Date on Database

Writings 2000–2006

C. J. Date

Apress®

Date on Database: Writings 2000–2006

Copyright © 2006 by C. J. Date

All rights reserved. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage or retrieval system, without the prior written permission of the copyright owner and the publisher.

ISBN-13 (pbk): 978-1-59059-746-0

ISBN-10 (pbk): 1-59059-746-X

Printed and bound in the United States of America 9 8 7 6 5 4 3 2 1

Trademarked names may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, we use the names only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

Lead Editor: Jonathan Gennick
Editorial Board: Steve Anglin, Ewan Buckingham, Gary Cornell, Jason Gilmore, Jonathan Gennick, Jonathan Hassell, James Huddleston, Chris Mills, Matthew Moodie, Dominic Shakeshaft, Jim Sumser, Keir Thomas, Matt Wade
Project Manager: Tracy Brown Collins
Copy Edit Manager: Nicole LeClerc
Assistant Production Director: Kari Brooks-Copony
Production Editor: Kelly Winquist
Compositor: Susan Glinert
Proofreader: Lori Bring
Indexer: C. J. Date
Cover Designer: Kurt Krames
Manufacturing Director: Tom Debolski

Distributed to the book trade worldwide by Springer-Verlag New York, Inc., 233 Spring Street, 6th Floor, New York, NY 10013. Phone 1-800-SPRINGER, fax 201-348-4505, e-mail `orders-ny@springer-sbm.com`, or visit `http://www.springeronline.com`.

For information on translations, please contact Apress directly at 2560 Ninth Street, Suite 219, Berkeley, CA 94710. Phone 510-549-5930, fax 510-549-5939, e-mail `info@apress.com`, or visit `http://www.apress.com`.

The information in this book is distributed on an “as is” basis, without warranty. Although every precaution has been taken in the preparation of this work, neither the author(s) nor Apress shall have any liability to any person or entity with respect to any loss or damage caused or alleged to be caused directly or indirectly by the information contained in this work.

Contents

About the Author xix
Acknowledgments xxiii
Preface xxv

PART 1 Some Preliminaries

CHAPTER 1 Edgar F. Codd: A Tribute and Personal Memoir 3

Background 3
Database Contributions 4
Other Contributions 8
Personal Memories 9
Concluding Remarks 10

CHAPTER 2 An Interview with Chris Date 13

CHAPTER 3 Good Writing Does Matter 23

Preamble 23
Exhibit A 24
Restriction Conditions in DB2 27
A Suggested Replacement for Exhibit A 28
Exhibit B 29
Exhibit C 31
Concluding Remarks 31

PART 2 And Now for Something Completely Different

CHAPTER 4 On the Notion of Logical Difference 35

Preamble 35
Why Is It Important to Think Precisely? 35
Logical Differences Are Big Differences 36
Some Things Are Much the Same 37
Object Orientation 39
The Unified Modeling Language 41
Miscellaneous Examples 42
Concluding Remarks 43
Acknowledgments 43
Appendix A: Some Quotes from the Literature 43

CHAPTER 5 On the Logical Difference Between Model and Implementation 51

Terms and Definitions 52
What Are Models For? 53
Definitions from the Literature 54
Examples of Confusion: Performance Issues 55
Examples of Confusion: Many Different Data Structures 58
Examples of Confusion: Objects 62
Examples of Confusion: Miscellaneous 64
Consequences of Such Confusions 66
Concluding Remarks 67
Acknowledgments 67

CHAPTER 6 On the Logical Differences Between Types, Values, and Variables 69

Types Are Fundamental 71
More on Types 73
Values and Variables 76
Values and Variables Can Be Arbitrarily Complex 77
Variables Are Updatable, Values Aren't 78

Pseudovariables 79
Variables Have Addresses, Values Don't 82
Relation Values and Variables 84
Examples of Confusion: Values *vs.* Variables 87
Logical Sameness 91
A Hypothesis 92
Examples of Confusion: Types *vs.* Values and/or Variables 92
Consequences of Such Confusions 97
Concluding Remarks 100

PART 3 **Relational Database Management**

CHAPTER 7 Why We Need Type BOOLEAN 105

CHAPTER 8 What First Normal Form Really Means 107

Some Preliminaries 108
"Data Value Atomicity" 109
Relation-Valued Attributes 113
Domains Can Contain *Anything!* 115
Relation-Valued Attributes Make Outer Join Unnecessary 116
Relation-Valued Attributes in Base Tables?—The Bad News 118
Relation-Valued Attributes in Base Tables?—The Good News 121
A Remark on "NF² Relations" 126
A Definition 127
Concluding Remarks 130
Acknowledgments 132
Appendix A: Terminology 132
Appendix B: *The Information Principle* 133
Appendix C: Definitions from the Literature 135
Appendix D: So What About "Multi-Value Systems"? 136
Appendix E: Formal Definitions 137

