treated by simply referring to the use of Boolean operators.

The emphasis placed on problems of information *technology* becomes visible, furthermore, through two chapters entitled 'Database Management Systems' and 'Computer and Communications Hardware'; chapters, again, which contain in themselves valuable information but which contribute to the overall impression that in this book *too much* is being offered in a sometimes less than transparent structuring framework.

The contents are completed by chapters entitled 'User Interface', 'Evaluation of Systems and Software' and 'Ideal Systems'. The perusal of these chapters is certainly valuable to anyone having to do with the shaping of Information Retrieval Systems. One obtains a rapid overview of the many things one must think of and, on the other hand, of the possibilities one has for realizing specific wishes.

The chapter on 'Ideal Systems' is certainly worthy of discussion. Yet one may well doubt whether at the present stage of developments the moment has already come to speak not only of desirable improvements but also, as if deciding on absolute values, of ideal systems. Such doubts are only increased by the fact that the emphasis in this chapter is placed, in the end, on the improvement, the optimization, of the *retrieval component*, with the fact tending to recede into the background that the real problem lies in achieving a proper balance between documentation language, indexing principle selected, data bank design and retrieval language used. Omissions in any one of these fields, - particularly in those of the documentation language and/or indexing principle used - can only to a very limited extent be compensated for by improvements in those of the other components. Specifically, as one knows, it has meanwhile been generally accepted by now that Boolean linkages, while admittedly constituting powerful retrieval instruments, fall far short of permitting a solution of any and all problems of precision and recall.

As to the formal presentation of the book there are a number of peculiar features to be noted. Thus, e.g., the table of contents contains no page numbers, and the alphabetic index likewise does not refer to pages but instead to chapter subdivisions. The text is supplemented by a glossary of close to 30 pages which gives brief definitions of the terms listed and points out the context in which they appear in the text in each case.

Of particular value, undoubtely, is the nearly 50 pages long bibliography, access to which is furnished also by a name index and an alphabetic subject index, both all the more useful in view of the fact that the bibliography is subdivided into six sections not broken down according to any criteria that are discernible to this reviewer.

To make an attempt at an overall evaluation: this is beyond doubt a book that was written with a high degree of competence in the subject field concerned and that therefore has a great deal of valuable information to offer. To a teacher it can furnish, in addition, sound didactic suggestions at many points as to the manner in which various subjects can be presented in courses and/or classes. As particularly well done I regard e.g. the presentation of p.50-58 of the processing of documents for indexing purposes, for their storage in data banks and for the generating of various inverted lists for retrieval purposes. Nevertheless the impression remains that the author has tried to present too much in one book and could not bring himself to decide, for the sake of brevity of presentation, to rather impose limits on himself at one point or another.

Winfried Gödert

Prof.W.Gödert, FHS Hamburg, FB Bibliothekswesen, Grindelhof 30, D-2000 Hamburg 13

WEINBERG, Bella Hass (Ed.): **Indexing – The State of the Art and the State of Our Ignorance.** Medford, NJ: Learned Inform.,Inc.1989. 134p. ISBN 0-938734-32-6

This book presents the papers read at the 20th annual convention of the American Society of Indexers on 13 May 1988 in New York City. Its title already strikingly indicates how acutely the situation prevailing in the indexing field is reflected here. The book claims to be a status report on the subject of indexing, particularly the indexing of books. To this claim it lives up one hundred percent. It can be recommended both to the novice in this field and to the oldtimer who frequently is forced to concentrate overly much on his or her daily work. It is recommendable not in the last place because of the valuable references to the latest literature usually given by the various authors.

The excellent index to this book, compiled by Bettie Jane THIRD, contributes a great deal to the book's value. In this index one finds practically everything realized which the authors of the book declare, at one point or another, to be desirable for an index.

The editor made an excellent choice by placing the contribution by Mr. WELLISCH – an appeal of well-nigh insurpassable frankness to the conscience of the profession – at the beginning of the book, as it were as an introduction, admonition and overview. All too often, ignorance and a lack of self-criticism cause distrust and superciliousness to be sown against the indexer, whose work is alleged to be mechanizable or predicted to become so in the near future. A few authors in this book, too, must feel 'touché' by Wellisch's expostulations.

