Applied Mathematics for Database Professionals

Lex de Haan and Toon Koppelaars

Apress®

Applied Mathematics for Database Professionals

Copyright © 2007 by Lex de Haan and Toon Koppelaars

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: 978-1-59059-745-3

ISBN-10: 1-59059-745-1

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
Technical Reviewers: Chris Date, Cary Millsap
Editorial Board: Steve Anglin, Ewan Buckingham, Gary Cornell, Jonathan Gennick, Jason Gilmore, Jonathan Hassell, Chris Mills, Matthew Moodie, Jeffrey Pepper, Ben Renow-Clarke, Dominic Shakeshaft, Matt Wade, Tom Welsh
Project Manager: Tracy Brown Collins
Copy Edit Manager: Nicole Flores
Copy Editor: Susannah Davidson Pfalzer
Assistant Production Director: Kari Brooks-Copony
Production Editor: Kelly Winquist
Compositor: Dina Quan
Proofreader: April Eddy
Indexer: Brenda Miller
Artist: April Milne
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 2855 Telegraph Avenue, Suite 600, Berkeley, CA 94705. 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.

The source code for this book is available to readers at `http://www.apress.com` in the Source Code/Download section. You will need to answer questions pertaining to this book in order to successfully download the code.

Lex de Haan
1954–2006

"Tall and narrow, with a lot of good stuff on the upper floor"
"A missing value"

Contents at a Glance

Foreword xv
About the Authors xvii
About the Technical Reviewers xix
Acknowledgments xxi
Preface xxiii
Introduction xxv

PART 1 The Mathematics

CHAPTER 1 Logic: Introduction 3
CHAPTER 2 Set Theory: Introduction 23
CHAPTER 3 Some More Logic 47
CHAPTER 4 Relations and Functions 67

PART 2 The Application

CHAPTER 5 Tables and Database States 91
CHAPTER 6 Tuple, Table, and Database Predicates 117
CHAPTER 7 Specifying Database Designs 139
CHAPTER 8 Specifying State Transition Constraints 185
CHAPTER 9 Data Retrieval 199
CHAPTER 10 Data Manipulation 221

PART 3 The Implementation

CHAPTER 11 Implementing Database Designs in Oracle 241
CHAPTER 12 Summary and Conclusions 305

PART 4 Appendixes

APPENDIX A Formal Definition of Example Database 311
APPENDIX B Symbols 333
APPENDIX C Bibliography 335
APPENDIX D Nulls and Three (or More) Valued Logic 337
APPENDIX E Answers to Selected Exercises 347

INDEX 367

Contents

Foreword xv
About the Authors xvii
About the Technical Reviewers xix
Acknowledgments xxi
Preface xxiii
Introduction xxv

PART 1 ■■■ The Mathematics

■CHAPTER 1 Logic: Introduction 3

The History of Logic 4
Values, Variables, and Types 5
Propositions and Predicates 5
Logical Connectives 8
Simple and Compound Predicates 9
Using Parentheses and Operator Precedence Rules 10
Truth Tables 11
Implication 14
Predicate Strength 15
Going a Little Further 15
Functional Completeness 16
Special Predicate Categories 17
Tautologies and Contradictions 17
Modus Ponens and Modus Tollens 18
Logical Equivalences and Rewrite Rules 18
Rewrite Rules 19
Using Existing Rewrite Rules to Prove New Ones 20
Chapter Summary 21
Exercises 22

CHAPTER 2 Set Theory: Introduction 23

Sets and Elements 24
Methods to Specify Sets 26
Enumerative Method 26
Predicative Method 26
Substitutive Method 27
Hybrid Method 27
Venn Diagrams 28
Cardinality and Singleton Sets 29
Singleton Sets 29
The Choose Operator 30
Subsets 30
Union, Intersection, and Difference 31
Properties of Set Operators 33
Set Operators and Disjoint Sets 33
Set Operators and the Empty Set 34
Powersets and Partitions 35
Union of a Set of Sets 36
Partitions 37
Ordered Pairs and Cartesian Product 38
Ordered Pairs 38
Cartesian Product 39
Sum Operator 40
Some Convenient Shorthand Set Notations 41
Chapter Summary 41
Exercises 42

CHAPTER 3 Some More Logic 47

Algebraic Properties 47
Identity 48
Commutativity 48
Associativity 48
Distributivity 49
Reflexivity 49
Transitivity 49
De Morgan Laws 50
Idempotence 50
Double Negation (or Involution) 50
Absorption 50

Quantifiers 51
Quantifiers and Finite Sets 54
Quantification Over the Empty Set 54
Nesting Quantifiers 55
Distributive Properties of Quantifiers 57
Negation of Quantifiers 57
Rewrite Rules with Quantifiers 58
Normal Forms 59
Conjunctive Normal Form 59
Disjunctive Normal Form 60
Finding the Normal Form for a Given Predicate 61
Chapter Summary 63
Exercises 64

CHAPTER 4 Relations and Functions 67

Binary Relations 68
Ordered Pairs and Cartesian Product Revisited 68
Binary Relations 69
Functions 70
Domain and Range of Functions 71
Identity Function 74
Subset of a Function 74
Operations on Functions 75
Union, Intersection, and Difference 75
Limitation of a Function 76
Set Functions 77
Characterizations 78
External Predicates 79
The Generalized Product of a Set Function 79
A Preview of Constraint Specification 81
Function Composition 82
Chapter Summary 84
Exercises 85

