

# C O N T E N T S

## PART I

### Database Fundamentals 1

#### Chapter 1 Database Overview 2

Minicase: Teck Information Systems, 2

1.1 Database Systems, 3

1.2 The Database Approach: Shareability and Cooperation, 6

*Shareability of Data, 6*

*Centralized Data Management and Control, 6*  
*Adaptability, 7*

1.3 File Processing versus the Database Approach, 7

*Database Approach Summary, 18*

1.4 Roles, 18

*Database Administrator, 18*

*Systems Analyst, 19*

*Programmers, 19*

*Users, 20*

*The Database Management System (DBMS), 20*

1.5 Summary, 27

Episode 1: Community Hospital:  
Background Information, 29

#### Chapter 2 Data: The Foundation of the Database 32

Minicase: Suburban Hospital, 32

2.1 Data versus Information, 34

2.2 The Real World versus the Abstract World, 34

*The User's Logical World, 35*

*The Programmer's Data World: The Data Dictionary, 38*

*The Physical Stored Data World: Data as Seen by Programs and Disk Drives, 42*

#### 2.3 Data Element and Field Characters, 43

*Name, 43*

*Type, 44*

*Representation, 44*

*Length, 44*

*Origin: Real versus Virtual, 44*

*Domain, 44*

*Value, 45*

*Data Independence, 45*

*Key versus Nonkey, 46*

*Data Integrity, 48*

*Data Security, 51*

#### 2.4 Relationships and Associations: Relating Entities, 53

*One-to-One Relationships, 54*

*One-to-Many Relationships, 55*

*Many-to-Many Relationships, 57*

*Cycles, 58*

*Loops, 60*

*Modeling Record/Entity Relationships, 60*

#### 2.5 Data Views or Perspectives: A Taxonomy of Database Types, 63

*Historical Perspective, 64*

*The Conceptual View, 64*

*The External View, 65*

*The Physical View, 65*

#### 2.6 Data Models, 65

*The Relational Data Model, 66*



*The Hierarchical Data Model, 67*

*The Network Data Model, 70*

## 2.7 Summary, 71

Episode 2: Community Hospital Adopts  
the Database Approach, 73

## Chapter 3 Database Structures: Implementing Relationships 77

Minicase: ICU Library, 77

### 3.1 Pointers, 81

*Pointer Types, 82*

*Implementation of Pointers, 88*

*Location, 88*

### 3.2 Data Structures: Relationships among Records, 91

### 3.3 Simple Lists, 92

*Using a Simple List to Sequentially Process  
Records Based on a Primary Key, 93*

*Adding Records to a Simple List, 94*

*Deleting Records from a Simple List, 96*

*Using a Simple List to Process Secondary  
Keys, 98*

*How to Find the Head of a List, 99*

### 3.4 Rings, 102

### 3.5 Two-Way Lists and Rings, 104

### 3.6 Multilists, 106

### 3.7 Inverted Data Structures, 108

### 3.8 Hierarchical Data Structures, 109

*Terminology, 110*

*Physical Representation of Trees, 112*

*Tree for Indexing: Binary and B-Trees, 114*

### 3.9 Network Data Structures, 118

*Simple Network, 121*

*Complex Networks, 121*

### 3.10 Multilists, Secondary Keys, and Set Types: A Comparison, 128

### 3.11 Summary, 129

## PART II Database Design 133

### Chapter 4 An Introduction to Database Design: The Conceptual Database and Normalization 134

Minicase: Fred's Pharmacy, 134

#### 4.1 Conceptual Database Design: An Introduction, 138

#### 4.2 Step 1. Determine the Scope of the Design: What Are the Boundaries? 139

#### 4.3 Step 2. Determine the Relevant Transaction the Database Must Support: What Does the Organization Do? 140

#### 4.4 Step 3. Determine the Business Rules: Constraints and Domains, 140

#### 4.5 Step 4. Determine the Entities: What Are the Organization's Objects? 142

#### 4.6 Step 5. Determine the Identifier of Each Entity: Finding a Primary Key, 143

#### 4.7 Step 6. Draw the DSD and/or E-R Diagram: Graphical Database Models, 144

#### 4.8 Step 7. Ensure that the Conceptual Design Satisfies the Business Rules and Transactions' Needs: How Good a Model Do We Have? 146

#### 4.9 Step 8. Add Attributes to Entities: A Difficult, Error-Prone Step, 148

*Modification Anomalies, 149*

*Functional Dependencies, 151*

*Eliminating Modification Anomalies: Rules for  
Assigning Attributes to Entities, 156*