WELLISCH, Hans H: *Literature of Indexing*. Regrettably, indexing still is not yet recognized as a professional task that must be learned if good results are to be achieved. It is still widely assumed (as borne out by a quotation) that literacy and possession of a marker pencil suffice for preparing a book index: English is the dominating language in the modern literature on indexing and abstracting. Most procedures described here are based on word or word stem matching or on proximity and freqency measures, whereas German and Russian authors have a greater preference, in comparison, for the linguistic approach. The few genuine innovations are to be found in the field of truly scientific research on the fun-

damentals of indexing. - Automatic Indexing is a favorite topic in the literature, yet the results are not readily discernible to the naked eye outside the strictly controlled laboratory situation. It has not appreciably advanced beyond the primitive form constituted by Luhn's KWIC Index, the only fully automatic indexing method that works, except that it is neither quick (for the user), nor is it indexing in the sense of *concept* indication, for it is not independent of the vagaries of verbal expression. – Vast, too, is the literature on Artificial Intelligence and its application for automatic indexing. English, German and Russian titles account for roughly equal portions of this literature. But the purportedly successful applications take place in extremely limited domains in the hope that the methods can be extrapolated to cope with real-world problems in information retrieval.

Ignorance with respect to indexing is of a twofold nature. In the one case we are talking of the widespread and – because of the information explosion – practically unavoidable ignorance of what has already been written by other authors. Editors and review boards might bring some relief here. The second type of ignorance is more serious and even dangerous: it is encountered in the

"utopia of a brave new world in which all or most questions will be dealt with by machines, which will also execute the subsequent search and retrieval of information unfailingly and to the satisfaction of the inquirer. The authors of such utopian prognostications seem to be woefully ignorant of the mounting evidence on the limits to totally automatic information retrieval, including indexing as its most crucial element. In view of the inherent unpredictability of our verbal expression of concepts, which is not entirely reducible to rules, how can we ever hope to pin down an information need and to match it automatically with equally variable and forever changing expressions used by those who want to convey information to others? It is the ignorance of these fundamental facts that is so blatantly displayed in the writings of those who would convince us that the solution of the problems of information retrieval will be found in the automatic matching of the words of an inquirer with those recorded in retrievable form by other people. Sprucing up this basic contention with the bells and whistles of frequency, probability, and even (most fallacious, if not fraudulent) measures of relevance does not make it less of a dangerous illusion. It may become a menace both to indexers and index users, because publishers and other sectors of the information community are enthusiastic about the idea. The indexers may be encouraged by publishers to produce their desktop indexes, which hardly deserve that name, by merely tagging so-called "significant" words in the texts. On the other hand, only human indexing, fallible as we know it to be, will not provide the answer to the problems of information retrieval, and all our work on automation as an aid to information retrieval is not doomed to failure. The answer lies in finding the right balance between what machines can do best, always keeping in mind that machines can recognize symbols, but only human beings can recognize meaning including that which is not explicitly expressed by symbols."

What expectations do we vest in the future indexing literature? Still too little is known about the utilization of indexes and the usefulness of various kinds of index entries. More should become known about failures of indexing software so as to accelerate progress in this field toward better, simpler and more user-friendly programs. We also need more studies on the economics of computer-aided indexing. Will humanly-produced indexes carrying a higher price tag force publishers to omit indexes even more often than now, or will they turn to machine-produced indexes of little value? A new retrospect will be due in 1993 on the 25th birthday of the American Society of Indexers when, as we hope, we will have a little more knowledge and a little less ignorance.

