An Introduction to Database Systems

VOLUME I

Fachbereichsbibliothek Informatik
TU Darmstadt
59405438

Fourth Edition

C. J. DATE

Technische Hochschule Darmstadt
FACHBEREICH INFORMATIK
BIBLIOTHEK
Inventar-Nr.: 6395
Sachgebiete:
Standort:

ADDISON-WESLEY PUBLISHING COMPANY

Reading, Massachusetts • Menlo Park, California
Don Mills, Ontario • Wokingham, England • Amsterdam • Bonn • Sydney
Singapore • Tokyo • Madrid • Bogota • Santiago • San Juan

Contents

PART 1
Basic Concepts 1

1 An Overview of Database Management 3

1.1 An introductory example 3
1.2 What is a database system? 5
1.3 Operational data 9
1.4 Why database? 12
1.5 Data independence 15
1.6 Relational systems and others 19
Exercises 21
References and bibliography 23
Answers to selected exercises 25

2 An Architecture for a Database System 29

2.1 Purpose 29
2.2 The three levels of the architecture 29
2.3 The external level 32
2.4 The conceptual level 35
2.5 The internal level 36
2.6 Mappings 37
2.7 The database administrator 37
2.8 The database management system 40
2.9 Data communications 40
Exercises 41
References and bibliography 41

3 The Internal Level 45

3.1 Introduction 45
3.2 Database access: an overview 47
3.3 Page sets and files 51
3.4 Indexing 58
3.5 Hashing 67
3.6 Pointer chains 72
3.7 Compression techniques 75
3.8 Concluding remarks 80
Exercises 80
References and bibliography 82
Answers to selected exercises 88

PART 2
A Sample Relational System 93

4 An Overview of DB2 95

4.1 Background 95
4.2 Relational databases 96
4.3 The SQL language 100
4.4 Major system components 107
References and bibliography 113

5 Data Definition 117

5.1 Introduction 117
5.2 Base tables 117
5.3 Indexes 121
5.4 Concluding remarks 123
Exercises 123
Answers to selected exercises 124

6 Data Manipulation 127

6.1 Introduction 127
6.2 Simple queries 128
6.3 Join queries 132
6.4 Built-in functions 136
6.5 Advanced features 139
6.6 Update operations 147
6.7 Concluding remarks 151
Exercises 154
Answers to selected exercises 157

7 The System Catalog 165

7.1 Introduction 165
7.2 Querying the catalog 167
7.3 Updating the catalog 168
Exercises 169
Answers to selected exercises 170

8 Views 173

8.1 Introduction 173
8.2 View definition 176
8.3 DML operations on views 178
8.4 Logical data independence 182
8.5 Advantages of views 184
Exercises 185
Answers to selected exercises 185

9 Embedded SQL 189

9.1 Introduction 189
9.2 Operations not involving cursors 192
9.3 Operations involving cursors 193
9.4 A comprehensive example 196
9.5 Dynamic SQL 199
9.6 Concluding remarks 201
Exercises 202
Answers to selected exercises 203

PART 3
Relational Database Management 207

10 An Overview of INGRES 209

10.1 Background 209
10.2 Data definition 210
10.3 Data manipulation: retrieval operations 214
10.4 Data manipulation: update operations 221
10.5 Views 223
10.6 Embedded QUEL 224
Exercises 227
References and bibliography 228
Answers to selected exercises 228

11 Relational Data Structure 233

11.1 An introductory example 233
11.2 Domains 235
11.3 Relations 239
11.4 Relational databases 245
References and bibliography 245

12 Relational Integrity Rules 249

12.1 Introduction 249
12.2 Primary keys 249
12.3 Foreign keys 250
12.4 The two integrity rules 252
12.5 Implications of the two integrity rules 253
References and bibliography 255

13 Relational Algebra 257

13.1 Introduction 257
13.2 A syntax for the relational algebra 260
13.3 Traditional set operations 262
13.4 Special relational operations 264
13.5 Examples 270
13.6 Concluding remarks 270
Exercises 272
References and bibliography 273
Answers to selected exercises 276

