Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

An Implementation of Domains and Keys in SQL

A Thesis

presented in fulfilment of

the requirements for the degree

Master of Philosophy

at

Massey University

John Lindsay Alexander

1987

ABSTRACT

The relational Data Base Management System model has steadily acquired increasing acceptance over the years since it was first introduced in 1970, yet few - if any - of the many relational products currently available support the fundamental concepts of domains and the integrity constraints of primary and foreign keys. Over more recent years the SQL Relational Database Sub-Language has found most favour among users and vendors alike, and a standard for SQL has recently been produced. This standard provides no support for domains or foreign keys, and only indirect support for primary keys.

This thesis first reviews MURDER, the relational dabatase management system used for teaching purposes at Massey, and then describes an implementation of the domain and key concepts, highlighting some of the problem areas still to be resolved. Also described is an implementation of the query and update facilities of SQL, including some extensions which it is claimed increase its functionality. Finally, refinements to the language definition are suggested, to remove some redundancies and ambiguities.

Acknowledgements

I wish to express my gratitude to

my supervisors June Verner and Chris Phillips, for their encouragement, advice and support;

Elizabeth Kemp, for her patient forbearance when the system kept changing, and for many helpful suggestions;

the many students, known and unknown, who never failed to activate the most obscure and insidious bugs; and

my wife Shirley, without whose acceptance of the unsocial habits such an undertaking inevitably develops, nothing would have been possible.

CONTENTS

page

List of Illustrations iv

List of Algorithms v

Chapter 1 Introduction 1
1.1 Overview 1
1.1.1 Background 1
1.1.2 Database Teaching at Massey . . . 2
1.1.3 The MURDER System 3
1.2 Project Outline 4
1.3 Syntax 6

Chapter 2 Relational Database Management Systems . . . 7
2.1 Example Database 7
2.2 Informal Introduction 8
2.3 Formal Definition 10
2.4 Practical Problems 15
2.4.1 Domains 15
2.4.2 Keys 17
2.4.3 Assignment 21
2.4.4 Null Values 25
2.4.5 Indexes 27

Chapter 3 MURDER 31
3.1 Overview of Implementation 31
3.2 Relation Definition 38
3.3 Additional Features 45
3.3.1 Utility Statements 45
3.3.2 Experimental Statements 47
3.4 Evaluation 57

Chapter 4 MURDER Algebra 61
4.1 Projection 63
4.2 Selection 66
4.3 Union 68
4.4 Difference 69
4.5 Intersection 70
4.6 Product 71
4.7 Join 73
4.7.1 Theta-join 73
4.7.2 Natural Join 75
4.7.3 Other Joins 78
4.8 Quotient 79
4.9 Updates 82
4.10 Discussion 83
4.11 Efficiency 85

Chapter 5 MURDER Calculus 88
5.1 Relational Calculus 88
5.1.1 The Target List 89
5.1.2 Simple Qualification 92
5.1.3 Range Expressions 92
5.1.4 Further Qualification 93
5.2 Codd's Algorithm 97
5.3 MURDER Implementation 102
5.4 Optimization 104
5.4.1 Eliminating Redundant Operations . 104
5.4.2 Eliminating Redundant Computation 107
5.4.3 Reducing Relation Size 107
5.5 Discussion 110

Chapter 6 Introduction to SQL 114
6.1 DSL Alpha 114
6.1.1 Retrieval 114
6.1.2 Updates 117
6.1.3 Other Facilities 119
6.2 SEQUEL 121
6.3 SEQUEL2 and SQL 124
6.3.1 Simple Queries 126
6.3.2 Nested Queries 129
6.3.3 Grouping 133
6.3.4 Indexes 135
6.3.5 Views 137
6.3.6 Updates 139

Chapter 7 MURDER SQL 141
7.1 The Basic SQL Expression 141
7.2 The WHERE Clause 147
7.2.1 Simple Booleans 148
7.2.2 The EXISTS Sub-query 150
7.2.3 The IN Sub-query 153
7.3 The GROUP Clause 161
7.3.1 Simple Grouping 161
7.3.2 The HAVING Clause 165
7.4 Binary Queries 172
7.5 Modifications to the Calculus 174
7.6 Updates 177

Chapter 8 Analysis and Conclusions 183
8.1 Domains 183
8.2 Integrity 184
8.3 SQL 191
8.3.1 Suitability of Calculus 191
8.3.2 Omissions in MURDER 193
8.3.3 Suggested Extensions 196

Appendix A Algorithms 206
B SEQUEL Syntax 209
C SEQUEL2 Query Syntax 211
D MURDER Syntax 214
Bibliography 221

ILLUSTRATIONS

Figure 2.1 SPJ Database 7
2.2 SUPPLIER Relation 9
3.1 Page Structure 31
3.2 Attribute Structure 34
3.3 Domain Structure 35
3.4 Relation Structure 36
3.5 Data Structures for SUPPLIER Relation . 37
3.6 Reference Structure 39
4.1 Expression Node Structure 64
4.2 Data Structures for Projection 64
4.3 Data Structures for Selection 68
5.1 Calculus Node Structure 102
5.2 Data Structures for Calculus Query . . . 105
7.1 Data Structures for
GROUP BY Query Using COUNT Function . 164
7.2 Data Structures for
HAVING function θ function 166
8.1 Network Version of SPJ Database 188

ALGORITHMS

Algorithm 3.1 Find and Scan Procedures 33
4.1 Projection 66
4.2 Selection 68
4.3 Union 69
4.4 Difference 70
4.5 Intersection 71
4.6 Product 72
4.7 Join 77
4.8 Quotient 82
5.1 Codd's Algorithm 98
5.2 Calculus Universe 111
7.1 High-level GROUP Logic 164
7.2 Empty 172
7.3 Revised High-level MODIFY 179
7.4 High-level UPDATE 180
7.5 MODIFY Outline 181