Macmillan Computer Science Series

Consulting Editor

Professor F. H. Sumner, University of Manchester

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

S. T. Allworth and R. N. Zobel, *Introduction to Real-time Software Design, second edition*

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*

Philip Barker, *Author Languages for CAL*

A. N. Barrett and A. L. Mackay, *Spatial Structure and the Microcomputer*

R. E. Berry, B. A. E. Meekings and M. D. Soren, *A Book on C, second edition*

P. Beynon-Davies, *Information Systems Development*

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

B. G. Blundell and C. N. Daskalakis, *Using and Administering an Apollo Network*

B. G. Blundell, C. N. Daskalakis, N. A. E. Heyes and T. P. Hopkins, *An Introductory Guide to Silvar Lisco and HILO Simulators*

T. B. Boffey, *Graph Theory in Operations Research*

Richard Bornat, *Understanding and Writing Compilers*

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 occam 2 on the Transputer*

J. K. Buckle, *Software Configuration Management*

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

J. C. Cluley, *Interfacing to Microprocessors*

J. C. Cluley, *Introduction to Low Level Programming for Microprocessors*

Robert Cole, *Computer Communications, second edition*

Derek Coleman, *A Structured Programming Approach to Data*

Eric Davalo and Patrick Naim, *Neural Networks*

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

C. Delannoy, *Turbo Pascal Programming*

Tim Denvir, *Introduction to Discrete Mathematics for Software Engineering*

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

J. S. Florentin, *Microprogrammed Systems Designs*

A. B. Fontaine and F. Barrand, *80286 and 80386 Microprocessors*

J. B. Gosling, *Design of Arithmetic Units for Digital Computers*

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*

(continued overleaf)

Roger Hutton, *COBOL 85 Programming*
 Roger Hutton, *Z80 Assembly Language Programming for Students*
 Roland N. Ibbett and Nigel P. Topham, *Architecture of High Performance Computers, Volume I*
 Roland N. Ibbett and Nigel P. Topham, *Architecture of High Performance Computers, Volume II*
 Patrick Jaulent, *The 68000 - Hardware and Software*
 P. Jaulent, L. Baticle and P. Pillot, *68020-30 Microprocessors and their Coprocessors*
 M. J. King and J. P. Pardoe, *Program Design Using JSP - A Practical Introduction*
 E. V. Krishnamurthy, *Introductory Theory of Computer Science*
 V. P. Lane, *Security of Computer Based Information Systems*
 David Lightfoot, *Formal Specification Using Z*
 A. M. Lister and R. D. Eager, *Fundamentals of Operating Systems, fourth edition*
 Elizabeth Lynch, *Understanding SQL*
 Tom Manns and Michael Coleman, *Software Quality Assurance*
 A. Mével and T. Guéguen, *Smalltalk-80*
 R. J. Mitchell, *Microcomputer Systems Using the STE Bus*
 R. J. Mitchell, *Modula-2 Applied*
 Y. Nishinuma and R. Espesser, *UNIX - First contact*
 Pim Oets, *MS-DOS and PC-DOS - A Practical Guide, second edition*
 Pham Thu Quang and C. Chartier-Kastler, *MERISE in Practice*
 A. J. Pilavakis, *UNIX Workshop*
 Christian Queinnec, *LISP*
 E. J. Redfern, *Introduction to Pascal for Computational Mathematics*
 Gordon Reece, *Microcomputer Modelling by Finite Differences*
 W. P. Salman, O. Tisserand and B. Toulout, *FORTH*
 L. E. Scales, *Introduction to Non-Linear Optimization*
 Peter S. Sell, *Expert Systems - A Practical Introduction*
 A. G. Sutcliffe, *Human-Computer Interface Design*
 Colin J. Theaker and Graham R. Brookes, *A Practical Course on Operating Systems*
 M. R. Tolhurst *et al.*, *Open Systems Interconnection*
 J-M. Trio, *8086-8088 Architecture and Programming*
 A. J. Tyrrell, *COBOL from Pascal*
 M. J. Usher, *Information Theory for Information Technologists*
 B. S. Walker, *Understanding Microprocessors*
 Colin Walls, *Programming Dedicated Microprocessors*
 I. R. Wilson and A. M. Addyman, *A Practical Introduction to Pascal - with BS6192, second edition*

Data Structures, Files and Databases

Carl Bamford

*Division of Computing and Mathematical Sciences
Liverpool Polytechnic*

and Paul Curran

Quantac Image Processing Ltd

Second Edition

M
MACMILLAN
EDUCATION

© Carl Bamford and Paul Curran 1987, 1991

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, 33-4 Alfred Place, London WC1E 7DP.

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

First edition 1987
Second edition 1991

Published by
MACMILLAN EDUCATION LTD
Houndmills, Basingstoke, Hampshire RG21 2XS
and London
Companies and representatives
throughout the world

British Library Cataloguing in Publication Data
Bamford, Carl

Data structures, files and databases. – 2nd ed.

