

Relational Database:

Selected Writings

C. J. Date

Relational Database:

Selected Writings

C. J. Date



ADDISON-WESLEY PUBLISHING COMPANY

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

Library of Congress Cataloging-in-Publication Data

Date, C. J.

Relational database.

Bibliography: p.

Includes index.

1. Data base management. I. Title.
QA76.9.D3D3724 1986 005.75'6 85-22907
ISBN 0-201-14196-5

Copyright © 1986 by Addison-Wesley Publishing Company, Inc.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. Printed in the United States of America. Published simultaneously in Canada.

Preface

The purpose of this book is to bring together into one place all those of my database papers from the period 1974–1985 that seem to me to be worth preserving. Most of those papers were originally published in one of the professional database journals or in the proceedings of one of the technical database conferences (see below), and I am grateful to all concerned for permission to republish them in this fashion. I have of course taken the opportunity to make a number of small improvements and corrections in the versions printed here, but apart from such minor changes the papers are all essentially as originally published. However, I have also included a couple of papers that have not previously been published at all.

In a collection of this kind there is inevitably a certain unevenness of style—some of the papers are fairly formal in nature, others (marked generally by the use of the first person singular) much less so. There is also inevitably a certain amount of repetition. I have edited out the worst overlaps, but otherwise have generally tried to keep each paper reasonably self-contained.

The book is arranged into four principal parts:

- I. Relational Database Management
- II. Relational vs. Nonrelational Systems
- III. The SQL Language
- IV. Database Design

There is also an afterword, containing the text of an interview I gave to *Data Base Newsletter* in 1983.

Part I consists of a collection of papers on the relational model and relational systems in general. Part II is concerned with various aspects of the differences between relational and nonrelational systems. Part III contains a detailed critique of what is rapidly becoming the de facto standard in the relational field, namely the "Structured Query Language" SQL, and proposes a number of extensions and revisions to that language. Finally, Part IV presents a detailed methodology for relational database design. The sequence of papers within each part is very largely arbitrary.

Acknowledgments: As usual, I am delighted to acknowledge my debt to the many people involved, directly or indirectly, in the production of this book. First, as already stated, I am grateful to the copyright holders of the original papers for permission to republish those papers in this manner. Second, I am also grateful to the numerous reviewers of those original papers; individual acknowledgments to those reviewers appear within the papers themselves. Third, I am very pleased to acknowledge the encouragement I received from numerous close friends and colleagues while working on this book. Last, I am (as always) grateful to my editor, Elydia Siegel, and to the staff at Addison-Wesley for their assistance and professional expertise.

Saratoga, California

C.J.D.

Contents

PART I RELATIONAL DATABASE MANAGEMENT

CHAPTER 1

Relational Database: An Overview (AN INTRODUCTION TO RELATIONAL SYSTEM CONCEPTS, USING SQL/DS AS A CONCRETE EXAMPLE) **3**

Abstract	3
Comments on Republication	4
Introduction	4
User Perspective	6
System Structure	7
Data Manipulation	9
Views	12
Security	15
Conclusion	16
Further Reading	17

CHAPTER 2

An Informal Definition of the Relational Model **21**

Abstract	21
Comments on Republication	21

Relational Data Structure.....	22
Relational Data Integrity.....	25
Relational Data Manipulation	27
Summary	31

CHAPTER 3

Why Every Relation Should Have Exactly One Primary Key 33

Abstract	33
Comments on Republication	33
Definitions	34
A Note on Terminology.....	35
Why Primary Keys Are Important	35
The Current X3H2 Position.....	36
Position Statement.....	36
Justification	36

CHAPTER 4

Referential Integrity 41

Abstract	41
Comments on Republication	41
Introduction	42
Structure of the Paper	43
Referential Integrity in the Relational Model.....	44
A Language Proposal for General Integrity Rules	49
Key Constraints.....	51
Referential Constraints	52
The Reference Graph	58
Referential Integrity in DBTG.....	59
Summary	61
Acknowledgments.....	62
References.....	62

