

FUNDAMENTALS OF

# Database Systems

SIXTH EDITION

**Ramez Elmasri**

*Department of Computer Science and Engineering  
The University of Texas at Arlington*

**Shamkant B. Navathe**

*College of Computing  
Georgia Institute of Technology*

**Addison-Wesley**

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

# Contents

## ■ part 1

### Introduction to Databases ■

#### chapter 1 Databases and Database Users 3

- 1.1 Introduction 4
- 1.2 An Example 6
- 1.3 Characteristics of the Database Approach 9
- 1.4 Actors on the Scene 14
- 1.5 Workers behind the Scene 16
- 1.6 Advantages of Using the DBMS Approach 17
- 1.7 A Brief History of Database Applications 23
- 1.8 When Not to Use a DBMS 26
- 1.9 Summary 27
- Review Questions 27
- Exercises 28
- Selected Bibliography 28

#### chapter 2 Database System Concepts and Architecture 29

- 2.1 Data Models, Schemas, and Instances 30
- 2.2 Three-Schema Architecture and Data Independence 33
- 2.3 Database Languages and Interfaces 36
- 2.4 The Database System Environment 40
- 2.5 Centralized and Client/Server Architectures for DBMSs 44
- 2.6 Classification of Database Management Systems 49
- 2.7 Summary 52
- Review Questions 53
- Exercises 54
- Selected Bibliography 55

## ■ part 2

### The Relational Data Model and SQL ■

#### chapter 3 The Relational Data Model and Relational Database Constraints 59

- 3.1 Relational Model Concepts 60
- 3.2 Relational Model Constraints and Relational Database Schemas 67
- 3.3 Update Operations, Transactions, and Dealing with Constraint Violations 75
- 3.4 Summary 79
- Review Questions 80
- Exercises 80
- Selected Bibliography 85

#### chapter 4 Basic SQL 87

- 4.1 SQL Data Definition and Data Types 89
- 4.2 Specifying Constraints in SQL 94
- 4.3 Basic Retrieval Queries in SQL 97
- 4.4 INSERT, DELETE, and UPDATE Statements in SQL 107
- 4.5 Additional Features of SQL 110
- 4.6 Summary 111
- Review Questions 112
- Exercises 112
- Selected Bibliography 114

#### chapter 5 More SQL: Complex Queries, Triggers, Views, and Schema Modification 115

- 5.1 More Complex SQL Retrieval Queries 115
- 5.2 Specifying Constraints as Assertions and Actions as Triggers 131
- 5.3 Views (Virtual Tables) in SQL 133
- 5.4 Schema Change Statements in SQL 137
- 5.5 Summary 139
- Review Questions 141
- Exercises 141
- Selected Bibliography 143

## chapter **6** The Relational Algebra and Relational Calculus 145

- 6.1 Unary Relational Operations: SELECT and PROJECT 147
- 6.2 Relational Algebra Operations from Set Theory 152
- 6.3 Binary Relational Operations: JOIN and DIVISION 157
- 6.4 Additional Relational Operations 165
- 6.5 Examples of Queries in Relational Algebra 171
- 6.6 The Tuple Relational Calculus 174
- 6.7 The Domain Relational Calculus 183
- 6.8 Summary 185
- Review Questions 186
- Exercises 187
- Laboratory Exercises 192
- Selected Bibliography 194

## ■ part **3**

### Conceptual Modeling and Database Design ■

## chapter **7** Data Modeling Using the Entity-Relationship (ER) Model 199

- 7.1 Using High-Level Conceptual Data Models for Database Design 200
- 7.2 A Sample Database Application 202
- 7.3 Entity Types, Entity Sets, Attributes, and Keys 203
- 7.4 Relationship Types, Relationship Sets, Roles,  
and Structural Constraints 212
- 7.5 Weak Entity Types 219
- 7.6 Refining the ER Design for the COMPANY Database 220
- 7.7 ER Diagrams, Naming Conventions, and Design Issues 221
- 7.8 Example of Other Notation: UML Class Diagrams 226
- 7.9 Relationship Types of Degree Higher than Two 228
- 7.10 Summary 232
- Review Questions 234
- Exercises 234
- Laboratory Exercises 241
- Selected Bibliography 243

