Computer Science Series

Consulting Editor:

Professor F.H. Sumner, University of Manchester

A. Abdellatif, J. Le Bihan, M. Limame, *Oracle - A User's Guide*

Ian O. Angell, *High-resolution Computer Graphics Using C*

Ian O. Angell and Gareth Griffith, *High-resolution Computer Graphics Using FORTRAN 77*

Ian O. Angell and Gareth Griffith, *High-resolution Computer Graphics Using Pascal*

M. Azmoodeh, *Abstract Data Types and Algorithms, second edition*

C. Bamford and P. Curran, *Data Structures, Files and Databases, second edition*

P. Beynon-Davies, *Database Systems*

P. Beynon-Davies, *Information Systems Development, second edition*

G.M. Birtwistle, *Discrete Event Modelling on Simula*

Linda E.M. Brackenbury, *Design of VLSI Systems - A Practical Introduction*

Alan Bradley, *Peripherals for Computer Systems*

G.R. Brookes and A.J. Stewart, *Introduction to occam2 on the Transputer*

P.C. Capon and P.J. Jinks, *Compiler Engineering Using Pascal*

Robert Cole, *Computer Communications, second edition*

Eric Davalo and Patrick Naïm, *Neural Networks*

S.M. Deen, *Principles and Practice of Database Systems*

Joyce Duncan, Lesley Rackley and Alexandria Walker, *SSADM in Practice*

D. England *et al.*, *A Sun User's Guide, second edition*

Jean Ettinger, *Programming in C++*

J.S. Florentin, *Microprogrammed Systems Design*

A.B. Fontaine and F. Barrand, *80286 and 90386 Microprocessors*

Michel Gauthier, *Ada - A Professional Course*

M.G. Hartley, M. Healey and P.G. Depledge, *Mini and Microcomputer Systems*

J.A. Hewitt and R.J. Frank, *Software Engineering in Modula-2 - An Object-oriented Approach*

Patrick Jaulent, *The 68000 - Hardware and Software*

M.J. King and J.P. Pardoe, *Program Design Using JSP - A Practical Introduction, second edition*

Bernard Leguy, *Ada - A Programmer's Introduction*

M. Léonard, *Database Design Theory*

David Lightfoot, *Formal Specification Using Z*

A.M. Lister and R.D. Eager, *Fundamentals of Operating Systems, fifth edition*

Elizabeth Lynch, *Understanding SQL*

Tom Manns and Michael Coleman, *Software Quality Assurance, second edition*

G.P. McKeown and V.J. Rayward-Smith, *Mathematical Foundations for Computing*

B.A.E. Meekings, T.P. Kudrycki and M.D. Soren, *A book on C, third edition*

R.J. Mitchell, *C++ Object-oriented Programming*

R.J. Mitchell, *Microcomputer Systems Using the STE Bus*

R.J. Mitchell, *Modula-2 Applied*

Y. Nishinuma and R. Espesser, *UNIX - First Contact*

Ian Pratt, *Artificial Intelligence*

Pham Thu Quang and C. Chartier-Kastler, *MERISE in Practice*

E.J. Redfern, *Introduction to Pascal for Computational Mathematics*
F.D. Rolland, *Programming with VDM*
A.G. Sutcliffe, *Human-Computer Interface Design, second edition*
C.J. Theaker and G.R. Brookes, *Concepts of Operating Systems*
M. Thorin, *Real-time Transaction Processing*
M.R. Tolhurst *et al.*, *Open Systems Interconnection*
A.J. Tyrrell, *Eiffel Object-Oriented Programming*
I.R. Wilson and A.M. Addyman, *A Practical Introduction to Pascal, second edition*

Other titles

Ian O. Angell and Dimitrios Tsoubelis, *Advanced Graphics on VGA and XGA Cards
Using Borland C++*
N. Frude, *A Guide to SPSS/PC+, second edition*
Peter Grossman, *Discrete Mathematics for Computing*
Percy Mett, *Introduction to Computing*
Tony Royce, *COBOL - An Introduction*
Tony Royce, *Structured COBOL - An Introduction*

Database Systems

Paul Beynon-Davies

*Department of Computer Studies
University of Glamorgan*

© P. Beynon-Davies, 1996

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No paragraph of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, 90 Tottenham Court Road, London W1P 9HE.

