CONTENTS

Preface xix

Part I The Context of Database Management 1
An Overview of Part One 1

Chapter 1 The Database Environment and Development Process 2
Learning Objectives 2
Introduction 2
Basic Concepts and Definitions 4
Data 4
Data Versus Information 5
Metadata 6
Traditional File Processing Systems 7
Disadvantages of File Processing Systems 7
Program-Data Dependence 7
Duplication of Data 8
Limited Data Sharing 8
Lengthy Development Times 8
Excessive Program Maintenance 8
The Database Approach 8
Data Models 8
Entities 8
Relationships 9
Relational Databases 10
Database Management Systems 10
Advantages of the Database Approach 11
Program-Data Independence 11
Planned Data Redundancy 11
Improved Data Consistency 11
Improved Data Sharing 11
Increased Productivity of Application Development 12
Enforcement of Standards 12
Improved Data Quality 12
Improved Data Accessibility and Responsiveness 13
Reduced Program Maintenance 13
Improved Decision Support 13
Cautions About Database Benefits 13
Costs and Risks of the Database Approach 13
New, Specialized Personnel 14
Installation and Management Cost and Complexity 14
Conversion Costs 14
Need for Explicit Backup and Recovery 14
Organizational Conflict 14
Components of the Database Environment 14
The Database Development Process 16
Contents

Part II  Database Analysis  43

An Overview of Part Two  43

Chapter 2  Modeling Data in the Organization  45

Learning Objectives  45
Introduction  45
The E-R Model: An Overview  47
Sample E-R Diagram  47
E-R Model Notation  49
Modeling the Rules of the Organization  50
Data Names and Definitions  50
DATA NAMES  51
DATA DEFINITIONS  51
GOOD DATA DEFINITIONS  51
Modeling Entities and Attributes  53
Entities  53
ENTITY TYPE VERSUS ENTITY INSTANCE  53
ENTITY TYPE VERSUS SYSTEM INPUT, OUTPUT, OR USER  54
STRONG VERSUS WEAK ENTITY TYPES  55
NAMING AND DEFINING ENTITY TYPES  55
Attributes 57
- Required Versus Optional Attributes 57
- Simple Versus Composite Attributes 58
- Single-Valued Versus Multivalued Attributes 59
- Stored Versus Derived Attributes 59
- Identifier Attribute 59
- Naming and Defining Attributes 61

Modeling Relationships 62
- Basic Concepts and Definitions in Relationships 63
  - Attributes on Relationships 63
  - Associative Entities 65
- Degree of a Relationship 66
  - Unary Relationship 67
  - Binary Relationship 67
  - Ternary Relationship 68
- Attributes or Entity? 69
- Cardinality Constraints 71
  - Minimum Cardinality 71
  - Maximum Cardinality 72
- Some Examples of Relationships and Their Cardinalities 72
  - A Ternary Relationship 73
- Modeling Time-Dependent Data 74
- Modeling Multiple Relationships Between Entity Types 75
- Naming and Defining Relationships 76

E-R Modeling Example: Pine Valley Furniture Company 78

Chapter 3 The Enhanced E-R Model 93

Learning Objectives 93

Introduction 93

Representing Supertypes and Subtypes 94

Basic Concepts and Notation 94

An Example of a Supertype/Subtype Relationship 95

Attribute Inheritance 96

When to Use Supertype/Subtype Relationships 97

Representing Specialization and Generalization 97

Generalization 98

Specialization 99

Combining Specialization and Generalization 100

Specifying Constraints in Supertype/Subtype Relationships 100

Specifying Completeness Constraints 101

Total Specialization Rule 101

Partial Specialization Rule 101
Specifying Disjointness Constraints 101
  DISJOINT RULE 102
  OVERLAP RULE 102
Defining Subtype Discriminators 102
  DISJOINT SUBTYPES 103
  OVERLAPPING SUBTYPES 103
Defining Supertype/Subtype Hierarchies 105
  AN EXAMPLE OF A SUPERTYPE/SUBTYPE HIERARCHY 105
  SUMMARY OF SUPERTYPE/SUBTYPE HIERARCHIES 106
EER Modeling Example: Pine Valley Furniture Company 106
Packaged Data Models 110
  A Revised Data Modeling Process with Packaged Data Models 110
Summary of Prepackaged Data Models 112
  Summary 112 • Key Terms 113 • Review Questions 113 •
  Problems and Exercises 114 • References 116 • Further
  Reading 117 • Web Resources 117

Part III Database Design 119
An Overview of Part Three 119
Chapter 4 Logical Database Design and the Relational Model 121
  Learning Objectives 121
  Introduction 121
  The Relational Data Model 122
    Basic Definitions 122
      RELATIONAL DATA STRUCTURE 123
      RELATIONAL KEYS 123
      PROPERTIES OF RELATIONS 124
      REMOVING MULTIVALUED ATTRIBUTES FROM TABLES 124
  Sample Database 124
  Integrity Constraints 126
    Domain Constraints 126
    Entity Integrity 126
    Referential Integrity 128
  Creating Relational Tables 129
    Well-Structured Relations 130
  Transforming EER Diagrams into Relations 131
    Step 1: Map Regular Entities 132
      COMPOSITE ATTRIBUTES 132
      MULTIVALUED ATTRIBUTES 133
    Step 2: Map Weak Entities 133
      WHEN TO CREATE A SURROGATE KEY 134
    Step 3: Map Binary Relationships 135
      MAP BINARY ONE-TO-MANY RELATIONSHIPS 135
      MAP BINARY MANY-TO-MANY RELATIONSHIPS 135
      MAP BINARY ONE-TO-ONE RELATIONSHIPS 136
    Step 4: Map Associative Entities 137
      IDENTIFIER NOT ASSIGNED 137
      IDENTIFIER ASSIGNED 138
Step 5: Map Unary Relationships 139
  UNARY ONE-TO-MANY RELATIONSHIPS 139
  UNARY MANY-TO-MANY RELATIONSHIPS 139
Step 6: Map Ternary (and n-ary) Relationships 141
Step 7: Map Supertype/Subtype Relationships 141
Summary of EER-to-Relational Transformations 143

