

BOOK REVIEWS

COMPUTERS IN LANGUAGE RESEARCH 2. (Trends in Linguistics, Studies and monographs 21)

Walter A. Sedelow, Jr. and Sally Yeates Sedelow, Editors

Berlin: Mouton Publishers, 1983, vii+301 pp.
ISBN 90-279-3009-0; DM 126.00, \$57.30

The editors of this book are both at The University of Kansas, Lawrence. He is a professor of computer science and sociology, and she is a professor of computer science and linguistics. The influence of all these disciplines can be found in this book, which is in the Trends in Linguistics series edited by Werner Winter.

The editors of the volume being reviewed here argue that the amount of information processing capability that has been brought to bear on the science of language is very small. The volume is divided into two parts. The first deals with formalization in literary and discourse analysis; the second with notating the language of music and the rhythms of speech.

In the first paper, "Computer Criticism", John Smith explains that computer applications for language and literature studies have either involved the production of aids with which most of us are familiar (e.g., dictionaries or concordances) or have involved the use of the computer for the analysis of specific works. He explains, correctly I think, that the second offers a new methodology that has been largely overlooked because of lack of experience on the part of those who have traditionally undertaken literary analysis. Smith correctly points out that the computer can only augment the critics' judgment, not replace it, but he makes a convincing case that such augmentation is of value.

Yorick Wilks offers a paper on the relationship between the development of artificial intelligence and the development of machine translation: "Machine Translation and the Artificial Intelligence Paradigm of Language Processes". He discusses three "intractables" of natural language that many felt would be cleared up by now: word sense ambiguity, case ambiguity, and referential ambiguity. The paper presents examples from artificial intelligence which make progress toward solving these problems.

Bertram Bruce contributes a paper, entitled "Belief Systems and Language Understanding", that represents an effort to use the computer to cope with intentionality. It begins with an explanation of the role of belief in natural language understanding and reviews two important components of a belief system: social actions and patterns of behavior. The paper deals with what is, to me at least, one of the most fascinating topics in computational linguistics, **credal regression**, that is, beliefs about beliefs.

The second part of the book deals with encoding and opens with a short paper by the editors in which they discuss the premise that the distinction between number

symbolism and non-number symbolism may be just as important as the distinction between spoken and written language. The first paper in the section is "The DARMS (Digital-Alternate Representation of Musical Scores) Project: Implementation of an Artificial Intelligence Language for the Representation of Music". The relationship between music, language, and mathematics has been demonstrated many times, even before the introduction of the computer. However, until modern times there was no means of preserving musical performance and therefore no means of analyzing or representing it in any systematic way. The DARMS project is an attempt to specify the representation of music notation as a consistent and complete artificial language, namely DARM, and to implement a set of DARMS translator programs, or compilers. The article presents an historical evolution of the project, a condensed grammatical description of the language, and representative psycholinguistic problems encountered during implementation. One of the interesting conclusions that can be drawn is that our understanding of the semantic structure of musical notation is nearly as impoverished as our understanding of the notation of natural language.

The final paper in the text is "Pausology" by Daniel C. O'Connell and Sabine Kowal. According to the authors, this is a topic that, until recently, has received little attention because of the lack of accurate and reasonable, inexpensive voice recording equipment. While the paper does provide a comprehensive history of the topic, it has far less to do with computing than it does with pausology. Much of the research documented in the paper undoubtedly made use of computers as research tools, but there is little direct description of those efforts.

What is the overall effect of these papers, which have been grouped together under the title *Trends in Linguistics: Computers in Language Research, Volume 2?* The overall quality of the papers is good, with some, of course, seeming stronger than others. The papers are generally well researched and well written. However, this reader had a great deal of difficulty ascertaining the overall effect the volume was designed to have, if any. As a collection, the volume seemed incohesive. While the articles in the first section do share the common topic of language, Bertram Bruce's "Belief Systems and Natural Language Understanding" seems only incidentally related to the other two in that section. The situation in section two, which is entitled "Notating the language of music, and the (pause) rhythms of speech", is even more dramatic. The article on the DARMS project and the article on pausology may be related in significant ways, but the editors never make that relationship clear.