Any person who does any unauthorised act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

First published 1996 by
MACMILLAN PRESS LTD
Houndmills, Basingstoke, Hampshire RG21 6XS
and London
Companies and representatives
throughout the world

ISBN 978-0-333-63667-1 ISBN 978-1-349-13722-0 (eBook)
DOI 10.1007/978-1-349-13722-0

A catalogue record for this book is available
from the British Library.

10	9	8	7	6	5	4	3	2	1
05	04	03	02	01	00	99	98	97	96

Contents

<i>Preface</i>	xi
1 Databases as Abstract Machines	1
1.1 Introduction	1
1.2 Universe of Discourse	1
1.3 Integrity	3
1.4 Database Functions	5
1.5 Formalisms	7
1.6 Multi-User Access	8
1.7 Conclusion	8
1.8 Reference	9
1.9 Exercises	9
2 Key Concepts	10
2.1 Introduction	10
2.2 Piecemeal Development	10
2.3 Databases and Database Management Systems	11
2.4 What is a Database Management System?	13
2.5 Data Models	14
2.6 Database System	16
2.7 A Brief History of Database Software	16
2.8 Database Development.	18
2.9 Conclusion	19
2.10 References	19
Part 1 Data Models	21
3 Relational Data Model	23
3.1 Introduction	23
3.2 Data Definition	23
3.3 Data Manipulation	27
3.4 Data Integrity	37
3.5 Conclusion	40
3.6 References	40
3.7 Exercises	40
4 Classic Data Models	42
4.1 Introduction	42

4.2	The Hierarchical Data Model	42
4.3	The Network Data Model	44
4.4	Comparison of Classic Data Models with Relational Data Model	47
4.5	Conclusion	48
4.6	Reference	48
4.7	Exercises	48
5	Object-Oriented Data Model	49
5.1	Introduction	49
5.2	Extended Relational Approaches	50
5.3	Components of an OO Data Model	52
5.4	Conclusion	58
5.5	References	58
5.6	Exercises	58
Part 2	Database Management	59
6	DBMS Interface - SQL	61
6.1	Introduction	61
6.2	The History of SQL	61
6.3	University Database	62
6.4	Data Definition	63
6.5	Data Manipulation	66
6.6	Data Integrity	73
6.7	Additions in SQL2	75
6.8	Conclusion	78
6.9	References	79
6.10	Exercises	79
7	DBMS Kernel	80
7.1	Introduction	80
7.2	Physical Storage	80
7.3	Kernel Components	81
7.4	Key Functions of the Kernel	83
7.5	File Organisation	83
7.6	Access Mechanisms	86
7.7	Transaction Management	88
7.8	Dictionary Management	92
7.9	Query Management	93
7.10	Backup and Recovery	95
7.11	Conclusion	96

8	DBMS Toolkit	97
8.1	Introduction	97
8.2	SQL Interfaces	97
8.3	4GLs and SQL	102
8.4	End-User Interfaces	103
8.5	Conclusion	106
8.6	References	106
9	Distributed Database Systems	107
9.1	Introduction	107
9.2	Objectives	107
9.3	Relational Database Systems	108
9.4	Advantages of Distribution	108
9.5	Types of Distributed Database System	109
9.6	Distributed DBMS	110
9.7	Client-Server Database Systems	111
9.8	Conclusion	114
9.9	References	114
10	Two Commercial DBMS	116
10.1	Introduction	116
10.2	A Relational DBMS - ORACLE	116
10.3	An Object-Oriented DBMS - O ₂	121
10.4	Conclusion	125
10.5	References	126
Part 3	Database Development	127
11	Data and Database Administration	129
11.1	Introduction	129
11.2	Data Administration	129
11.3	The Development of Database Systems	131
11.4	Database Administration	134
11.5	Data Control	135
11.6	Data Dictionaries	141
11.7	Conclusion	142
11.8	References	143
12	Normalisation	144
12.1	Introduction	144
12.2	Why Normalise?	144