**chapter 8 The Enhanced Entity-Relationship (EER) Model 245**

- 8.1 Subclasses, Superclasses, and Inheritance 246
- 8.2 Specialization and Generalization 248
- 8.3 Constraints and Characteristics of Specialization and Generalization Hierarchies 251
- 8.4 Modeling of UNION Types Using Categories 258
- 8.5 A Sample UNIVERSITY EER Schema, Design Choices, and Formal Definitions 260
- 8.6 Example of Other Notation: Representing Specialization and Generalization in UML Class Diagrams 265
- 8.7 Data Abstraction, Knowledge Representation, and Ontology Concepts 267
- 8.8 Summary 273
- Review Questions 273
- Exercises 274
- Laboratory Exercises 281
- Selected Bibliography 284

**chapter 9 Relational Database Design by ER- and EER-to-Relational Mapping 285**

- 9.1 Relational Database Design Using ER-to-Relational Mapping 286
- 9.2 Mapping EER Model Constructs to Relations 294
- 9.3 Summary 299
- Review Questions 299
- Exercises 299
- Laboratory Exercises 301
- Selected Bibliography 302

**chapter 10 Practical Database Design Methodology and Use of UML Diagrams 303**

- 10.1 The Role of Information Systems in Organizations 304
- 10.2 The Database Design and Implementation Process 309
- 10.3 Use of UML Diagrams as an Aid to Database Design Specification 328
- 10.4 Rational Rose: A UML-Based Design Tool 337
- 10.5 Automated Database Design Tools 342

10.6 Summary	345
Review Questions	347
Selected Bibliography	348

## ■ part 4

### Object, Object-Relational, and XML: Concepts, Models, Languages, and Standards ■

#### chapter **11** Object and Object-Relational Databases **353**

11.1 Overview of Object Database Concepts	355
11.2 Object-Relational Features: Object Database Extensions to SQL	369
11.3 The ODMG Object Model and the Object Definition Language ODL	376
11.4 Object Database Conceptual Design	395
11.5 The Object Query Language OQL	398
11.6 Overview of the C++ Language Binding in the ODMG Standard	407
11.7 Summary	408
Review Questions	409
Exercises	411
Selected Bibliography	412

#### chapter **12** XML: Extensible Markup Language **415**

12.1 Structured, Semistructured, and Unstructured Data	416
12.2 XML Hierarchical (Tree) Data Model	420
12.3 XML Documents, DTD, and XML Schema	423
12.4 Storing and Extracting XML Documents from Databases	431
12.5 XML Languages	432
12.6 Extracting XML Documents from Relational Databases	436
12.7 Summary	442
Review Questions	442
Exercises	443
Selected Bibliography	443

## ■ part 5

### Database Programming Techniques ■

#### chapter **13** Introduction to SQL Programming Techniques 447

- 13.1 Database Programming: Techniques and Issues 448
- 13.2 Embedded SQL, Dynamic SQL, and SQLJ 451
- 13.3 Database Programming with Function Calls: SQL/CLI and JDBC 464
- 13.4 Database Stored Procedures and SQL/PSM 473
- 13.5 Comparing the Three Approaches 476
- 13.6 Summary 477
- Review Questions 478
- Exercises 478
- Selected Bibliography 479

#### chapter **14** Web Database Programming Using PHP 481

- 14.1 A Simple PHP Example 482
- 14.2 Overview of Basic Features of PHP 484
- 14.3 Overview of PHP Database Programming 491
- 14.4 Summary 496
- Review Questions 496
- Exercises 497
- Selected Bibliography 497

## ■ part 6

### Database Design Theory and Normalization ■

#### chapter **15** Basics of Functional Dependencies and Normalization for Relational Databases 501

- 15.1 Informal Design Guidelines for Relation Schemas 503
- 15.2 Functional Dependencies 513
- 15.3 Normal Forms Based on Primary Keys 516
- 15.4 General Definitions of Second and Third Normal Forms 525
- 15.5 Boyce-Codd Normal Form 529

15.6 Multivalued Dependency and Fourth Normal Form	531
15.7 Join Dependencies and Fifth Normal Form	534
15.8 Summary	535
Review Questions	536
Exercises	537
Laboratory Exercises	542
Selected Bibliography	542

## chapter **16** Relational Database Design Algorithms and Further Dependencies **543**

16.1 Further Topics in Functional Dependencies: Inference Rules, Equivalence, and Minimal Cover	545
16.2 Properties of Relational Decompositions	551
16.3 Algorithms for Relational Database Schema Design	557
16.4 About Nulls, Dangling Tuples, and Alternative Relational Designs	563
16.5 Further Discussion of Multivalued Dependencies and 4NF	567
16.6 Other Dependencies and Normal Forms	571
16.7 Summary	575
Review Questions	576
Exercises	576
Laboratory Exercises	578
Selected Bibliography	579

