information technology

research development applications

volume 3
number 2
april 1984

Butterworths

EDITOR

Professor C J van Rijsbergen
Department of Computer Science, University College, Dublin 4, Ireland

US EDITOR

Professor W B Croft
Department of Computer Science, University of Massachusetts, Amherst, MA 01003, USA

BOOK REVIEWS EDITOR

Dr C D Paice
Department of Computer Studies, Lancaster University, Lancaster UK

EDITORIAL ADVISORY BOARD

Dr T Addis
Department of Computer Science, Brunel University, Uxbridge, Middx. UB8 3PH, UK

D H Barlow
Kent-Barlow Information Associates, Century House, 79 Great Titchfield Street, London W1P 7FN, UK

Dr H P Frei
Institut fur Informatik, ETH-Zentrum, CH-8092 Zurich, Switzerland

Dr D J Harper
Department of Computer Science, University College, Dublin 4, Ireland

Dr P Higginson
Department of Computer Science, University College, London WC1E 6BT, UK

Dr J Lloyd
Department of Computer Science, The University of Melbourne, Parkville, Victoria 3052, Australia

R D Mannix
Unilever Research Laboratory, Bebington, Wirral, Merseyside L63 3JW, UK

Professor M J McGill
Ohio College Library Center, 6565 Frantz Road, Dublin, Ohio, USA

Professor D R McGregor
Department of Computer Science, University of Strathclyde, Livingstone Tower, Richmond Street, Glasgow G1 1XH, UK

Mr N Naffah
INRIA, Bulltransac, 127, rue Ampère, 91300 Massy, France

Professor W M Newman
14 Denning Road, London NW3, UK

Professor R N Oddy
School of Information Studies, Syracuse University, 113 Euclid Avenue, Syracuse, NY 13210, USA

Dr T Radecki
Department of Computer Science, Louisiana State University, Baton Rouge, LA 70803–4020, USA

Dr S E Robertson
Centre for Information Science, The City University, London EC1V 0HB, UK

Professor G Salton
Department of Computer Science, Cornell University, Ithaca, NY 14850, USA

Dr H J Schek
Department of Informatics and Office Systems, Technische Hochschule Darmstadt, Alexanderstrasse 24, 61 Darmstadt, FRG

D Wallis
Reckitt H & Colman, Dansom Lane, Hull HU8 7DS, UK

Dr E J Yannakoudakis
Department of Computer Science, University of Bradford, West Yorkshire BD7 1DP, UK

Professor C T Yu
Department of Information Engineering, University of Illinois, Chicago, IL 60680, USA

Publishers, Editorial, Advertisement, Production and Reprint Offices:

Butterworth Scientific Limited,
PO Box 63, Westbury House, Bury Street,
Guildford, Surrey GU2 5BH, UK
Telephone: Guildford (0483) 31261
Telex: 859556 SCITEC G

Publishing Director: C J Rawlins
Production Manager:
B Marshall

Subscription enquiries and orders should be sent to:

Quadrant Subscription Services Limited,
Oakfields House, Perrymount Road,
Haywards Heath, Sussex RH16 3DH, UK
Telephone: Haywards Heath (0444) 459188

Annual Subscription price:

UK: £55.00 Overseas: £55.00 Single copies: £16.50. The journal is published every three months. Prices include packing and delivery by surface mail. Airmail prices available on request.

Copies of the journal sent to Canada and the United States of America are airspeeded at no extra cost. Second class postage paid of New York, NY.
Postmaster send all address corrections to c/o Expediters of the printed Word Limited, Suite 1217, 527 Madison Avenue, New York, NY 10022, USA.

This journal is listed in Current Contents.

US Mailing Agents:

Expediters of the printed Word Limited,
Suite 1217, 527 Madison Avenue, New York,
NY 10022, USA

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 written permission of the Publisher.

