Contents

SRINIVAS COLLEGE OF P.G. MANAGEMENT STUDIES
MANGALORE
575 001.

Preface xv

Chapter 1
Basic Concepts 1

1.1 Data Modeling for a Database 2
1.1.1 Entities and Their Attributes 3
1.1.2 Relationships 7

1.2 Records and Files 7

1.3 Abstraction and Data Integration 9

1.4 The Three-Level Architecture Proposal for a DBMS 14
1.4.1 Mapping between Views 15
1.4.2 Data Independence 17

1.5 Components of a DBMS 20
1.5.1 Classification of DBMS Users 20
1.5.2 DBMS Facilities 21
1.5.3 Structure of a DBMS 23
1.5.4 Database Access 27

1.6 Advantages and Disadvantages of a DBMS 27
1.6.1 Advantages of a DBMS 27
1.6.2 Disadvantages of a DBMS 29

1.7 Summary 30
Key Terms 31
Exercises 32
Bibliographic Notes 33
Bibliography 33

Chapter 2
Data Models 34

2.1 Introduction 35

2.2 Data Associations 35
2.2.1 Entities, Attributes, and Associations 36
2.2.2 Relationships among Entities 41
2.2.3 Representation of Associations and Relationships 44

2.3 Data Models Classification 45

2.4 Entity-Relationship Model 46
2.4.1 Entities 47
2.4.2 Relationships 49
2.4.3 Representation of Entities 51
2.4.4 Representation of Relationship Set 52
2.4.5 Generalization and Aggregation 53

2.5 A Comparative Example 57

2.6 Relational Data Model 59

2.7 Network Data Model 63

2.8 Hierarchical Model 67

2.9 A Comparison 69

2.10 Summary 72
Key Terms 72
Exercises 72

6.6 Concluding Remarks 333

6.7 Summary 335
Key Terms 335
Exercises 336
Bibliographic Notes 338
Bibliography 338

Chapter 7
Synthesis Approach and Higher Order Normal Form 340

7.1 Problems in the Decomposition Approach 341

7.2 Normalization through Synthesis 342

7.2.1 Functional Dependencies and Semantics 342
7.2.2 Semantics of Nonfunctional Relationships 343
7.2.3 Synthesis Approach 343
7.2.4 Synthesis Algorithm 345

7.3 Multivalued Dependency 348

7.3.1 MVD and Normalization 352
7.3.2 Axioms for Functional and Multivalued Dependencies 353
7.3.3 Closure under MVDS 354
7.3.4 Fourth Normal Form 357
7.3.5 Lossless Join Decomposition into Fourth Normal Form 357
7.3.6 Enforceability of Dependencies in the Fourth Normal Form 360

7.4 Normalization Using Join Dependency: Fifth Normal Form 361

7.4.1 Join Dependencies 364
7.4.2 Project-Join Normal Form 366

7.5 Domain Key Normal Form 368

7.6 Summary 372
Key Terms 373
Exercises 373
Bibliographic Notes 373
Bibliography 374

Chapter 8
The Network Model 376

8.1 The Network Data Model 377

8.1.1 Expressing Relationship: The DBTG Set 378
8.1.2 Multiple Level Set Construct 380
8.1.3 Complex Multilevel Set Construct 382

8.2 DBTG Set Construct and Restrictions 384

8.2.1 Implementation of the DBTG Set Construct 385

8.3 Expressing an M:N Relationship in DBTG 386

8.4 Cycles in DBTG 391

8.4.1 Set Involving Only One Type of Record 391
8.4.2 Sets Involving Different Record Types in a Cycle 395

8.5 Data Description in the Network Model 395

8.5.1 Record 395
8.5.2 Set 397
8.5.3 Order of Members in a Set 398
8.5.4 Set Membership 399
8.5.5 Structural Constraint 402
8.5.6 Set Selection 402
8.5.7 Singular Sets 403
8.5.8 Area 403

8.6 Schema and Subschema 403

8.7 DBTG Data Manipulation Facility 405

8.7.1 Run Unit 405
8.7.2 Currency Indicators 406
8.7.3 Database Status Registers 406
8.7.4 Record Templates 407
8.7.5 DML Commands 407

8.8 Database Manipulation 407

8.8.1 Operations on Records 410
8.8.2 Operations on Sets 413

8.9 Concluding Remarks 421

8.10 Summary 421

Key Terms 422
Exercises 422
Bibliographic Notes 425
Bibliography 426

Chapter 9
The Hierarchical Data Model 427

9.1 The Tree Concept 428

9.2 Hierarchical Data Model 430

9.2.1 Replication vs. Virtual Record 435
9.2.2 Expressing a Many-to-Many Relationship 436
9.2.3 Another Example of a Many-to-Many Relationship 438