– (Computer science series)

I. Title II. Curran, Paul III. Series

005.74

ISBN 978-0-333-56043-3

ISBN 978-1-349-12615-6 (eBook)

DOI 10.1007/978-1-349-12615-6

Contents

<i>Preface</i>	<i>ix</i>
1 Data Structures I	1
1.1 Introduction	1
1.2 Data types and conversions	2
1.3 Lists	4
1.4 Hierarchical structures	6
1.5 Tables	7
1.6 Variant records	15
1.7 Summary	16
Exercises	16
2 Serial and Sequential Data Files	18
2.1 Introduction	18
2.2 Files of simple data types	19
2.3 Text files	19
2.4 File buffers in Pascal	22
2.5 Files of records	24
2.6 Sequential files and applications	25
2.7 Summary	36
Exercises	37
3 Data Structures II	38
3.1 Pointers	38
3.2 Access vectors	38
3.3 Stacks	42
3.4 Queues	50
3.5 Summary	54
Exercises	54
4 Dynamic Data Structures	56
4.1 Linked lists	56
4.2 Free storage lists	67
4.3 Stacks and queues	69
4.4 Dummy entries and circular chains	70

4.5	Sets	72
4.6	Priority queues	73
4.7	Circular chains and doubly linked lists	73
4.8	Summary	75
	Exercises	76
5	Non-serial File Access	77
5.1	Introduction	77
5.2	Direct files	77
5.3	Hashed random files	79
5.4	Indexes	80
5.5	Inverted files	89
5.6	Summary	89
	Exercises	90
6	Recursion and Trees	91
6.1	Recursion	91
6.2	Indirect recursion in Pascal	96
6.3	Trees	97
6.4	General trees	115
6.5	Multi-way trees	118
6.6	B-trees	119
6.7	Summary	125
	Exercises	125
7	Graphs, Networks and Relations	127
7.1	Graphs and networks	127
7.2	Applications of graphs	133
7.3	Implementations of graphs	135
7.4	Relations	145
7.5	Summary	148
	Exercises	148
8	Design Considerations	150
8.1	Problem analysis	150
8.2	Abstract data types and formal specifications	150
8.3	Data structure design	152
8.4	Internal data structures or external file?	153
8.5	Summary	154
9	Database Management Systems	155
9.1	Introduction	155
9.2	Limitations of conventional filing systems	155
9.3	What is a database?	156
9.4	DBMS software	156

9.5	Conceptual, logical and physical models	157
9.6	Database terminology	158
9.7	Summary	160
10	Database Design	162
10.1	Introduction	162
10.2	Conceptual models	162
10.3	To develop a data model diagram	162
10.4	Functional analysis approach to conceptual modelling	166
10.5	Normalisation	168
10.6	Summary	177
	Exercises	177
11	Network Database Models	179
11.1	Introduction	179
11.2	The CODASYL database model	179
11.3	The IMAGE database model	189
11.4	Summary	193
	Exercises	194
12	Relational Database Models	196
12.1	Introduction	196
12.2	Relational database terminology	196
12.3	Relational database operations	198
12.4	Translation of conceptual model to a logical relational model	202
12.5	Using relational operators to access a database	202
12.6	SQL/DS relational database	203
12.7	dBase II database	204
12.8	Implementation considerations	205
12.9	Summary	206
	Exercises	206
13	SQL	209
13.1	Creating the database, tables and indexes	210
13.2	Altering the database structure	212
13.3	Adding, deleting and updating data	212
13.4	Retrieving data from one table	215
13.5	Retrieval from more than one table	219
13.6	Grouped data	220
13.7	Controlling user access	221
13.8	Interactive and host-language access	223
13.9	Criticism of SQL	225
13.10	Summary	226
	Exercises	227

14 Fourth Generation Languages	229
14.1 What is a 4GL?	229
14.2 4GL environments and 4GL programs	231
14.3 Reports	231
14.4 Screen forms	239
14.5 Menus	241
14.6 Additional tools	243
14.7 Prototyping	245
14.8 Summary	246
Exercises	247
 <i>Appendix A: Review of Data Types in Pascal</i>	 249
 <i>Appendix B: Text Files in Pascal</i>	 256
 <i>Bibliography</i>	 259
 <i>Index</i>	 261

Preface

The first edition of this text has been used successfully on introductory courses in data structures, file structures and databases for second-year undergraduates and HND students in computer studies. It covers a wide range of material appropriate to this level, and assumes the reader has completed a first course in programming, Pascal being used to illustrate the topics covered. Although it is intended for undergraduate/HND students, its practical approach makes it suitable for professional programming practitioners.

For the second edition, we have retained the same theme: that the design of structures for data and processes should be considered first at an abstract level and then developed using concrete representations which take account of efficiency. The changes we have made relate to databases, where we have added more details of the language SQL and discussed fourth generation languages, topics which are increasingly important.

Data structures are treated informally, with an abstract view of each structure examined separately from its concrete implementation, preparing the reader for a more formal study of abstraction at a later date. We have emphasised the use of procedures and functions to implement the operations on the structures, thus hiding the implementation details to some extent from the main program, an important technique which is used extensively in object-oriented programming.

Standard Pascal is used for the examples on data structures and serial files. In the discussions of random access files, an extended version is introduced which is typical of many implementations with built-in routines to handle such files.

The later chapters of the book provide an introduction to databases. A practical design method which produces a conceptual data model is discussed. Such a model provides an abstract view of the data before a logical model is derived for a particular implementation. Both network and relational database models are discussed as possible implementations. With the increasing importance of the relational model and the adoption of SQL as a standard language, we have provided a fuller account of SQL in a separate chapter. Fourth generation languages are discussed in the final chapter, contrasting their non-procedural features with the procedural style of third generation languages such as Pascal.

We should like to thank the people who have helped in the preparation of this text; in particular, the reviewers whose constructive comments led to improvements in the presentation of the material, and the readers who made valuable comments on the first edition.