THOMAS, Dorothy: *Book Indexing Principles and Standards.* The deeper cause of the widely varying quality of present-day indexes is to be found in the widespread misconception that indexing is merely a matter of extracting words from the book text. Rather, it is also a matter of competently translating antiquated, idiosyncratic and paraphrasing expressions into currently valid terminologies. – Writing a good index is just as much of an art as the writing of good texts. But misconception is compounded when the stylistic rules of text writing are unreflectingly applied to index writing, too, as often happens in ignorance of the requirements imposed on an index. Here, style and format are confused with principles. To be sure, we have excellent compilations at our disposal of the principles by which an indexer should let himself be guided in his work (a pertinent summary and a good bibliography are included in the paper), but a standardization of this process is impossible. All too great are the differences between the various books and the requirements of the index users. A major obstacle to fruitful communication of indexers among one another is the lack of a consensus on the use and meaning of their own technical terms, a shortcoming for the overcoming of which efforts have now been started. – The indexer's real world will always be marked by conflicts with the publishers concerning the trade-off between printing costs an the one hand and index quality on the other hand.

FETTERS, Linda K.: *Indexing Software.* Software available for indexing is divided into two main groups: "Indexer controlled" and "software controlled". The first group includes: a) stand-alone indexing programs, intended for book indexes, b) embedded indexing programs, which process terms embedded in machine readable texts, and c) hypertext programs, which link information in one file to related information in another file.- The second group comprises: a) text retrieval programs, which allow for the retrieval of words, phrases in machine readable texts, b) automatic indexing software, which generates an index based on human-produced lists of terms to be included or excluded, c) expert systems, which perform work guided by rules based on human experience.

In the case of a machine readable text of any form the software offers the possibility of extracting words or phrases directly from the text (e.g. after a a previous marking or after a frequency count) and of using these parts of the text as entries for the index. This software might in particulalr stimulate the indexing of a given company's own business records. Indexers might find a new field of work where a company's internal reports are to be brought into machine-readable form and where hypertext databanks are to be built up. – Looking at the indexing problem from a purely technological perspective, L.Fetters regards it as practically solved ("Now, there are indexing programs to meet every need and style") and considers the mere extracting of words from the texts of books as being equal to human indexing. Accordingly, the indexer is regarded as becoming dispensable before long.

MILSTEAD, Jessica L.: *Database Design: Indexing Applications".*
This paper is intended to aid indexers in deciding whether a generalized data base approach is more appropriate than the use of specialized indexing or bibliographic software for a given indexing task. Word processing is often not a viable alternative. But is can serve for editing the output from one of the other software packages and format it so flexibly as is generally not possible with one of the other commercial systems mentioned. The strengths and weaknesses of the various approaches are described from a practitioner's point of view. Surprising failures are reported on as well. The indexer in quest of advice will find many useful suggestions in this contribution for the selection of his software and the shaping of his databank.

PRESCHEL, Barbara M.: *Indexing for Print, Online, and CD-ROM.* Although print indexes and electronic indexes to online data bases share many characteristics, each has certain unique capabilities and felicities. Differences exist particularly with regard to browsability, proximity of the full texts, pre- and post-coordination of terms, precision of terms, depth and specificity of indexing, use of controlled vocabulary terms, and customization of index point of view. To offset its inflexibility, the printed index also has a few advantages which the electronic indexes lack, quite apart from its easy accessibility for the user. These advantages assure it of its continuing (and possibly even increasing) importance in practice. Even an experienced indexer will profit from this contribution when starting to index for one of the three storage media.

BRENNER, Everett H.: *Vocabulary Control.* Present day information systems still have major shortcomings, and all efforts should be directed at the development of suitable software with the aim of a fully automatic information system. The queries in the original wording and the texts in natural language should constitute the only input into such a system, and a machine program (fantastically efficient in that case) should completely replace the human operator in the analysis of texts and the processing of queries. Brenner dissociates himself emphatically from the view taken elsewhere by E.Svenonius that in order to overcome the current shortcoming we should first and foremost improve thesauri and indexing with them. The comments made on the subject of vocabulary control are essentially limited to the remarks that it is more expensive than the processing of natural language and that for the time being it is still regarded as necessary, if on a vastly reduced scale. – Brenner bases these hopes in large measure on work done by himself in which he dispensed with part of the indexing operations and which nevertheless furnished – to him – satisfactory results. One regrets the absence of information on the kinds of queries employed here, on the exact kind of fully automatic procedure he places his hopes in, and on the reason why these hopes are justified. Brenner's satisfaction with his own results and his hope for future fully automatic procedures are only explainable to this reviewer by assuming that the requirements Brenner imposes on an information system are correspondingly low and seem to be largely reduced to mere word matching (cf. in this connection the remarks by Wellisch).