*Adding Link Attributes, 169*

#### 4.10 Step 9. Add Anticipated Attributes: Back from the Future, 171

#### 4.11 Step 10. Revise and Refine the Design, 171



- 4.12 Step 11. Identify User Views: The Database through the End User's Eyes, 171
- 4.13 A Bottom-Up Alternative, 172
- 4.14 Summary, 176
- Episode 3: Barb Redesigns Her Database, 181

## **Chapter 5** **Logical and Physical Database Design** **193**

- 5.1 Relational Database Design, 194
  - Mapping to the Logical Database*, 194
  - Physical Design of Relational Database*, 196
- 5.2 Network Database Design, 198
  - Mapping to the Logical Database*, 198
  - Physical Design of Network Database*, 202
- 5.3 The Hierarchical Model, 210
  - Mapping to the Logical Database*, 210
  - Hierarchical (IMS) Physical Database Design*, 216
- 5.4 Summary, 218

## **PART III** **The Relational Model** **223**

### **Chapter 6** **An Introduction to the Relational Model: Concepts and Manipulation via Relational Algebra and Calculus and Query-By-Example** **224**

- 6.1 Features of the Relational Model, 225
- 6.2 A Historical Perspective, 227
  - Relational Calculus Implementations*, 227
  - Relational Algebra Implementations*, 228
  - Transform-Oriented Implementations*, 228
  - Graphics-Oriented Implementations*, 228

### **6.3 Terminology and Concepts, 229**

- Relations*, 230
- Occurrences of Relations*, 230
- Base and View Tables*, 230
- Relational Rules*, 231

### **6.4 Relational Assertions, 231**

- Referential Integrity Assertions*, 231
- Primary Key Assertions*, 232

### **6.5 Relational Algebra, 232**

- The SELECT Operator*, 233
- The PROJECT Operator*, 238
- The JOIN Operator*, 240
- Natural Joins and Equijoins*, 241
- The UNION Operator*, 244
- The INTERSECTION Operator*, 247
- The DIFFERENCE Operator*, 247
- The PRODUCT Operator*, 249
- Nesting Operators*, 251

### **6.6 Relational Calculus, 256**

### **6.7 End User Interaction with DB2: QBE, 257**

- Identifying the Desired Tables*, 258
- Retrieval*, 258
- Retrieval of Selected Rows: Implementing Algebra's SELECT Operator*, 259
- Adding New Rows: The I. Operator*, 266
- Changing Existing Rows: The U. Operator*, 266
- Deleting Rows: The D. Operator*, 268

### **6.8 Summary, 269**

## **Chapter 7** **An Introduction to SQL Using Database 2 (DB2)** **273**

### **7.1 DB2 Concepts, 275**

- Table Types*, 275
- Objects*, 276

### **7.2 The Data Description Language, 278**

- Defining Objects: The CREATE Statement*, 278



- Deleting Objects: The DROP Statement, 286*
  - Changing a Table's Structure: The ALTER TABLE Statement, 287*
  - 7.3 Data Retrieval from a Single Table: The SELECT-FROM-WHERE Statement, 288
    - Predicates, 289*
    - SQL and the Relational PROJECT, 292*
    - Built-in SQL Functions, 296*
    - Concatenation, 299*
    - DATE and TIME Functions, 299*
    - SQL and the Relational Select: SQL with Predicates, 300*
    - Sorting Queries: the ORDER BY Clause, 304*
    - Simple Queries with Calculations and Literals, 305*
    - Table Aliases, 306*
    - Retrieval of NULL Values, 306*
    - SQL and UNION: The SQL UNION Statement, 306*
  - 7.4 Querying Multiple Tables: SQL and Join, 309
  - 7.5 Nested Select-from-Where Blocks: Subselects, 313
  - 7.6 Queries Using All or Any, 317
  - 7.7 Queries Using GROUP BY, 320
  - 7.8 Adding Rows to a Table: The Insert Statement, 322
  - 7.9 Deleting Rows from a Table: The Delete Statement, 323
  - 7.10 Changing a Table's Row Values: The Update Statement, 324
  - 7.11 Deleting Columns from Tables, 325
  - 7.12 Use of EXISTS and NOT EXISTS, 326
  - 7.13 Correlated subqueries, 329
  - 7.14 Creating Views: Logical Tables and Domain Constraints, 329
  - 7.15 Security: User-IDs, GRANT, and REVOKE, 333
  - 7.16 Data Integrity: Entity Integrity, Referential Integrity, and Currency Control, 335
    - Entity Integrity, 335*
    - Specifying Referential Integrity Rules, 335*
    - Concurrency Control, 337*
  - 7.17 Querying the System Tables, 339
  - 7.18 ANSI/IBM SQL Differences, 340
  - 7.19 Summary, 340
- Chapter 8**  
**Embedded SQL: COBOL and 4GLs** **345**
- 8.1 The Embedded SQL Environment, 346
    - Create and Test the SQL Portion of the Host Program, 352*
    - Declare the DB2 Tables and the Associated Host Structures and Set up the Communications Area: DATA DIVISION Requirements, 352*
    - Embed the Debugged SQL Statements into the Host Program, 361*
    - Precompile the Host Program, 363*
    - Compile and Link-Edit the Host Program, 364*
    - Bind the Host Program: Creating the Application Plan, 364*
    - Execute the Host Program, 365*
  - 8.2 Single-Row Processing: Embedded SQL without Cursors, 365
    - Retrieval of One Row: Host Structures with the INTO Clause, 365*
    - Handling NULL Values: The NULL Indicator, 366*
    - Updating a Row, 368*
    - Adding a Row, 369*
  - 8.3 Multiple Row Processing: Embedded SQL with and without Cursors, 371
    - Retrieval of Multiple Rows, 371*
    - Updating Multiple Rows, 375*
    - Deleting Multiple Rows, 379*
  - 8.4 Dynamic SQL, 380