CHAPTER 9 A Sweet Disorder 139

SQL Tables 139
The "Select *" Problem 140
Duplicate Column Names 142
Missing Column Names 143

Implications for the UNION Operator 146
Type *vs.* Degree 148
Which Union? 148
Relation Types 150
Predicates Revisited 154
Where Did Column Ordering Come From? 155
Concluding Remarks 156
Acknowledgments 156
Appendix A: Row Comparisons 157

CHAPTER 10 Double Trouble, Double Trouble 159

The Cat Food Example 160
Expression Transformation 162
Rows Represent Propositions 167
The Conclusion So Far 167
Rats and Ratlets 168
The Pennies Example 168
Bag Theory 170
The Performance Issue 170
Some More SQL Problems 171
Concluding Remarks 172
Appendix A: A Tuple-Bag Algebra? 174
Appendix B: More on Counting Pennies 178

CHAPTER 11 Multiple Assignment 181

A Little History 182
Assigning to Several Variables at Once 186
A Multiple Assignment Example 189
Why We Need Multiple Assignment 192
Why Integrity Checking Must Be Immediate 195
Assigning to a Variable More Than Once 197
Semantics 198
Syntax 201
Assigning to Several Components at Once 203
Concluding Remarks 205

Appendix A: Multiple Assignment in SQL 205
Overview 206
The SET Statement: Single Assignment 207
The SET Statement: Multiple Assignment 211
The SQL UPDATE Statement Revisited 214

CHAPTER 12 Data Redundancy and Database Design 217

The Running Example 219
Some Prerequisites 220
Database Design Is Predicate Design 222
A Brief Review of Normalization 222
Normalization Is Not Enough 226
Orthogonality (I) 227
Another Kind of Decomposition 229
Tuples *vs.* Propositions 231
Orthogonality (II) 234
What About Views? 237
Violating Orthogonality 238
But Isn't It All Just Common Sense? 240
Other Kinds of Redundancy 241
Managing Redundancy 251
1. Raw Design Only 251
2. Declare the Constraint 251
3. Use a View 252
4. Use a Snapshot 252
Concluding Remarks 253

CHAPTER 13 Data Redundancy and Database Design: Further Thoughts Number One 255

More on Predicates 256
Identity Decompositions 257
More on Propositions 258
More on Normalization 260
More on Dependencies 262
More on Dependency Preservation 264
Generalizing Key Constraints 267
Orthogonality Revisited 268
Concluding Remarks 270

CHAPTER 14 Tree-Structured Data 271

Definitions 271
A Tree-Based Sort Technique 273
How Treesort Works 274
Data Structures 275
Relational Operators 276
Implementing TREESORT 277
Integrity Constraints 278
Traversing the Nodes of a Tree 280
Bill-of-Materials Processing 283
Another Kind of Tree 291
What About Relation-Valued Attributes? 294
Concluding Remarks 297
Acknowledgments 297

CHAPTER 15 Twelve Rules for Business Rules 299

Assumptions 299
A Note on Terminology 300
Further Preliminary Remarks 301
Prescription 1: Executable Rules 302
Prescription 2: Declarative Rules 302
Prescription 3: Kinds of Rules 302
Prescription 4: Declaration Sequence *vs.* Execution Sequence 304
Prescription 5: The Rule Engine Is a DBMS 304
Prescription 6: The Rule Engine Is a *Relational* DBMS 305
Prescription 7: Rule Management 306
Prescription 8: Kinds of Constraints 307
Prescription 9: Extensibility 308
Prescription 10: Platform Independence 308
Prescription 11: No Subversion 308
Prescription 12: Full Automation 309
Summary 309
Acknowledgments 310
References 310

PART 4 SQL Database Management

CHAPTER 16 Two Remarks on SQL's UNION 313

UNION in the Relational Model 313
Data Type Conversion 314
Duplicate Rows 315
Concluding Remarks 317

CHAPTER 17 A Cure for Madness 319

The Syntax and Semantics of SELECT - FROM - WHERE 321
Model *vs.* Implementation 322
Nested Functions 323
Further Thoughts 324
A Final Observation 326