# ■ part **7**

## File Structures, Indexing, and Hashing ■

### chapter **17** Disk Storage, Basic File Structures, and Hashing **583**

17.1 Introduction	584
17.2 Secondary Storage Devices	587
17.3 Buffering of Blocks	593
17.4 Placing File Records on Disk	594
17.5 Operations on Files	599
17.6 Files of Unordered Records (Heap Files)	601
17.7 Files of Ordered Records (Sorted Files)	603
17.8 Hashing Techniques	606



- 17.9 Other Primary File Organizations 616
- 17.10 Parallelizing Disk Access Using RAID Technology 617
- 17.11 New Storage Systems 621
- 17.12 Summary 624
- Review Questions 625
- Exercises 626
- Selected Bibliography 630

chapter **18** Indexing Structures for Files 631

- 18.1 Types of Single-Level Ordered Indexes 632
- 18.2 Multilevel Indexes 643
- 18.3 Dynamic Multilevel Indexes Using B-Trees and B<sup>+</sup>-Trees 646
- 18.4 Indexes on Multiple Keys 660
- 18.5 Other Types of Indexes 663
- 18.6 Some General Issues Concerning Indexing 668
- 18.7 Summary 670
- Review Questions 671
- Exercises 672
- Selected Bibliography 674

■ **part 8**

**Query Processing and Optimization,  
and Database Tuning** ■

chapter **19** Algorithms for Query Processing  
and Optimization 679

- 19.1 Translating SQL Queries into Relational Algebra 681
- 19.2 Algorithms for External Sorting 682
- 19.3 Algorithms for SELECT and JOIN Operations 685
- 19.4 Algorithms for PROJECT and Set Operations 696
- 19.5 Implementing Aggregate Operations and OUTER JOINS 698
- 19.6 Combining Operations Using Pipelining 700
- 19.7 Using Heuristics in Query Optimization 700
- 19.8 Using Selectivity and Cost Estimates in Query Optimization 710
- 19.9 Overview of Query Optimization in Oracle 721
- 19.10 Semantic Query Optimization 722
- 19.11 Summary 723

Review Questions	723
Exercises	724
Selected Bibliography	725

## chapter **20** Physical Database Design and Tuning **727**

20.1 Physical Database Design in Relational Databases	727
20.2 An Overview of Database Tuning in Relational Systems	733
20.3 Summary	739
Review Questions	739
Selected Bibliography	740

# ■ part **9**

## Transaction Processing, Concurrency Control, and Recovery ■

### chapter **21** Introduction to Transaction Processing Concepts and Theory **743**

21.1 Introduction to Transaction Processing	744
21.2 Transaction and System Concepts	751
21.3 Desirable Properties of Transactions	754
21.4 Characterizing Schedules Based on Recoverability	755
21.5 Characterizing Schedules Based on Serializability	759
21.6 Transaction Support in SQL	770
21.7 Summary	772
Review Questions	772
Exercises	773
Selected Bibliography	775

### chapter **22** Concurrency Control Techniques **777**

22.1 Two-Phase Locking Techniques for Concurrency Control	778
22.2 Concurrency Control Based on Timestamp Ordering	788
22.3 Multiversion Concurrency Control Techniques	791
22.4 Validation (Optimistic) Concurrency Control Techniques	794
22.5 Granularity of Data Items and Multiple Granularity Locking	795
22.6 Using Locks for Concurrency Control in Indexes	798
22.7 Other Concurrency Control Issues	800

22.8 Summary 802  
Review Questions 803  
Exercises 804  
Selected Bibliography 804

chapter **23 Database Recovery Techniques 807**

23.1 Recovery Concepts 808  
23.2 NO-UNDO/REDO Recovery Based on Deferred Update 815  
23.3 Recovery Techniques Based on Immediate Update 817  
23.4 Shadow Paging 820  
23.5 The ARIES Recovery Algorithm 821  
23.6 Recovery in Multidatabase Systems 825  
23.7 Database Backup and Recovery from Catastrophic Failures 826  
23.8 Summary 827  
Review Questions 828  
Exercises 829  
Selected Bibliography 832