Introduction to Normalization 144
Steps in Normalization 145
  Functional Dependencies and Keys 146
    DETERMINANTS 146
    CANDIDATE KEYS 146
Normalization Example: Pine Valley Furniture Company 148
Step 0: Represent the View in Tabular Form 148
Step 1: Convert to First Normal Form 149
  REMOVE REPEATING GROUPS 149
  SELECT THE PRIMARY KEY 149
  ANOMALIES IN 1NF 150
Step 2: Convert to Second Normal Form 150
Step 3: Convert to Third Normal Form 151
  REMOVING TRANSITIVE DEPENDENCIES 152
Determinants and Normalization 152
Merging Relations 153
An Example 154
View Integration Problems 154
  SYNONYMS 154
  HOMONYMS 155
  TRANSITIVE DEPENDENCIES 155
  SUPERTYPE/SUBTYPE RELATIONSHIPS 156
Summary 156 • Key Terms 157 • Review Questions 157 • Problems and Exercises 158 • References 165 • Further Reading 165 • Web Resources 165

Chapter 5 Physical Database Design and Performance 166
Learning Objectives 166
Introduction 166
The Physical Database Design Process 167
  Physical Database Design As a Basis for Regulatory Compliance 168
Designing Fields 169
  Choosing Data Types 169
    CODING TECHNIQUES 170
    HANDLING MISSING DATA 171
Denormalizing Data 171
  Denormalization 172
    OPPORTUNITIES FOR AND TYPES OF DENORMALIZATION 172
    DENORMALIZE WITH CAUTION 174
Designing Physical Database Files 175
File Organizations 176
  SEQUENTIAL FILE ORGANIZATIONS 178
  INDEXED FILE ORGANIZATIONS 179
  HASHED FILE ORGANIZATIONS 180
Chapter 7 Advanced SQL 232
Learning Objectives 232
Introduction 232
Processing Multiple Tables 233
   Equi-Join 234
   Natural Join 235
   Outer Join 236
   Sample Join Involving Four Tables 238
   Self-Join 239
   Subqueries 241
   Correlated Subqueries 246
   Using Derived Tables 247
   Combining Queries 248
Tips for Developing Queries 250
   Guidelines for Better Query Design 252
Ensuring Transaction Integrity 253
Data Dictionary Facilities 255
Triggers and Routines 256
   Triggers 256
   Routines 258
   Example Routine in Oracle's PL/SQL 259
Embedded SQL and Dynamic SQL 261
   Summary 262 • Key Terms 263 • Review Questions 263 • Problems and Exercises 264 • References 267 • Further Reading 267 • Web Resources 267

Chapter 8 Database Application Development 268
Learning Objectives 268
Introduction 268
Client/Server Architectures 269
Databases in a Two-Tier Architecture 271
   A VB.NET Example 272
   A Java Example 274
Three-Tier Architectures 275
Web Application Components 278
Databases in Three-Tier Applications 278
   A JSP Web Application 279
   An ASP.NET Example 282
Key Considerations in Three-Tier Applications 283
   Stored Procedures 284
   Transactions 285
   Database Connections 285
   Key Benefits of Three-Tier Applications 285
   Cloud Computing and Three-Tier Applications 287
Extensible Markup Language (XML) 288
   Storing XML Documents 290
   Retrieving XML Documents 290
Chapter 9 Data Warehousing 300

Learning Objectives 300

Introduction 300

Basic Concepts of Data Warehousing 302

A Brief History of Data Warehousing 303

The Need for Data Warehousing 303

Need for a Company-Wide View 303

Need to Separate Operational and Informational Systems 306

Data Warehouse Architectures 306

Independent Data Mart Data Warehousing Environment 306

Dependent Data Mart and Operational Data Store Architecture: A Three-Level Approach 308

Logical Data Mart and Real-Time Data Warehouse Architecture 310

Three-Layer Data Architecture 312

Role of the Enterprise Data Model 313

Role of Metadata 313

Some Characteristics of Data Warehouse Data 314

Status Versus Event Data 314

Transient Versus Periodic Data 315

An Example of Transient and Periodic Data 315

Transient Data 315

Periodic Data 316

Other Data Warehouse Changes 317

The Derived Data Layer 317

Characteristics of Derived Data 318

The Star Schema 318

Fact Tables and Dimension Tables 319

Example Star Schema 320

Surrogate Key 321

Grain of the Fact Table 322

Duration of the Database 322

Size of the Fact Table 323

Modeling Date and Time 324

Multiple Fact Tables 324

Hierarchies 325

Slowly Changing Dimensions 328

Determining Dimensions and Facts 329

Big Data and Columnar Databases 331

If You Knew SQL Like I NoSQL 332

The User Interface 333

Role of Metadata 334

SQL OLAP Querying 334