- 8.5 Automatic Error Handling: The Whenever Statement, 383
- 8.6 Transaction Processing Statements: Lock, Commit, and Rollback, 385
- 8.7 SQL and 4GLS, 386
- 8.8 Summary, 394
- Episode 4: A Purchase Order Program Is Written Using DB2, 397

## **PART IV**

### **The CODASYL Network Model 409**

#### **Chapter 9**

#### **The CODASYL Network Model: Theory and Implementation 410**

- Minicase: Microtech Luggage Company, 410
- 9.1 Historical Perspective, 413
- 9.2 The CODASYL Architecture, 414
  - Record Types: The Conceptual Files or Entities, 417*
  - Data Items: The Fields, or Attributes, 418*
  - Set Types: Named 1:M Relationships, 419*
  - Currency Indicators: Keeping Track of the Most Recently Accessed Logical Unit, 424*
- 9.3 Writing the Schema: Using the DDL to Describe the Conceptual Database, 426
  - The SCHEMA Entry, 431*
  - The Record Entries, 432*
  - The SET Entry: Defining Set Types, 436*
- 9.4 DDL Integrity Enforcement, 445
  - Domain Constraints, 445*
  - Entity Integrity, 446*
  - Referential Integrity Constraint Enforcement: Foreign Keys, 446*
  - Data Security: Passwords, 446*
  - Data (User) Views, 446*

- 9.5 Pre-1981 DDL Differences, 447
- 9.6 External Databases: Subschemas, 448
  - Designing the External Databases: Determining User Requirements, 448*
  - Defining the External Databases: The Subschema DDL, 452*
- 9.7 Summary, 453
- Episode 5: Defining the Schema Using IDMS, 457

## **Chapter 10**

### **The CODASYL Data Manipulation Language 473**

- Minicase: Microtech Luggage Company—Continued, 473
- 10.1 Invoking the Subschema: How to Access the Desired User View, 475
- 10.2 The Special Registers: Communicating with the DBMS, 476
  - DB-STATUS: Determining the Success or Failure of a DML Call, 476*
- 10.3 Housekeeping Commands: Ready and Finish, 478
- 10.4 Retrieval Commands: FIND and GET, 480
  - Random Access, 481*
  - Sequential Retrieval, 481*
  - Finding the Owner of a Set Occurrence: FIND OWNER, 496*
- 10.5 Adding New Records: The Store Command, 497
  - Storing Records that Are Owners Exclusively or Are Members of Singular Sets Only, 499*
  - Storing Records that Have Owners, 502*
- 10.6 Maintenance of Existing Records: The MODIFY, DISCONNECT, CONNECT, and ERASE Commands, 506
  - Changing Data Item Values: The MODIFY Command, 506*



