Information Technology: Research and Development (1982), *1* (173–175)

BOOK REVIEWS

B. Christie, *Face to File Communication.* Chichester: J. Wiley and Sons Ltd. 1981. 306 pp. ISBN: 0 471 27939 0.

This is a book by a psychologist about the way people, particularly managers, use information, particularly of the electronically coded sort. It is one of the first volumes in a Wiley series on Information Processing. The material is largely based on Dr. Christie's experimental work on psychological aspects of communication carried out over a number of years in consultancy situations with a variety of clients.

It is important to classify this book properly. There is of course a literature on man-machine interaction; on the ergonomics of topics such as keyboard design and effective information displays. This book is not of that kind and should not be dismissed, as it might be by some, as another on 'human factors'. It should be seen rather as an early contribution to what will become an increasingly important field. The new cheap 'personal' computing, convergent with communication systems and office technology, will be inextricably involved with the human activities of management and other professions. Human activities are ill structured and integral to the organization of the institutions in which they take place. It follows that the technique of this computing cannot be separated from a study of the use and meaning of information by and for people. Only a small number of people are now fluent or even literate in both the technical and social aspects of this new subject matter.

The material in the book is divided into four main sections. The first develops some *basic concepts* of information. A useful distinction is made between what are called Type *A* and Type *B* communication. Type *A* is direct person to person. Type *B*, the one the book is mainly concerned with, involves some kind of intermediate store, either paper or electronic. There is a discussion of the value of information; also a little on the new kinds of information system — videotex, computer conferencing, word processing — which pose the kinds of problem under review. The second section looks at *behaviour* in relation to information. It is this section which has the most substantial experimental basis, resting on work done in the offices of the EEC in Brussels, a place one imagines where more information is shuffled than anywhere on Earth. Responses to widely circulated questionnaires are subjected to the statistical apparatus of factor analysis to expose the 'deep structure' of office activity. This leads the author to propose a rather formidable 'first law of information behaviour': 'if any two items are selected from the universe of information behaviour items, and if the population observed is not selected artificially, then the population regression between these two items will be monotone and with positive or zero sign'. The main result seems to be that there are some people in offices who use information a lot and these tend to be the same people who generate a lot. There is also material on how people seek information taking a route through a 'behaviour tree'. The next section deals with managers' *attitudes* to office automation, particularly in its effect on 'social presence', by which is meant a range of informal, non-verbal factors in communication. Lastly there is a section on *managing change* and resistance to it. The author believes that one effect of office

automation will be that managerial work will be more stressful since automation of routine will lead to more time spent on demanding face-to-face relationships. He concludes, a little surprisingly, with an advocacy of hypnotism (a particular interest of his) as a way of reducing stress.

In a field exposed to much variable quality journalism an attempt to provide a base of carefully analysed data is to be welcomed. There is, however, curiously in such an exciting subject, little vivid sense of real people wrestling with problems using real machines. The attempts at law formulation seem at best premature. Careful sympathetic observation of particular cases might seem preferable to analysis of variance of data from large statistical surveys.

It seems to this reviewer that a theory of information use will have to be embedded in a theory of organization, and if this is of a cybernetic kind it will have to confront problems of the characteristic multiple references (and hence meanings) of information used by people; also its relation to organizational power. There is a great deal to be done. We should be grateful to Dr. Christie for leading the way.

R. H. Anderton
Department of Systems
University of Lancaster
UK

N. J. Belkin and R. N. Oddy, *Design Study for an Anomalous State of Knowledge based Information Retrieval System.* British Library R. & D. Report no. 5547. September 1979. 226 pp.

Seeing that the *raison d'être* of any information retrieval system is to satisfy the needs of its users, why are actual systems nearly always designed around the data to be accessed, rather than the user who is to be served? Is it because the data are permanent and tractable, whereas the user's need is ephemeral and unreliable? No doubt there is truth in this, but the authors of this report detect a more fundamental problem. When someone has a need for information, it is in principle impossible for him (or her) to express quite what it is that is lacking. He can explain the need only by describing the context of his ignorance, expressing it as a kind of semantic structure. And yet it is a structure which contains a gap, an inconsistency, an anomaly: such a structure the authors term an 'anomalous state of knowledge' (ASK). In looking for a novel approach to IR, they not only shift the emphasis to the user: they also wish to detect the anomaly which lies at the heart of the user's 'problem statement'. This requires that items in the database must also be represented as semantic structures: an item is likely to satisfy the user's need if it contains a substructure in which the user's ASK is resolved.

This report describes a design study based on structural analyses carried out on 27 verbal and eight written problem statements obtained from actual users of the University of London's Central Information Services. For comparison, 32 author-produced abstracts were also analysed. *Ad hoc* scores were assigned for three levels of association (adjacent words, words in the same sentence, words in adjacent sentences), and the scores were summed for each pair of non-trivial word-stems. The 40 highest totals, divided into three ranges (strong, medium and weak), were displayed as a graph structure. Each representation was sent to the original user or author, whose views on the correctness of the analysis were elicited by means of a short questionnaire. The response was good, and the results were felt to be

encouraging, although most respondents felt that some strong associations had been undervalued in the analysis.

The crucial questions in all this are whether it is possible to pinpoint the actual anomaly in an ASK, and, if so, whether the anomaly can be exploited in some workable retrieval strategy. The authors make some suggestions. For instance, if a strongly bound group of concepts is linked rather weakly to another strongly bound group, then perhaps it is the weak link between the groups which is the anomaly? An item in the database which has the two groups with a strong link between may be what the user needs.