12.3	Stages of Normalisation	145
12.4	The Normalisation Oath	150
12.5	Determinacy Diagramming	151
12.6	Pragmatics of Determinacy Diagramming	153
12.7	Accommodation	154
12.8	Determinacy Diagramming and the Normal Forms	157
12.9	Conclusion	160
12.10	References	160
12.11	Exercises	160
13	Entity-Relationship Diagramming	162
13.1	Introduction	162
13.2	Data Models	162
13.3	Rules Defining a Relationship	166
13.4	Extensions to Data Modelling	170
13.5	Pragmatics of Data Modelling	172
13.6	Accommodation to Relational Schema	176
13.7	Data Dictionaries	178
13.8	Schema Integration	179
13.9	Conclusion	182
13.10	References	183
13.11	Exercises	183
14	Object Modelling	185
14.1	Introduction	185
14.2	A Universe of Discourse	185
14.3	Object Models	186
14.4	Structural Abstraction	187
14.5	Generalisation	187
14.6	Aggregation	192
14.7	Behavioural Abstraction	196
14.8	Object Life Histories	198
14.9	Composing an Object Model	201
14.10	Abstraction	203
14.11	Conclusion	205
14.12	References	205
14.13	Exercises	206
15	Physical Database Design	207
15.1	Introduction	207
15.2	Inputs to and Outputs from Physical Design	207
15.3	University Short Courses	208

15.4	Logical Design	209
15.5	Volume Analysis	212
15.6	Usage Analysis	212
15.7	Access Requirements	214
15.8	Integrity Analysis	214
15.9	Defining Performance	215
15.10	Performance Decisions	216
15.11	Distributed Database Design	220
15.12	Conclusion	220
15.13	References	221
15.14	Exercises	221
Part 4	New Developments	223
16	Parallel Computers for Database Systems By Frank Sumner	225
16.1	Historical Background	225
16.2	The Special Problems of Commercial Data Processing	226
16.3	Today's Computers	227
16.4	Multi Processor Systems (Parallel Computing)	227
16.5	Symmetric Multi Processors	231
16.6	Shared Nothing Systems	232
16.7	Interconnected Arrays	234
16.8	Globally Distributed Shared Virtual Memory	235
16.9	Conclusion	237
16.10	References	237
17	Intelligent Databases	238
17.1	Introduction	238
17.2	Intelligence	238
17.3	Database Systems and PSSs	239
17.4	Logic and Databases	240
17.5	Data Definition in Datalog	240
17.6	Data Manipulation in Datalog	243
17.7	Data Integrity in Datalog	244
17.8	Datalog and the Relational Data Model	246
17.9	Differences between Datalog and Prolog	248
17.10	Extensions to Pure Datalog	248
17.11	Intelligent Databases and Business Applications	250
17.12	Conclusion	251
17.13	References	252
17.14	Exercises	252

18	Database Systems and Complex Applications	253
18.1	Introduction	253
18.2	Hypermedia Systems	253
18.3	Geographical Information Systems	259
18.4	Conclusion	263
18.5	References	264
18.6	Exercises	265
	<i>Sample Solutions</i>	267
	<i>Index</i>	277

Preface

The main aim of this work is to provide one readable text of essential core material for most higher education and commercial courses on database systems. The current volume is designed to form a comprehensive text on modern database technology and the development of database systems.

It is undoubtedly true that database systems hold a prominent place in most contemporary approaches to the development of information systems. It is this practical emphasis on the use of database systems for information systems work that distinguishes the current volume from other, more theoretical, texts on the subject. This work therefore forms a companion volume to *Information Systems Development: an introduction to information systems engineering*, also published by Macmillan Press.

The text is built from the author's experiences of running a number of academic and commercial courses on database technology for some seven years or so.

The text is organised into a number of parts:

1. The introductory chapters set the scene for the core of the text. We begin with a description of the key features of a database system. We then move on to define some key concepts.
2. Part one explores a number of contemporary architectures for database systems. Because of its current dominance, particular emphasis is made of the relational data model.
3. Part two provides a description of the major elements of a contemporary relational database management system: interface, kernel and toolkit. The issue of distributed database systems is also addressed. The part concludes by reviewing two existing DBMS: one relational; the other object-oriented.
4. Part three presents a discussion of the major techniques utilised in the design and implementation of database systems. The issue of administering data in organisations is also considered.
5. Part four considers new applications for databases. It consists of three chapters which look at areas having a significant effect on the functionality of database systems: the issue of parallelism (by Professor Frank Sumner), the issue of embedded 'intelligence' and handling complex data.

Finally, the author is keen to receive feedback on the current text. Any comments or suggestions should be addressed to the author, care of the publisher. My thanks go to Professor Frank Sumner of the University of Manchester for kindly producing chapter 16 and to Malcolm Stewart at Macmillan for his support in completing this project.