<i>Removing a Member Record from a Set Occurrence: The DISCONNECT Command, 508</i>	
<i>Inserting a Record into a Set Occurrence: The CONNECT Command, 509</i>	
<i>Deleting Records: The ERASE Command, 512</i>	
10.7 Transaction Processing Commands: COMMIT and ROLLBACK, 515	
<i>Establishing a Quiet Point: The COMMIT Command, 515</i>	
<i>Removing Updates Since the Last Quiet Point: The ROLLBACK Command, 515</i>	
<i>Use of COMMIT and ROLLBACK with Transaction Processing, 516</i>	
10.8 Retaining Currency: Ensuring that the DBMS Does Not Update One or More Currency Indicators, 516	
10.9 Declaratives and the USE Command, 520	
10.10 Concurrency Control, 523	
10.11 Pre-1981 CODASYL DML Differences, 526	
10.12 TAB: A Microcomputer Implementation of IDMS, 527	
10.13 Summary, 527	
Episode 6: A Report Is Written Using IDMS/R, 530	
 <b>PART V</b>	
<b>The Hierarchial Data Model</b>	<b>543</b>
 <b>Chapter 11</b>	
<b>The Hierarchial Data Model and IMS</b>	<b>544</b>
11.1 Historical Perspective, 544	
11.2 IMS Concepts and Terminology, 546	
<i>Segments and Segment Occurrences, 547</i>	
<i>Relationships, 548</i>	
<i>Hierarchic Sequence: Hierarchic Order plus KSF, 548</i>	
<i>Physical and Logical Databases, 553</i>	
11.3 Implementing Relationships Using Pointers, 556	
<i>Hierarchical Pointers, 556</i>	
<i>First-Child Pointers, 557</i>	
<i>Twin Pointers, 558</i>	
<i>Summary, 558</i>	
11.4 Logical Relationships, 559	
11.5 IMS Access Methods, 561	
<i>HSAM, 561</i>	
<i>HISAM, 562</i>	
<i>HDAM, 562</i>	
<i>HIDAM, 562</i>	
<i>MSDB, 563</i>	
<i>DEDB, 563</i>	
<i>INDEX, 563</i>	
11.6 Defining the Physical Database: The DBDGEN Utility, 563	
<i>Specifying the PDB Name, 564</i>	
<i>Defining the Data Set, 564</i>	
<i>Defining the Segments and Their Hierarchic Order, 566</i>	
<i>Defining the Fields, 567</i>	
<i>Defining a HIDAM Index Database, 568</i>	
11.7 Defining the IMS Logical (External) Databases: The PSBGEN Utility, 569	
11.8 Secondary Indexes, 572	
11.9 Fast Path, 574	
<i>The Main Storage Data Base, 574</i>	
<i>The Data Entry Data Base, 574</i>	
11.10 Summary, 574	
Episode 7: Defining the IMS Conceptual Database, 579	
 <b>Chapter 12</b>	
<b>Manipulating the Hierarchial Model: DL/I</b>	<b>586</b>
12.1 The ENTRY and GOBACK	



- Statements: Establishing and Terminating Communications with IMS, 588
- 12.2 The PCB Mask: The Application/IMS Communications Area, 588
- The DBD Name, 591*
  - The Segment Level Number, 591*
  - The Status Code, 592*
  - The Processing Options, 592*
  - The Reserved for IMS Field, 593*
  - The Segment Name, 593*
  - The Length of the Key Feedback Area, 593*
  - The Number of Sensitive Segments, 593*
  - The Key Feedback Area, 593*
- 12.3 The I/O Areas: Record Definitions (UWAS) for IMS Segment Types, 594
- 12.4 The DL/I COBOL Call, 594
- Function, 594*
  - PCB Mask Name, 596*
  - I/O Name, 596*
  - Segment Search Arguments (SSAs), 596*
- 12.5 Kinds of Calls, 599
- Unqualified Calls, 599*
  - Qualified Calls, 599*
- 12.6 Direct Retrieval: The GET UNIQUE (GU) Function, 600
- Direct Retrieval of the First Root: The Unqualified GU Call, 600*
  - Direct Retrieval of the First Occurrence of Any Segment: An Unqualified SSA Call, 600*
  - Direct Retrieval of a Specified Occurrence: A Qualified SSA Call, 601*
  - GU Status Codes, 601*
  - GU with Multiple SSAs, 601*
  - GU with Command Codes, 602*
- 12.7 Sequential Retrieval: The GET NEXT (GN) Function, 603
- GN Status Codes, 603*
  - Unqualified Calls: Retrieval of Every Segment in a Database, 604*
  - Unqualified SSA Calls: Retrieving All the Segments of a Given Type, 606*
  - Qualified SSA Calls: Accessing Specific Segment Occurrences for a Given Type, 607*
- 12.8 Sequential Retrieval of Children: The GET NEXT within Parent (GNP) Function, 607
- GNP Status Codes, 608*
  - Establishing Parentage, 608*
  - Unqualified Calls: Retrieval of All Children, 608*
  - Unqualified SSA Calls: Retrieving All the Children of a Specific Type, 610*
  - Qualified SSA Calls: Retrieval of Specified Child Occurrences, 613*
- 12.9 Adding New Segments: The INSERT (ISRT) Function, 614
- The ISRT Status Codes, 616*
  - Inserting a Child for an Existing Parent, 616*
  - Adding a Child When the Parent Hasn't Been Stored Yet: The Path Call, 617*
- 12.10 Updating Segments: The REPLACE (REPL) Function, 617
- REPL Status Codes, 618*
- 12.11 Deleting Segments: The DELETE (DLET) Function, 619
- 12.12 Data Integrity: The CHECKPOINT (CHKP) Function and IMS Logs, 619
- 12.13 Rolling Back Updates: The ROLL and ROLB Functions, 620
- 12.14 IMSVS86: IMS for the Microcomputer, 620
- 12.15 Summary, 621
- Episode 8: A Program that Uses DL/I to Print an IV Order Report, 624