CHAPTER 18 Why Three- and Four-Valued Logic Don't Work 329

Some Basic Concepts 329
Tautologies, Identities, and Contradictions 331
De Morgan's Laws 332
Codd's Three-Valued Logic 332
The Unknown Truth Value and A-marks 334
How Many 3VLs Are There? 335
Codd's Four-Valued Logics 336
4VL Number 1 337
4VL Number 2 339
4VL Number 3 340
Some Questions of Intuition 340

PART 5 Further Relational Misconceptions

CHAPTER 19 There's Only One Relational Model 345

Celko's Preamble 347
"Chris Date = No Duplicates, No NULLs" 347
"E. F. Codd, RM Version I" 349
"E. F. Codd, RM Version II" 350
"SQL-92 = Duplicates, One NULL" 350
"Duplicates, One NULL, Non-1NF Tables" 351
"Rick Snodgrass = Temporal SQL" 352
How Many Relational Models? 353
Concluding Remarks 354
Technical Correspondence 355

CHAPTER 20 The Relational Model Is Very Much Alive! 361

Some Specific Issues 361
Concluding Remarks 365
Technical Correspondence 365

CHAPTER 21 What Do You Mean, "Post-Relational"? 369

What About Object Technology? 370
Further Points Arising 370
What About Object/Relational Technology? 374
Technical Correspondence 375

CHAPTER 22 A Database Disconnect 383

The SMART Paper 383
The Version 8 Paper 384
Conclusion 386

PART 6 Subtyping and Inheritance

CHAPTER 23 Is a Circle an Ellipse? 389

Preamble 389
A Little Background 392
The Main Argument 392
Software Engineering *vs.* Conceptual Modeling? 397
What About "Colored Circles"? 398
A Type Inheritance Proposal 399
Technical Correspondence 400

CHAPTER 24 What Does Substitutability Really Mean? 409

What Is LSP? 410
The Liskov Substitution Principle 414
A Behavioral Notion of Subtyping 417
The Liskov/Wing Paper: Assumptions 418
The Liskov/Wing Paper: Objectives 419
The Liskov/Wing Inheritance Model 421
The Running Example 427
Preconditions and Postconditions 428
Covariance and Contravariance 429
Extending the Example 429
Extension and Constraint Subtypes 430
Substitutability Defined 432

PART 7 Relational *vs.* Nonrelational Systems

CHAPTER 25 Models, Models, Everywhere, Nor Any Time to Think 437

"Modeling in the UML" 438
What *Does* "Model" Mean? 440
Miscellaneous Observations 442
Concluding Remarks 444
Postscript 444

CHAPTER 26 Basic Concepts in UML: A Request for Clarification 445

Some Immediate Discoveries 445
The OCL Book 447
The UML Book 451
Types, Classes, and Domains 452
Further Puzzles 453
Concluding Remarks 455

CHAPTER 27 A Comparison Between ODMG and *The Third Manifesto* 457

Preamble 457
Overview 460
Objects 460
Atomic Objects 460
Structured Objects 462
Collection Objects 462
Literals 464
Types, Classes, and Interfaces 465
Inheritance 466
Object Definition Language 468
Object Interchange Format 468
Object Query Language 468
Summary 469
RM Prescriptions 469
RM Proscriptions 473
OO Prescriptions 474
OO Proscriptions 475
RM Very Strong Suggestions 475
OO Very Strong Suggestions 476
IM Prescriptions 477
References 479

CHAPTER 28 An Overview and Analysis of Proposals Based on the TSQL2 Approach 481

TSQL2 and the SQL Standard 482
TSQL2 Tables 483
The Central Idea 486

Temporal Upward Compatibility 489
Current, Sequenced, and Nonsequenced Operations 493
Current Operations 494
Sequenced Operations 495
Nonsequenced Operations 498
Accessing the Hidden Columns 498
Data Definition Statements 499
Valid-Time Tables 499
Transaction-Time Tables 501
Bitemporal Tables 502
Statement Modifiers Are Flawed 502
Consequences of Hidden Columns 508
Lack of Generality 509
Imprecise Specification 510
Concluding Remarks 511
References 511

APPENDIX **The Role of the Trade Press in Educating the Professional Community: A Case Study** 515

Overview 515
The Saga Continues 516
Conclusions 517
Exhibit A 518
Exhibit B 519
Exhibit C 520
Exhibit D 521
Exhibit E 521
Exhibit F 522
Exhibit G 522
Exhibit H 523
Exhibit I 524
Exhibit J 524
Postscript 525

INDEX 527