This is a significant and thought-provoking report, but it really represents no more than a preliminary skirmish. Before even the basic approach can be regarded as well-founded, several questions need answering. How sensitive is the graph-building process to the method used for scoring and summing associations? Will it be necessary (as I suspect) for some local syntactic analysis to be introduced? How reliably will it be possible to detect and represent anomalies in ASKs? Will it be possible to show significant correspondence between an anomaly in an ASK and a non-anomalous region in the structure representing a pertinent item from the database? I for one look forward with interest to seeing these questions answered.

C. D. Paice
Department of Computer Studies
University of Lancaster
UK

ERRATA

Vol 1 No 1 January 1982

Probability of Relevance: A Unification of Two Competing Models for Document Retrieval

S. E. ROBERTSON, M. E. MARON AND W. S. COOPER

p. 15 Formula 11 should read

$$\sum_{\substack{1\leq k\leq K \\ 1\leq m\leq M}} [P(b_k, d_m, R) \log P(b_k, d_m, R) + P(b_k, d_m, \bar{R}) \log P(b_k, d_m, \bar{R})].$$

p. 17 The formula marked (4) should not be numbered.

Information Technology: Research and Development

Aims and Scope

Information Technology: Research and Development is concerned with research into, and development of the use of technology for storing and accessing information. It will cover both theoretical and practical details of information storage and retrieval. In doing so it will attempt to bridge the gap between researcher and practitioner. The intention is to bring together in one journal papers dealing with research into the principles and methods of information storage and retrieval, and the development of implementation of technology for effective access. The journal is intended to cover a wide range of topics from the conceptual problems inherent in the automatic organisation of information to their resolution in new technology, and from basic theory to the embodiment of ideas in operational systems. It will cover such topics as theories of indexing and searching models for processes and structures in information retrieval, methodology of experimental studies and new developments in retrieval technology. It will also publish papers on the implementation details of specific systems, new technology, computer algorithms for information storage and retrieval, and new developments such as Viewdata, videodisk systems, word processing, graphical displays, etc. Papers on the management of libraries or information collections, or concerning traditional manual methods would not normally be appropriate.

Instructions to authors

Original manuscripts that fit the aims, scope and policies of *Information Technology: Research and Development* may be submitted to the Editor.

Manuscripts should be submitted in triplicate (original and two copies). They should be typed, double-spaced, single-side, on International Standard Size A4 paper (or the nearest equivalent standard size in the USA) with a left-hand margin of 40 mm.

The title of the paper together with the name(s) and affiliation(s) of the author(s), and an abstract of 150–250 words should be given on a separate title page. The title should be repeated on page one of the manuscript.

All papers submitted will be refereed and must be written to a high standard of English. If referees require alterations to or a revision of an otherwise acceptable manuscript the author(s) will be responsible for re-typing the paper. Papers up to 10,000 words in length will be accepted by the Editor but those over 5,000 words may be subject to editorial revision.

Tables and illustrations should be given on separate sheets with their location noted in the text. Graphs and diagrams should be drawn in black ink in a form suitable for reproduction without retouching.

Citations in the text should be given either with the date in parentheses:

'. . . as suggested by Hawkins (1980) . . .'

or with name(s) and date in parentheses:

'. . . recently it has been shown (Manola and Hsiao, 1979) that . . .'

Use the form '(Manola et al., 1975)' when there are more than two authors, but list all authors in the references. Quotations of more than one line of text from cited works should be indented and the citation should include the page number of the quotation, e.g. (Heaps, 1965:56).

To aid anonymous refereeing authors should cite their own work impersonally, e.g. '. . . Smith (1979) has suggested . . .' rather than 'I have recently suggested (Smith, 1979) that . . .'.

Footnotes should be kept to a minimum (where they cannot be avoided completely) and should be listed on a separate sheet with their location in the text marked by Arabic numerals.

References should be listed in one alphabetical sequencing following the text in the form shown below:

(a) *journal articles*
(spell out journal titles completely and give issue number when a volume is not continuously paged)
Yu, C.T. and Salton, G (1977) Effective information retrieval using term accuracy. *Communications of the ACM* 20, 135–142.

(b) *monographs*
Heaps, H.S. (1978) *Information Retrieval: Computational and theoretical aspects*. New York: Academic Press.

(c) *papers from conference proceedings, etc.*
Hawkins, C.J.B. (1980) Developments in raw technology. *Microprocessors and intelligence*. (L.J. Anthony, ed.) pp 1–14. London: ASLIB.

(d) *research report*
Manola, F. and Hsiao, D.K. (1973) *A model for keyword based file structures and access*. Washington D.C.: Naval Research Laboratory (NRL Memorandum Report 2544).

Galley proofs will be supplied to the senior author of a paper but only errors in typesetting may be corrected at this stage. No changes in content will be permitted once the paper is set in type: consequently the author(s) should ensure that the paper is submitted in its final form.

The senior author will receive 20 complimentary reprints of the paper after publication. Additional reprints may be ordered: inquiries should be addressed directly to Butterworth & Co., *not* to the Editor.

Manuscripts will be accepted on the understanding that their content is original and that the manuscript has not been submitted for publication elsewhere. However, papers delivered at conferences and meetings may be acceptable if they are significantly extended or altered from their original presentation as a result of further work.

Vol 1 No 2 April 1982

Information Technology: Research and Development

Contents

Printed in Great Britain at the Alden Press, Oxford

ISSN 0144-817X