9.3 Data Definition 441

9.4 Data Manipulation 443

9.4.1 User Work Area in the HDM 443
9.4.2 Basic Data Manipulation 444
9.4.3 Sequential Retrieval 444
9.4.4 Sequential Retrieval within a Hierarchy 446

9.5 Updates 447

9.5.1 Insert 448
9.5.2 Modification and Deletion 449
9.5.3 Updates of Virtual Records 450

9.6 Implementation of the Hierarchical Database 451

9.7 Additional Features of the Hierarchical DML 453

9.8 Concluding Remarks 455

9.9 Summary 456
Key Terms 457
Exercises 457
Bibliographic Notes 459
Bibliography 459

Chapter 10
Query Processing 460

10.1 Introduction 461
10.2 An Example 463
10.3 General Strategies for Query Processing 467
10.3.1 Query Representation 467
10.3.2 General Processing Strategies 469
10.4 Transformation into an Equivalent Expression 470
10.5 Expected Size of Relations in the Response 475
10.5.1 Selections 476
10.5.2 Projection 477
10.5.3 Join 478
10.6 Statistics in Estimation 480
10.7 Query Improvement 480
10.8 Query Evaluation 486
10.8.1 One-Variable Expressions 486
10.8.2 Two-Variable Expressions 488
10.8.3 N-Variable Expressions 496
10.8.4 Access Plan 501
10.9 Evaluation of Calculus Expressions 502
10.10 View Processing 503
10.11 A Typical Query Processor 505
10.12 Summary 507
Key Terms 507
Exercises 507
Bibliographic Notes 508
Bibliography 509

Chapter 11
Recovery 510

11.1 Reliability 511
11.1.1 Types of Failures 513
11.1.2 Types of Errors in Database Systems and Possible Detection Schemes 515
11.1.3 Audit Trails 516
11.1.4 Recovery Schemes 516
11.2 Transactions 517
11.2.1 States of a Transaction 521
11.2.2 Properties of a Transaction 523

SRINIVAS COLLEGE OF P.G. MANAGEMENT STUDIES
MANGALORE 575 001.

11.2.3 Failure Anticipation and Recovery 523

11.3 Recovery in a Centralized DBMS 524

11.3.1 Logs 524
11.3.2 Checkpoints 526
11.3.3 Archival Database and Implementation of the Storage Hierarchy of a Database System 529
11.3.4 Do, Undo, and Redo 530

11.4 Reflecting Updates to the Database and Recovery 532

11.4.1 Update in Place 533
11.4.2 Indirect Update and Careful Replacement 538

11.5 Buffer Management, Virtual Memory, and Recovery 544

11.6 Other Logging Schemes 546

11.7 Cost Comparison 547

11.8 Disaster Recovery 548

11.9 Summary 549
Key Terms 550
Exercises 551
Bibliographic Notes 552
Bibliography 552

Chapter 12
Concurrency Management 554

12.1 Introduction 555

12.1.1 Lost Update Problem 557
12.1.2 Inconsistent Read Problem 558
12.1.3 The Phantom Phenomenon 560
12.1.4 Semantics of Concurrent Transactions 561

12.2 Serializability 562

12.2.1 Precedence Graph 565
12.2.2 Serializability Algorithm: Read-before-Write Protocol 567
12.2.3 Serializability Algorithm: Read-Only and Write-Only Protocols 569

12.3 Concurrency Control 569

12.4 Locking Scheme 570

12.4.1 Two-Phase Locking 574
12.4.2 Granularity of Locking 576
12.4.3 Hierarchy of Locks and Intention-Mode Locking 576
12.4.4 Tree-Locking Protocol 581
12.4.5 DAG Database Storage Structure 582

12.5 Timestamp-Based Order 583

12.6 Optimistic Scheduling 588

12.7 Multiversion Techniques 591

12.8 Deadlock and Its Resolution 594

12.8.1 Deadlock Detection and Recovery 594
12.8.2 Deadlock Avoidance 599

12.9 Atomicity, Concurrency, and Recovery 603

12.10 Summary 603
Key Terms 605
Exercises 605
Bibliographic Notes 607
Bibliography 608

Chapter 13
Database Security, Integrity, and Control 610

13.1 Introduction 611

13.2 Security and Integrity Threats 612

13.3 Defense Mechanisms 614

13.3.1 Security Policies 615
13.3.2 Authorization 619
13.3.3 Identification and Authentication 625
13.3.4 Views/Subschemes in Security Enforcement 626
13.3.5 Distributed Systems 626
13.3.6 Cryptography and Encryption 627