ANDERSON, JAMES, D.: *Indexing and Classification: File Organization and Display for Information Retrieval.* As far as file organization is concerned there is no essential difference between indexing and classification, as both are based on the same basic operation. The function of the thesaurus in this connection is based on the term groupings inherent to it in the sense of generic, partitive and associative relations. – Texts may likewise be grouped by attributes, i.e. on the basis of keywords, assigned descriptors, citations, author names, etc. Grouping on the basis of such text attributes can thereupon be performed by a variety of algorithmic procedures. – In view of the vast quantity of printed material and the many advantages of searching in printed indexes the question of the display has lost none of its topical interest. A differentiation is made between the look-up and the browsing functions of a printed index. The former requires a predictable localization in the index of the object or information looked for, whereupon the searcher should be granted the possibility to conduct, through browsing, a more specific search. Of essential importance here is the question to what extent the details of a title can be presented sufficiently predictably and clearly through consistent use of subheadings under a lead term and through a syntax of entries. In this respect there are major differences between the various techniques existing. – The displays of relational indexes, keyword indexes (KWIC, KWOC, KWAC and permuted indexes) and subject heading indexes are compared among one another. They permit a good overview for browsing even when the syntax is selected "ad hoc" and is not predictable. - Computers have provided new index production techniques in case indexing has taken place accordingly. This is the case in the various variants of "String Indexing", in which the terms for a text are linked with each other. SPINDEX, NEPHIS, PRECIS and POPSI are pertinent examples. The latter system uses faceted index strings, thus permitting a particularly wide variety of printed indexes and almost any building block sequence required in display. Faceted indexing also controls in an optimally reliable manner the selection of the essence in the indexing itself and captures the essential conceptual structure of a field. One must warn against formal, nonrealistic groupings which not only are of no use to the searcher but even lead him or her astray. The continuing importance of printed indexes is far too little realized, and one still knows far too little about the possibilities offering themselves here.
CROFT, Bruce W.: *Automatic Indexing.* The variants practiced today are looking at different evidence for relevance. This fact has given occasion to the experimental project "I³R", in which it is being tried to utilize the sum of the results from various sources for relevance testing. The intellectual process taking place when a user looks at the results from various sources and makes his or her selection is to be simulated by machine. The relevance of a document is best assessed by looking at evidence from all available sources. The basis for these experiments is formed by the knowledge base consisting of the docu-

ments with their desriptors somehow assigned to the texts or already occurring in them naturally. The search strategy is to make use of the links between documents and terms. The goal consists in "knowledge-based retrieval", in which the representation of the query and the documents are built by natural language processing. The knowledge base will be used to see which things are related to which other things". But it is impossible to build thesauri, controlled vocabularies and knowlege bases automatically. – On the prospects for success of such an approach aiming at full automation striking remarks were made, in this reviewer's opinion, by Wellisch in his introductory contribution.

SARACEVIC, Tefko: *Indexing, Searching and Relevance.* How do searchers perform searches, and how do user context and type of questions affect the outcome of searches? This study was about people – users and searchers - and about processes – question asking and searching. 40 users with one question each, 36 outside professional searchers and 3 project searchers participated in this study. Each query was processed by 5 different searchers. The queries were directed in each case to the most appropriate DIALOG files. The union of the nine searches each for each single question was sent to the users for evaluation. They had to indicate in how far each reply was relevant and to evaluate their usefulness in five different measures. The results were expressed in relevance odds that the retrieved answer would be relevant as opposed to not relevant in user judgment.