CHAPTER 5

On the Performance of Relational Database Systems 65

Abstract	65
Comments on Republication	65
Introduction	66

Performance of an Individual Transaction	67
Performance of the Overall System	71
Further Considerations	72
References	76

CHAPTER 6

Some Relational Myths Exploded (AN EXAMINATION OF SOME POPULAR MISCONCEPTIONS CONCERNING RELATIONAL DATABASE MANAGEMENT SYSTEMS)

	77
Abstract	77
Comments on Republication	77
Introduction	78
Myths 1-26	80
Conclusion	122
Acknowledgments	122
References	122

CHAPTER 7

A Formal Definition of the Relational Model **125**

Abstract	125
Comments on Republication	125
Introduction	126
Relational Database	127
Relational Operations	132
Relational Algebra: Additional Operators	137
Relational Rules	139
Concluding Remarks	140
Acknowledgments	140
References	140

CHAPTER 8

The Relational Model and Its Interpretation **143**

Abstract	143
Introduction	143
The Argument	145
Conclusion	150
References	150

PART II RELATIONAL VS. NONRELATIONAL SYSTEMS

CHAPTER 9

Interactive Support for Nonprogrammers:

The Relational and Network Approaches

153

Abstract	153
Comments on Republication	154
Introduction	154
The Two Approaches	155
Types of Database User and Usage	160
Status of Support for Nonprogrammers	162
Questions Concerning Owner-Coupled Sets	167
A Discipline	168
Conclusion	170
Historical Note	171
Acknowledgment	171
References	171
Appendix: A Comparative Example	173

CHAPTER 10

The Relational and Network Approaches:

Comparison of the Application Programming Interfaces

179

Abstract	179
Comments on Republication	180
Introduction	180
Simplicity	181
Uniformity	190
Completeness	192
Data Independence	195
Integrity and Security	198
Summary	200
Acknowledgment	200
References	201

CHAPTER 11

An Introduction to the Unified Database Language (UDL)

203

Abstract	203
Comments on Republication	204
Introduction	205

An Introductory Example.....	206
Direct Reference.....	207
The Approach to Commonality	210
Declarative Language	214
Cursors and Cursor States	219
Bound Cursors.....	221
Manipulative Language.....	222
Summary	227
Acknowledgments.....	228
References.....	228
Appendix A: PL/I-UDL Syntax (Simplified)	229
Appendix B: A CODASYL Version of the Enrollment Problem	233
Appendix C: A Note on Performance.....	236

CHAPTER 12

Why Is It So Difficult to Provide a Relational Interface to IMS?

Abstract	241
Comments on Republication	241
Introduction	242
Some Fundamental Incompatibilities	243
Defining a Flat Record View of IMS Data	247
What About Bypassing DL/I?.....	253
Concluding Remarks.....	255
Acknowledgments.....	257

PART III THE SQL LANGUAGE

CHAPTER 13

Some Principles of Good Language Design (WITH ESPECIAL REFERENCE TO THE DESIGN OF DATABASE LANGUAGES)

Abstract	261
Comments on Republication	261
Introduction	262
Orthogonality	262
Language Definition.....	265
Expressions	266
Miscellaneous Points	267
References.....	268

CHAPTER 14

A Critique of the SQL Database Language	269
Abstract	269
Comments on Republication	270
Introduction	270
Lack of Orthogonality: Expressions	271
Lack of Orthogonality: Built-in Functions	280
Lack of Orthogonality: Miscellaneous Items	287
Formal Definition	296
Mismatch with Host Languages	299
Missing Function	301
Mistakes	302
Aspects of the Relational Model Not Supported	305
Summary and Conclusions	308
Acknowledgments	309
References	309
Appendix: SQL Strongpoints	310

CHAPTER 15

Null Values in Database Management	313
Abstract	313
Comments on Republication	313
Introduction	314
A Scheme for Handling Nulls	316
Effect on the Relational Algebra	320
Additional Operators	321
Consequences of the Foregoing Treatment	322
Implementation Anomalies	327
Default Values	331
Conclusion	333
Acknowledgments	334
References	334