It is, of course, possible that a logical connection does exist between the materials in the text and this reviewer simply failed to perceive it. If that is the case, then I would argue that the editors should have made more

effort, especially in transitional sections, to make these connections clear.

It is also the volume's incohesiveness that makes this reviewer question the audience for the volume. Each paper seems directed at a specific audience and it is difficult to find a more general audience that would find the text of interest. One might argue that it could be used as an introduction to all the topics presented, but in this reader's opinion the papers are at a level that requires a background of some sophistication, though more in linguistics than in computing. I am at a loss to suggest any course in which the book might be used as a text. I think there are a number of quality papers in this volume; perhaps it can be best used to reference the included papers on an individual basis.

Brian D. Monahan

Department of Computer and Information Sciences
Iona College
New Rochelle, NY 10801

INFORMATION RETRIEVAL EXPERIMENT

Karen Sparck Jones, Editor

London: Butterworths, 1981, vii+352 pp.
ISBN 0-408-10648-4; \$69.95

This book is essential reading for anyone interested in information retrieval. It describes in detail how to design and carry out experiments in information retrieval and how to analyze the results. Why should other workers in computational linguistics read this book – aside from the fact that, like all researchers, we are consumers of information retrieval? Because it includes a fascinating discussion of the philosophical and practical problems of performing experiments, tremendously valuable at this point when we are all supposed to be taking part in the new experimental computer science.

This book also provides a guide to the main stream of research in bibliographic information retrieval over the last twenty-five years. In this paradigm, documents are represented by lists of terms contained in the title, the abstract, or, occasionally, the whole text, along with frequency counts. Queries are represented by similar statistics and then documents are located by matching the word list from the query with the lists from the documents. Many refinements are possible: using phrases instead of single words, adding a thesaurus, controlling the indexing language, using feedback from the user. The experiments described in this book have helped to determine which of these strategies are of value and thus have been crucial in the development of the sophisticated systems of today.

There are fifteen chapters written by thirteen different scholars. The fact that Karen Sparck Jones could command the efforts of the best minds in information

retrieval testifies to the esteem in which she is held by this community. In spite of the multiplicity of authors, the book functions as a whole, with each chapter building on the preceding ones. The editor clearly kept a strong hand on every step of the writing process, through all three parts of the book.

The four chapters in Part I are concerned with fundamental issues in the design of experiments. Robertson sets the scene by discussing general methodology for retrieval system testing. Van Rijsbergen focuses on the issues involved in characterizing and measuring retrieval system effectiveness. Belkin looks at the other side of the problem – how to define human information needs and determine whether a system is satisfying them. Tague's chapter, the longest in the book, describes the practical decisions that have to be made at every stage in designing and carrying out an experiment. It could well serve as a handbook for the would-be experimenter.

The second part discusses the problems of applying this methodology in diverse practical situations. Some experiments are directed at evaluating the performance of systems that are actually in operation. Others are essentially laboratory tests designed to investigate alternative methods of retrieval effectiveness. Lancaster writes about the problems of "Evaluation within the environment of an operating information service". The system he is talking about is a largely manual one. Barraclough is trying to solve the same problems but in an automated, largely interactive system. Keen tackles the question of designing laboratory tests for manual systems, while Oddy looks at the same question for fully automatic systems. Heine defines some situations in which simulation tests can give the best results. Cooper argues for *gedanken* experiments as a substitute for, or at least a preliminary to, expensive tests with actual data.

The third and final part of the book contains descriptions of the major experiments in information retrieval over the last twenty-five years. The first chapter contains a survey by Sparck Jones of the development of retrieval system testing during this period. The second chapter, also by Sparck Jones, describes in detail the influential series of experiments at Cranfield. The final chapter, by Gerard Salton, describes his own experimental system, the SMART system, which has served as a testing ground for the most interesting developments in the United States for the last fifteen years.

The whole book is a fascinating blend of philosophy and practice centering on the problem of how to design and carry out carefully controlled, valid experiments on data that presents all the bewildering peculiarities of human language.

Martha Evens

Illinois Institute of Technology
Chicago, IL 60616