Agreement in the selection of the search term was only 27%. Only 17% was the degree of agreement in the responses to the nine searches. Hence each searcher had evidently retrieved a different portion of the file. Some responses were given only once, others several times. Only 3% of all relevant items were found four or five times, but the more frequently a text was found by the 5 searchers (optimally 5 times), the higher was the probability that this text was also classified as relevant.

The present design of online subject access, be it through library catalogs or online retrieval systems, does not accomodate human variability in searching (or indexing). This calls for radically different design principles and implementations in order to accomodate the observed patterns, interactions, and differences in human behavior, of which the overlap findings are one of the most important manifestations.

LIPETZ, Ben-Ami: *The Usefulness of Indexes.* The judgment as to the usefulness of an index is always subjective to a high degree with respect to the judging person and also to the purpose it serves and the time required for building it. Usefulness should also be regarded from the point of view of the estimated costs saved through finding or caused by non-finding. For index quality there is no measuring quantity, but fairly rational judgments on the suitability of an index for specific user groups are entirely possible. The judgment of the reviewers, librarians and indexers themselves might have an influence on all decisions concerning the index. Regrettably, a great many book reviewers pay no attention at all to the index of a book. - Publishers look at indexes from the point of view of whether they promote sales and to what extent the higher costs of the book they entail are thereby more than compensated for. – In the bibliographic database services, books are strongly neglected, because they sell their services by the mere number of publications covered and not by completeness of coverage. But in the future indexes will be included in online data bases and help the publishers increase sales of the books, because they will then be found far more frequently as relevant answers. Modern technology can contribute to a situation in which a great variety of different index forms tailored to the given needs can be brought into being with, on the whole, far greater usefulness and timeliness. Still being far removed from being an exact science, indexing deserves to be given more intensive efforts.

THOMAS, Dorothy: *History of the American Society of Indexers.* Origin and growth of the ASI are described, as well as the change in the tasks and the corresponding specialization of the Society's members. The history is rounded off by a summarization of the Society's past and present presidents, prizewinners and main conferences.

Reflecting as it does the currently large controversion of the results and experiments in the indexing field, this book is readable for anyone wishing to obtain an overview of the state of discussions. One feels transported back into the age of Galilei when it was an open question whether the falling speed of an object was or was not dependent of its weight. At that time, too, either side was able to adduce empirical evidence for its opinion, depending on the chosen experimental conditions. It will be a task of the future to find theoretical foundations for resolving the contradictions in the indexing field. Here, too, they are undoubtedly to be found in the differences in the prevailing working and testing conditions.

Robert Fugmann

Dr.R.Fugmann, Alte Poststr. 13, D-6270 Idstein

RICHTER, Noe: **Grammaire de l'indexation alphabétique.** (Grammar for alphabetical indexing). Le Mans: Bibliothèque de l'Université du Maine 1985. 4thed. 155p.

Noe Richters Einführung in den alphabetischen Schlagwortkatalog – dessen grundlegende Intentionen, Methoden, Verfahrensweisen, Instrumente und Fragestellungen – richtet sich auch an die (französichsprachigen) Praktiker der Sacherschließung, versteht sich aber vornehmlich als Lehrbuch für (französischsprachige Studenten des Bibliothekswesens[1]. Die vorliegende Monographie stellt eine wesentlich überarbeitete Auflage der 1984 erschienenen Auflage dar, der ihrerseits zwei Auflagen unter dem Titel "L'indexation alphabétique des matières" (April und Oktober 1980) vorausgegangen waren.

Die Einführung, die auch für den völlig unerfahrenen Leser verständlich sein will, vermittelt immer wieder elementare und häufig auch unumstrittene Grundkenntnisse – etwa in dem Vergleich zwischen systematischem Katalog und Schlagwortkatalog. Originalität ist dabei nicht beabsichtigt und darf angesichts der anvisierten Zielgruppe auch nicht erwartet werden. Der spezialisierte Leser, der vielleicht hier vergebens neue Erkenntnisse suchen wird, mag aber immerhin einige der didaktischen Vorzü-