CHAPTER 16

The Outer Join	335
Abstract	335
Comments on Republication	335
Introduction	336
A Note on Null Values	337
Definitions and Examples	338

Discussion and Further Examples.....	344
A Concrete Proposal.....	352
Summary of Syntax.....	361
A Comparison with Chamberlin's Proposal.....	362
Conclusion.....	365
Acknowledgments.....	366
References.....	366

CHAPTER 17

Updating Views

Abstract.....	367
Comments on Publication.....	367
Introduction.....	368
Scope of This Paper.....	369
What is View Updating For?.....	370
Some Preliminary Remarks.....	372
Base Tables, Views, and Relations.....	374
Restriction.....	377
Projection.....	379
Join.....	382
(PK,PK)-Join.....	382
Inner (PK,FK)-Join.....	386
Outer (PK,FK)-Join.....	391
Summary.....	394
References.....	394

CHAPTER 18

A Note on the Parts Explosion Problem

397

Abstract.....	397
Comments on Publication.....	398
Introduction.....	398
Tree-Structured Relations.....	400
Extending the SQL Cursor Mechanism.....	401
A PL/I Example.....	401
A COBOL Example.....	402
Discussion.....	403
Analysis.....	407
Extending the SQL Select Operation.....	407
Summary.....	411
Acknowledgment.....	412
References.....	412

PART IV DATABASE DESIGN**CHAPTER 19**

A Practical Approach to Database Design	417
Abstract	417
Comments on Republication	418
Section One: Prologue	418
Structure of the Paper	421
A Word of Warning	421
Examples	422
Disclaimer	424
Synthesis vs. Analysis	424
Section Two: The Synthesis Approach	425
Kernel Entities	429
A Digression: Recording Design Decisions	431
Associations	432
Designations	437
Characteristics	440
Properties	442
A Recipe for Primary Keys	445
A Recipe for Foreign Keys	445
The Synthetic Approach: A Checklist	447
Section Three: The Analytic Approach	448
Section Four: Miscellaneous Topics	453
Entity Classification	453
Alternate Keys	454
Composite Keys	455
Entity Supertypes and Subtypes	457
Domains	460
Overnormalization	461
Denormalization	462
Null Values	463
Vectors	464
Codes and Flags	466
Field Overloading	467
Individual User Views	469
Section Five: Epilogue	469
Acknowledgments	470

AFTERWORD: THE RELATIONAL FUTURE

Comments on Republication	471
Relational Systems Today.....	472
Defining the Relational Model.....	474
Relational DBMSs	475
Relational Languages	478
On IBM'S DB2	481
Answering Criticisms.....	484
Normalization.....	487
Database Machines	490

Index

493

PART

I

RELATIONAL DATABASE MANAGEMENT

Relational Database: An Overview

(AN INTRODUCTION TO RELATIONAL SYSTEM
CONCEPTS, USING SQL/DS AS A
CONCRETE EXAMPLE)

ABSTRACT

The aim of this paper is to present a brief and very informal description of relational database systems, using SQL/DS as a vehicle for purposes of illustration. Codd's pioneering paper on the relational model [3] was published in June 1970. SQL/DS, IBM's first fully supported relational database program product, was announced in January 1981 [10]. A secondary objective of the present paper is to give some guidance to the vast body of literature that has grown up on the subject in the intervening years (a Fur-

Originally published in *Proc. GUIDE 53*, Atlanta, Georgia (May 1981); *Proc. SEAS Spring Meeting*, Noordwijkerhout, Netherlands (April 1983); *Proc. 10th Australian Computer Conference*, Melbourne, Australia (September 1983); and elsewhere. Also published (in considerably revised form) as *Relational Data Base System Concepts Using SQL/DS*, copyright 1984 Auerbach Publishers Inc. Data Base Management.