13.4 Integrity 629
13.4.1 Domain or Data-Item Value Integrity Rules 630
13.4.2 Implicit and Data Dependency Constraints 632
13.4.3 Violation of Integrity Constraints and Corrective Action 634
13.4.4 A General Model of Integrity 634
13.4.5 Expressing Integrity Constraints 635

13.5 Statistical Databases 636

13.6 Auditing and Control 639

13.7 Summary 640
Key Terms 641
Exercises 641
Bibliographic Notes 642
Bibliography 643

Chapter 14
Database Design 644

14.1 The Organization and Its Information System 645

14.2 Phase I: Definition of the Problem 645

14.3 Phase II: Analysis of Existing System and Procedures 646

14.4 Phase III: Preliminary Design 648

14.5 Phase IV: Computing System Decision 649

14.6 Phase V: Final Design 650
14.6.1 Designing the Conceptual Database—Relational DBMS 650
14.6.2 Designing the Conceptual Database—Network DBMS 652
14.6.3 Designing the Conceptual Database—Hierarchical DBMS 652
14.6.4 Designing the Physical Database 652

14.7 Phase VI: Implementation and Testing 654

14.8 Phase VII: Operation and Tuning 655

14.9 Summary 655
Key Terms 656
Exercises 656
Bibliographic Notes 658
Bibliography 658

Chapter 15
Distributed Databases 660

15.1 Introduction 661
15.1.1 Advantages and Disadvantages of the DDBMS 662

15.2 Networks 663
15.2.1 Failures and Distributed Databases 668

15.3 Data Distribution 669
15.3.1 Fragmentation 672
15.3.2 Replication 676
15.3.3 Transparency 676
15.3.4 System Catalogs 677

15.4 Object Naming 678

15.5 Distributed Query Processing 679
15.5.1 Parallelism in Distributed Query Processing 681
15.5.2 Semijoin 681
15.5.3 Semijoin and Reduction of Relations 684
15.5.4 Concluding Remarks 686

15.6 Consistency 686

15.7 Concurrency Control 688
15.7.1 Distributed Locking 688
15.7.2 Timestamp-Based Concurrency Control 692

15.8 Distributed Commitment and Recovery 694
15.8.1 Two-Phase Commit 695
15.8.2 Recovery with Two-Phase Commit 698

15.9 Deadlocks in Distributed Systems 700
15.9.1 Deadloack Detection by Probe Computation 702

15.9.2 Deadlock Prevention 705

15.10 Security and Protection 705

15.11 Homogeneous and Heterogeneous Systems 705

15.11.1 The Homogeneous DDBMS 707
15.11.2 The Heterogeneous DDBMS 710

15.12 Summary 715
Key Terms 716
Exercises 716
Bibliographic Notes 718
Bibliography 719

Chapter 16
Current Topics in Database Research 721

16.1 What Is a Knowledge Base System? 722

16.2 Knowledge Base and Database Systems: A Comparison 723

16.3 Knowledge and Its Representation 724

16.3.1 Semantic Networks 726
16.3.2 First-Order Logic (Predicate Logic) 729
16.3.3 Frames 732
16.3.4 Rule-Based Systems (Production Systems) 733
16.3.5 Procedural Representation 736

16.4 Deductive Databases 737

16.5 Expert Systems 742

16.6 Expert Database Systems: Integration of Expert Systems in Database Applications 746

16.7 Object Approach 747

16.7.1 Concept of the Object 749
16.7.2 Names and Identity 750
16.7.3 Database and Identity 752
16.7.4 Implementation of Object Identifiers 752
16.7.5 Object Class and Instantiation 753
16.7.6 Inheritance 753

16.8 Object Databases 755

16.8.1 Pros and Cons of the Object Approach in Databases 757

16.9 Summary 758
Key Terms 759
Exercises 759
Bibliographic Notes 761
Bibliography 761

Chapter 17
Database Machines 766

17.1 Introduction 767

17.2 Database Machine Taxonomy 767

17.2.1 Backend Software Approach 768
17.2.2 Processor Associated with Memory or Intelligent Memory Approach 770
17.2.3 Special Hardware Approach 772

17.3 DBC/1012 Overview and Features 774

17.3.1 Operation of the DBC/1012 780
17.3.2 System Facilities of the DBC/1012 781

17.4 Summary 781
Key Terms 782
Bibliographic Notes 782
Bibliography 782

Appendix 3.1 Sequence 785
Appendix 3.2 Average Seek Distance Computation 789
Appendix 3.3 Rotational Delay (Latency) 791
Appendix 3.4 Probabilities of Access 793
Appendix 4.1 Formal Definitions of Some Relational Algebraic Operations 795
Appendix 4.2 Recursive Queries: Transitive Closure 799
Appendix 5.1 Syntax of SQL 801
Appendix 5.2 Syntax of QUEL 803
Index 805