PART 2 The Application

CHAPTER 5 Tables and Database States 91

Terminology 91
- Database Design 92
- Database Variable 92
- Database Universe 92
- Database State 93
- Database 93
- Database Management System (DBMS) 93
- Table Design 93
- Table Structure 93
- Table 94

Tables 94
- Formal Specification of a Table 94
- Shorthand Notation 96
- Table Construction 97

Database States 98
- Formal Representation of a Database State 98
- Database Skeleton 100

Operations on Tables 101
- Union, Intersection, and Difference 101
- Restriction 104
- Join 106
- Attribute Renaming 108
- Extension 109
- Aggregation 111

Chapter Summary 113

Exercises 114

CHAPTER 6 Tuple, Table, and Database Predicates 117

Tuple Predicates 118

Table Predicates 120

Database Predicates 124

A Few Remarks on Data Integrity Predicates 127

Common Patterns of Table and Database Predicates 127
Unique Identification Predicate 127
Subset Requirement Predicate 129
Specialization Predicate 132
Generalization 134
Tuple-in-Join Predicate 135
Chapter Summary 137
Exercises 138

CHAPTER 7 Specifying Database Designs 139

Documenting Databases and Constraints 140
The Layers Inside a Database Design 141
Top-Down View of a Database 141
Classification Schema for Constraints 142
Specifying the Example Database Design 143
Database Skeleton 144
Characterizations 147
Tuple Universes 151
Table Universes 156
Database Universe 167
Chapter Summary 182
Exercises 183

CHAPTER 8 Specifying State Transition Constraints 185

More Data Integrity Predicates 185
A Simple Example 186
State Transition Predicates 188
State Transition Constraints 190
State Transition Universe 190
Completing the Example Database Design 192
Chapter Summary 196
Exercises 197

CHAPTER 9 Data Retrieval 199

Formally Specifying Queries 199
Example Queries Over DB_UEX 201
A Remark on Negations 216
Chapter Summary 218
Exercises 219

CHAPTER 10 Data Manipulation 221

Formally Specifying Transactions 221
Example Transactions Over DB_UEX 226
Chapter Summary 236
Exercises 236

PART 3 The Implementation

CHAPTER 11 Implementing Database Designs in Oracle 241

Introduction 242
Window-on-Data Applications 243
Classifying Window-on-Data Application Code 243
Implementing Data Integrity Code 247
Alternative Implementation Strategies 248
Order of Preference 254
Implementing Table Structures 255
Implementing Attribute Constraints 259
Implementing Tuple Constraints 262
Table Constraint Implementation Issues 265
DI Code Execution Models 266
DI Code Serialization 284
Implementing Table Constraints 290
Implementing Database Constraints 291
Implementing Transition Constraints 296
Bringing Deferred Checking into the Picture 299
Why Deferred Checking? 299
Outline of Execution Model for Deferred Checking 300
The RuleGen Framework 303
Chapter Summary 304

CHAPTER 12 Summary and Conclusions 305

Summary 305
Conclusions 306

PART 4 Appendixes

APPENDIX A Formal Definition of Example Database 311

Bird's Eye Overview 312
Database Skeleton DB_S 313
Table Universe Definitions 314
Some Convenient Sets 315
Table Universe for EMP 315
Table Universe for SREP 317
Table Universe for MEMP 317
Table Universe for TERM 318
Table Universe for DEPT 319
Table Universe for GRD 319
Table Universe for CRS 321
Table Universe for OFFR 321
Table Universe for REG 322
Table Universe for HIST 323
Database Characterization DBCH 324
Database Universe DB_UEX 325
State Transition Universe TX_UEX 330

APPENDIX B Symbols 333

APPENDIX C Bibliography 335

Original Writings That Introduce the Methodology Demonstrated in This Book 335
Recommended Reading in the Area of the Underlying Mathematical Theories 335
Seminal Writings That Introduce the General Theory of Data Management 335
Recommended Reading on Relational Database Management 335
Research Papers on Implementing Data Integrity Constraints and Related Subjects 336
Previous Related Writings of the Authors 336

APPENDIX D Nulls and Three (or More) Valued Logic 337

To Be Applicable or Not 337
Inapplicable. 338
Not Yet Applicable 338
Nice to Know. 339
Implementation Guidelines. 340
Three (or More) Valued Logic 340
Unknown 340
Truth Tables of Three-Valued Logic 341
Missing Operators 343
Three-Valued Logic, Tautologies, and Rewrite Rules 343
Handling Three-Valued Logic. 344
Four-Valued Logic 345

APPENDIX E Answers to Selected Exercises 347

Chapter 1 Answers 347
Chapter 2 Answers 352
Chapter 3 Answers 353
Chapter 4 Answers 355
Chapter 5 Answers 356
Chapter 6 Answers 357
Chapter 7 Answers 358
Chapter 8 Answers 360
Chapter 9 Answers 361
Chapter 10 Answers 364

INDEX 367