829
Bibliographical Notes 831

### Chapter 22  Distributed Databases

22.1 Homogeneous and Heterogeneous Databases 833
22.2 Distributed Data Storage 834
22.3 Distributed Transactions 837
22.4 Commit Protocols 840
22.5 Concurrency Control in Distributed Databases 846
22.6 Availability 854
22.7 Distributed Query Processing 859
22.8 Heterogeneous Distributed Databases 862
22.9 Directory Systems 865
22.10 Summary 870
Exercises 873
Bibliographical Notes 876
PART 8  ■  OTHER TOPICS

Chapter 23  Advanced Application Development

23.1 Performance Tuning  881  23.5 Summary  900
23.2 Performance Benchmarks  891  Exercises  902
23.3 Standardization  895  Bibliographical Notes  903
23.4 Application Migration  899

Chapter 24  Advanced Data Types and New Applications

24.1 Motivation  905  24.5 Mobility and Personal Databases  922
24.2 Time in Databases  906  24.6 Summary  927
24.3 Spatial and Geographic Data  908  Exercises  929
24.4 Multimedia Databases  919  Bibliographical Notes  931

Chapter 25  Advanced Transaction Processing

25.1 Transaction-Processing Monitors  933  25.7 Transaction Management in
25.2 Transactional Workflows  938  Multidatabases  956
25.3 E-Commerce  944  25.8 Summary  959
25.4 Main-Memory Databases  947  Exercises  962
25.5 Real-Time Transaction Systems  949  Bibliographical Notes  964
25.6 Long-Duration Transactions  950

PART 9  ■  CASE STUDIES

Chapter 26  PostgreSQL

26.1 Introduction  967  26.5 Storage and Indexing  988
26.2 User Interfaces  968  26.6 Query Processing and
26.3 SQL Variations and Extensions  971  Optimization  991
26.4 Transaction Management in  26.7 System Architecture  994
PostgreSQL  979  Bibliographical Notes  995
## Chapter 27  Oracle

27.1 Database Design and Querying Tools 997  
27.2 SQL Variations and Extensions 999  
27.3 Storage and Indexing 1001  
27.4 Query Processing and Optimization 1010  
27.5 Concurrency Control and Recovery 1017  
27.6 System Architecture 1019  
27.7 Replication, Distribution, and External Data 1022  
27.8 Database Administration Tools 1024  
27.9 Data Mining 1025  
Bibliographical Notes 1026

## Chapter 28  IBM DB2 Universal Database

28.1 Overview 1027  
28.2 Database-Design Tools 1029  
28.3 SQL Variations and Extensions 1029  
28.4 Storage and Indexing 1034  
28.5 Multidimensional Clustering 1037  
28.6 Query Processing and Optimization 1040  
28.7 Materialized Query Tables 1045  
28.8 Autonomic Features in DB2 1047  
28.9 Tools and Utilities 1048  
28.10 Concurrency Control and Recovery 1050  
28.11 System Architecture 1052  
28.12 Replication, Distribution and External Data 1053  
28.13 Business Intelligence Features 1054  
Bibliographical Notes 1055

## Chapter 29  Microsoft SQL Server

29.1 Management, Design, and Querying Tools 1057  
29.2 SQL Variations and Extensions 1062  
29.3 Storage and Indexing 1066  
29.4 Query Processing and Optimization 1069  
29.5 Concurrency and Recovery 1074  
29.6 System Architecture 1078  
29.7 Data Access 1080  
29.8 Distributed Heterogeneous Query Processing 1081  
29.9 Replication 1082  
29.10 Server Programming in .NET 1084  
29.11 XML Support in SQL Server 2005 1089  
29.12 SQL Server Service Broker 1094  
29.13 Data Warehouse and Business Intelligence 1096  
Bibliographical Notes 1100

## PART 10  ■  APPENDICES

### Appendix A  Network Model (contents online)

A.1 Basic Concepts A1  
A.2 Data-Structure Diagrams A2  
A.3 The DBTG CODASYL Model A7  
A.4 DBTG Data-Retrieval Facility A13  
A.5 DBTG Update Facility A20  
A.6 DBTG Set-Processing Facility A22  
A.7 Mapping of Networks to Files A27  
A.8 Summary A31  
Exercises A32  
Bibliographical Notes A35
Appendix B  Hierarchical Model (contents online)

B.1 Basic Concepts  B1
B.2 Tree-Structure Diagrams  B2
B.3 Data-Retrieval Facility  B13
B.4 Update Facility  B18
B.5 Virtual Records  B21
B.6 Mapping of Hierarchies to Files  B22
B.7 The IMS Database System  B24
B.8 Summary  B25

Exercises  B26
Bibliographical Notes  B29

Appendix C  Advanced Relational Database Design (contents online)

C.1 Multivalued Dependencies  C1
C.2 Join Dependencies  C5
C.3 Domain-Key Normal Form  C8
C.4 Summary  C10

Exercises  C10
Bibliographical Notes  C11

Bibliography  1101

Index  1129