■ **part 10**

**Additional Database Topics:  
Security and Distribution ■**

chapter **24 Database Security 835**

24.1 Introduction to Database Security Issues 836  
24.2 Discretionary Access Control Based on Granting  
and Revoking Privileges 842  
24.3 Mandatory Access Control and Role-Based Access Control  
for Multilevel Security 847  
24.4 SQL Injection 855  
24.5 Introduction to Statistical Database Security 859  
24.6 Introduction to Flow Control 860  
24.7 Encryption and Public Key Infrastructures 862  
24.8 Privacy Issues and Preservation 866  
24.9 Challenges of Database Security 867  
24.10 Oracle Label-Based Security 868  
24.11 Summary 870

Review Questions	872
Exercises	873
Selected Bibliography	874

## chapter **25** Distributed Databases 877

25.1 Distributed Database Concepts	878
25.2 Types of Distributed Database Systems	883
25.3 Distributed Database Architectures	887
25.4 Data Fragmentation, Replication, and Allocation Techniques for Distributed Database Design	894
25.5 Query Processing and Optimization in Distributed Databases	901
25.6 Overview of Transaction Management in Distributed Databases	907
25.7 Overview of Concurrency Control and Recovery in Distributed Databases	909
25.8 Distributed Catalog Management	913
25.9 Current Trends in Distributed Databases	914
25.10 Distributed Databases in Oracle	915
25.11 Summary	919
Review Questions	921
Exercises	922
Selected Bibliography	924

# ■ part 11

## Advanced Database Models, Systems, and Applications ■

### chapter **26** Enhanced Data Models for Advanced Applications 931

26.1 Active Database Concepts and Triggers	933
26.2 Temporal Database Concepts	943
26.3 Spatial Database Concepts	957
26.4 Multimedia Database Concepts	965
26.5 Introduction to Deductive Databases	970
26.6 Summary	983
Review Questions	985
Exercises	986
Selected Bibliography	989

chapter **27** Introduction to Information Retrieval  
and Web Search 993

- 27.1 Information Retrieval (IR) Concepts 994
- 27.2 Retrieval Models 1001
- 27.3 Types of Queries in IR Systems 1007
- 27.4 Text Preprocessing 1009
- 27.5 Inverted Indexing 1012
- 27.6 Evaluation Measures of Search Relevance 1014
- 27.7 Web Search and Analysis 1018
- 27.8 Trends in Information Retrieval 1028
- 27.9 Summary 1030
- Review Questions 1031
- Selected Bibliography 1033

chapter **28** Data Mining Concepts 1035

- 28.1 Overview of Data Mining Technology 1036
- 28.2 Association Rules 1039
- 28.3 Classification 1051
- 28.4 Clustering 1054
- 28.5 Approaches to Other Data Mining Problems 1057
- 28.6 Applications of Data Mining 1060
- 28.7 Commercial Data Mining Tools 1060
- 28.8 Summary 1063
- Review Questions 1063
- Exercises 1064
- Selected Bibliography 1065

chapter **29** Overview of Data Warehousing  
and OLAP 1067

- 29.1 Introduction, Definitions, and Terminology 1067
- 29.2 Characteristics of Data Warehouses 1069
- 29.3 Data Modeling for Data Warehouses 1070
- 29.4 Building a Data Warehouse 1075
- 29.5 Typical Functionality of a Data Warehouse 1078
- 29.6 Data Warehouse versus Views 1079
- 29.7 Difficulties of Implementing Data Warehouses 1080

29.8 Summary	1081	
Review Questions	1081	
Selected Bibliography	1082	
appendix <b>A</b>	<b>Alternative Diagrammatic Notations for ER Models</b>	<b>1083</b>
appendix <b>B</b>	<b>Parameters of Disks</b>	<b>1087</b>
appendix <b>C</b>	<b>Overview of the QBE Language</b>	<b>1091</b>
C.1	Basic Retrievals in QBE	1091
C.2	Grouping, Aggregation, and Database Modification in QBE	1095
appendix <b>D</b>	<b>Overview of the Hierarchical Data Model</b> (located on the Companion Website at <a href="http://www.aw.com/elmasri">http://www.aw.com/elmasri</a> )	
appendix <b>E</b>	<b>Overview of the Network Data Model</b> (located on the Companion Website at <a href="http://www.aw.com/elmasri">http://www.aw.com/elmasri</a> )	
<b>Selected Bibliography</b>	<b>1099</b>	
<b>Index</b>	<b>1133</b>	