Readers who require copies of papers published in this journal may either purchase reprints or obtain permission to copy from the publisher at the following address: Butterworth Scientific Limited, PO Box 63, Westbury House, Bury Street, Guildford, Surrey GU2 5BH, UK. For readers in the USA, permission to copy is given on the condition that the copier pay the stated per-copy fee through the Copyright Clearance Center Inc., PO Box 765, Schenectady, NY 12301, USA, for copying beyond that permitted by Sections 107 and 108n of the US Copyright Law. Fees appear in the code printed at the foot of the first page of each paper.

Back issues prior to the current volume are available from: Wm Dawson & Sons Limited, Canon House, Folkestone, Kent CT19 5EE, UK.

ISSN 0144–817X

© 1984 Butterworth & Co. (Publishers) Ltd.

Typeset by Tech-Set, Gateshead, Tyne & Wear.
Printed by J.W. Arrowsmith Ltd., Winterstoke Road, Bristol, UK.

information technology

research development applications

vol 3 no 2
april 1984

Editor-in-Chief Professor C J van Rijsbergen

67 **Graphical interfaces for binary relationship databases.** Database interfaces suitable for the nonprogrammer are discussed by *F N Teskey, N Dixon* and *S C Holden*

78 **A nearest neighbour search algorithm for bibliographic retrieval from multilist files.** *Peter Willett* presents a best match search procedure and tests its efficiency

84 **An algorithm for weighted searching on a Boolean system.** *J D Bovey* and *S E Robertson* discuss the problem of turning a weighted search into a series of Boolean searches

88 **Maximum entropy and the optimal design of automated information retrieval systems.** *Paul Kantor* applies the maximum entropy principle to information storage and retrieval

Survey paper

95 **Information retrieval from classical databases from a signal-detection standpoint — a review.** *M Heine* applies a generalized form of the Swetsian schema

Technical note

113 A comparison of the cosine correlation and the modified probabilistic model. *Bruce Croft* draws some interesting conclusions about the cosine correlation

115 **Book reviews.** Systems design — libraries — security — videotex

120 **Calendar**

Executive Editor Corinne Stocker **Managing Editor** Amanda Harper

Notes for Authors

INFORMATION TECHNOLOGY is concerned with research into, development and application of technology for storing and accessing information. It will cover both theoretical and practical aspects of information storage and retrieval, and in doing so, will bridge the gap between researcher and practitioner.

All technical articles will be refereed and may be accepted without alterations, require amendment before publication and subsequent refereeing, or be rejected on grounds of either relevancy or content.

The submission of a technical article will imply that, if accepted for publication, the paper will not be published elsewhere in the same form, in any language, without the consent of the Managing Editor. It is the author's responsibility to obtain written permission to quote material that has appeared in another journal.

PREPARATION OF SCRIPTS

All articles should be written in clear and concise English. Spelling should follow the Concise Oxford Dictionary. The paper should be typed in double spacing, on one side only of A4 paper. Do not underline anything. Leave wide margins. Number every sheet. Four copies are required and these should be sent to the Editor.

Arrangement of papers

Papers should be arranged as follows:

1. Title
2. Strapline of 15–25 words incorporating the names of the authors
3. 50–100 word abstract outlining the purpose, scope and conclusions of the paper
4. The text, suitably divided under headings
5. Acknowledgements (if any)
6. References
7. Tables (each on a separate sheet)
8. Captions to illustrations (grouped on a separate sheet or sheets)
9. Figures, each on a separate sheet, containing no text

Style of text

Trade names should have an initial capital letter.

Subdivide your paper in the simplest way possible, consistent with clarity.

Avoid numbering sections.

SI units should be used. Where this is not possible, give a conversion factor in parentheses at the first mention of the non SI unit.

Mathematical and technical setting

Write mathematical, Greek and other symbols carefully by hand unless you have access to a special typewriter. Explain the meaning of all symbols in the text where they first occur, and identify Greek letters in the margin at their first mention with their names in words. If you use several symbols, a list of definitions – not necessarily for publication – will help the editors. Identify all vectors, tensors and matrices by ringing with a pencil. Ensure that symbols and abbreviations for units follow the latest editions of British Standard 1991, Parts 1–6.