14 Relational Calculus 281

14.1 Introduction 281
14.2 Tuple-oriented relational calculus 283
14.3 Examples 287
14.4 Relational calculus vs. relational algebra 290
14.5 Domain-oriented relational calculus 294
14.6 Query-By-Example 297
Exercises 301
References and bibliography 302
Answers to selected exercises 303

15 Relational Systems 313

15.1 The relational model: a summary 313
15.2 Essentiality 315

15.3 Definition of a relational system 320
15.4 Enforcing the relational integrity rules 322
15.5 Some current relational products 325
Exercises 327
References and bibliography 327
Answers to selected exercises 329

16 Query Optimization 333

16.1 Introduction 333
16.2 A simple example 334
16.3 The optimization process: an overview 335
16.4 Optimization in System R 340
16.5 Optimization in INGRES 343
16.6 Implementing the relational operators 345
References and bibliography 350

17 Further Normalization 361

17.1 Introduction 361
17.2 Functional dependence 364
17.3 First, second, and third normal forms 366
17.4 Boyce/Codd normal form 374
17.5 Good and bad decompositions 378
17.6 Fourth normal form 381
17.7 Fifth normal form 385
17.8 Concluding remarks 390
Exercises 392
References and bibliography 394
Answers to selected exercises 402

PART 4
The Database Environment 411

18 Recovery and Concurrency 413

18.1 Introduction 413
18.2 Transaction recovery 413
18.3 System and media recovery 417
18.4 Three concurrency problems 419
18.5 Locking 422
18.6 Concluding remarks 427
Exercises 428
References and bibliography 430
Answers to selected exercises 434

19 Security and Integrity 437

19.1 Introduction 437
19.2 Security: general considerations 438
19.3 Security in SQL 439
19.4 Other aspects of security 443
19.5 Integrity: general considerations 444
19.6 A hypothetical integrity language 446
19.7 Security and integrity in INGRES 450
19.8 Concluding remarks 453
Exercises 454
References and bibliography 455
Answers to selected exercises 457

20 The Database Product Family 461

20.1 Introduction 461
20.2 Data access 462
20.3 Data presentation 470
20.4 Application generation 477
Exercises 481
References and bibliography 481

PART 5
Nonrelational Systems 483

21 An Inverted List System: DATACOM/DB 487

21.1 Background 487
21.2 The inverted list model 488
21.3 An overview of DATACOM/DB 490
21.4 Data definition 492
21.5 Data manipulation 496
21.6 The Compound Boolean Selection feature 499
Exercises 501
References and bibliography 502

22 A Hierarchic System: IMS 503

22.1 Background 503
22.2 The hierarchic model 503
22.3 An overview of IMS 508

22.4 Data definition 511
22.5 Data manipulation 513
22.6 Storage structure 520
22.7 Logical databases 524
22.8 Secondary indexes 529
22.9 Concluding remarks 533
Exercises 536
References and bibliography 537
Answers to selected exercises 538

23 A Network System: IDMS 541

23.1 Background 541
23.2 The network model 542
23.3 An overview of IDMS 548
23.4 Data definition 549
23.5 Data manipulation 552
23.6 Storage structure 561
23.7 The Logical Record Facility 563
23.8 The Automatic System Facility 568
23.9 Concluding remarks 574
Exercises 577
References and bibliography 579
Answers to selected exercises 579

PART 6
Some Current Developments 585

24 Distributed Systems 587

24.1 Introduction 587
24.2 Objectives and rationale 589
24.3 Problems of distributed systems 591
24.4 Some sample systems 595
24.5 Other types of distributed system 600
24.6 Concluding remarks 604
Exercises 605
References and bibliography 605

25 Semantic Modeling 609

25.1 What is the problem? 609
25.2 The overall approach 610

25.3 The entity/relationship model 610
25.4 The extended relational model RM/T 615
25.5 Concluding remarks 621
Exercises 622
References and bibliography 622

List of Acronyms 625

Index 627