## **PART VI Database Management 633**

### **Chapter 13 Database Administration 634**

Minicase: Telecorp, 634

#### **13.1 Database Administration Functions, 636**

*Administrative Responsibilities, 637*

*Technical Responsibilities, 653*

#### **13.2 Concurrency Control, 660**

#### **13.3 Organizing the Database Administration Function, 662**

*Diffused Responsibility, 663*

*Staff Organization, 663*

*Functional Organization, 663*

*Line Organization, 665*

#### **13.4 Summary, 665**

### **Chapter 14 Database Efficiencies: Data Organization and Access 667**

Minicase: Tri-State Gas and Electric, 667

#### **14.1 Overview of Distributed Databases, 669**

#### **14.2 Distributed Database Goals, 674**

#### **14.3 Partitioning a Database, 675**

*Horizontal Partitioning, 675*

*Vertical Partitioning, 678*

#### **14.4 Replicating a Database, 679**

*The Master/Slave Technique, 679*

*The Two-Phase Commit, 679*

#### **14.5 Advantages of a Distributed Database, 680**

*More Efficient Processing, 680*

*Improved Accessibility, 680*

*Local Control of Data, 680*

*Improved System Reliability, 681*

#### **14.6 Disadvantages of a Distributed Database, 681**

*Updating Problems, 681*

*Query Resolution, 681*

*More Complex Recovery Procedures, 682*

*Performance Degradation, 682*

#### **14.7 The Role of Microcomputers in a Distributed Environment, 683**

*LAN DDBMs, 683*

*The Microcomputer as a Terminal, 684*

#### **14.8 Database Machines, 686**

#### **14.9 The 4GL Environment: INFORMIX-4GL and ORACLE, 688**

*INFORMIX-4GL, 688*

*ORACLE, 705*

#### **14.10 Summary, 707**

## **Appendix A: Disk-Based Input/Output 710**

### **A.1 The CPU, 710**

### **A.2 The Operating System, 711**

### **A.3 Channels, 712**

### **A.4 Access Methods, 713**

*Basic Access Methods, 714*

*Queued Access Methods: Blocking and Buffering, 714*

*Virtual Storage Access Method, 717*

### **A.5 Interface Equipment: Getting Data from Here to There, 717**

*Channels, 717*

*Controllers, 717*

*The Role of the Interface Equipment: Putting It All Together, 718*

### **A.6 The Disk System, 719**

*The Disk Drive or Spindle: The Device, 719*

*The Disk Pack: The Medium, 719*

*The Access Mechanism, 723*

*Data Storage, 731*

### **A.7 Track Capacities, 736**

### **A.8 Summary, 737**



## **Appendix B: File Organization and Access: Traditional Storage and Retrieval of Data 741**

- B.1 File Characteristics, 741
- B.2 File Organization versus File Access:  
An Introduction, 743
  - Access, 743
  - Organization, 745
- B.3 Sequential Organization and Access:  
Key Sequence Storage, 747
  - Sequential Organization, 748
  - Processing Sequential Files, 749
- B.4 Indexed Sequential Organization and  
Access: Flexibility with a Price, 750
  - Prime Data Area, 751
  - The Cylinder Index, 751
  - Track Indexes, 753
  - Overflow Areas, 754

- B.5 Direct Organization, 757
  - Physical Organization, 758
  - Direct Addressing and Processing, 760
  - Indirect Addressing and Processing, 760
- B.6 Virtual Storage Access Method, 767
  - VSAM Organizations, 769
  - Accessing a VSAM, KSDS: A Theoretical  
Approach, 771
  - Adding Records to a KSDS, 773
  - Advantages of VSAM, 775
- B.7 Summary, 775

## **Annotated Bibliography 778**

## **Glossary G-1**

## **Index I-1**