In particular, you should note the following points:

- Always put a zero before decimal quantities less than unity, e.g. 0.376 and not .376.
- Do not use the full stop as an ordinary multiplication sign.
- Use fractional indices instead of root signs, e.g. $(3x+9y)^{1/2}$ is preferable to $\surd(3x+9y)$.
- Lg x is assumed to mean $\log_{10} x$; $\log_e e$ is written as ln x.
- Write simple mathematical fractions in the text on one line, e.g.

  $L(N-1)/L(1)$ and not $\dfrac{L(N-1)}{L(1)}$
- Group numbers consisting of more than three numerals in threes separated by spaces; do not use commas.
- Indicate superscripts and subscripts clearly.

Tables

Number tables consecutively through the paper (with arabic numerals) referring to them in the text as Table 1, Table 2, etc., with a caption at the top of each table. Avoid the use of vertical rules. Tables should not duplicate results presented in graphs.

ILLUSTRATIONS

All illustrations should be identified with the author's name and the figure number marked in pencil.

Line illustrations

These should be drawn in black inks on tracing cloth or paper. Two photocopies should also be provided. For lettering, Standardgraph or similar stencils should be used (template size nos. 203/3.5 and 203/2.5) with the appropriate pen sizes. Lower-case (small) letters should be used throughout, with an initial capital letter for the first word only of a label. If suitable stencils are not available, a photocopy of the tracing should be lettered. Our studio will letter the original in ink in the standard style.

Illustrations should be drawn to either 172 mm or 344 mm width, including lettering. The maximum height at drawn size should not exceed 500 mm.

As all illustrations will be photographically reduced in size, lines must be drawn proportionately thicker and symbols larger than required in the printed version. Graph grid box lines should be drawn with a no. 0.3 Rotring or similar pen and the curves with pen no. 0.5. Two photocopies of all drawings should be supplied in addition to the originals. Use a selection of the following symbols on graphs:

X + ○ △ ▽ ■ ● ▲ ▼

The minimum amount of descriptive text should be used on graphs and drawings; label curves, points etc. with single-letter symbols, and place descriptive matter in the figure caption. Scale grids should not be used in graphs, unless required for actual measurements. A table is usually more satisfactory for recording data.

Illustrations showing lineprinter listings, plotter output, report forms and similar documents

If printouts are to be used as figures, provide clear, unmarked originals, preferably on unlined paper. Where several lines of figures are printed, include only the number of lines sufficient for the reader to interpret the nature of these figures. You should give a full explanation in the caption and text.

Photographs

Black and white glossy photographs (if not available, colour glossy photographs may be acceptable) should, where possible, be supplied unmounted, and should be labelled on the back in soft pencil with the author's name and the figure number. Two photocopies of photographs should also be supplied. A scale, where appropriate, should always be marked on photographs or included in the figure caption.

Figure captions

Captions should be typed on a separate sheet of paper.

REFERENCES

Indicate references to the literature in the text by superior arabic numerals that run consecutively through the paper. Where a reference is cited more than once in the text, use the same number each time. References should take the following form:

To journals:

1 **Barnes, J and Bergquist, B** 'Unite microprocessor hardware and software' *Electron Des.* (March 1976) pp 74–76

To books:

1 **McGlyn, DR** *Microprocessors: technology, architecture and applications* John Wiley, UK (1976)

Please ensure that references are complete, i.e. that they include, where relevant, author's name, article or book title, volume and issue number, publisher, date and page reference.

PROOFS

Correspondence and proofs for correction will be sent to the first named author, unless otherwise indicated.

The author will receive a copy of the page proofs for checking. It is essential that proofs be returned within three days. Authors may be charged for alterations in excess of 10% of the original setting.

REPRINTS

The principal author receives 25 reprints of the article, free of charge. Extra copies, minimum 100, can be ordered at any time from the Reprints Department, Butterworth Scientific Ltd, Guildford.

COPYRIGHT

Before publication, authors are requested to assign Copyright to Butterworth & Co (Publishers) Ltd. This allows Butterworth to sanction reprints and photocopies and to authorize the reprinting of complete issues or Volumes according to demand. Author's traditional rights will not be jeopardized by assigning Copyright in this manner, as they will retain the right to reuse and a veto over third-party publications.