COMPUTATIONAL MODELS OF DISCOURSE

The MIT Press Series in Artificial Intelligence

Artificial Intelligence: An MIT Perspective, Volume I: Expert Problem Solving, Natural Language Understanding, Intelligent Computer Coaches, Representation and Learning edited by Patrick Henry Winston and Richard Henry Brown, 1979

Artificial Intelligence: An MIT Perspective, Volume II: Understanding Vision, Manipulation, Computer Design, Symbol Manipulation edited by Patrick Henry Winston and Richard Henry Brown, 1979

NETL: A System for Representing and Using Real-World Knowledge by Scott Fahlman, 1979

The Interpretation of Visual Motion by Shimon Ullman, 1979

A Theory of Syntactic Recognition for Natural Language by Mitchell P. Marcus, 1980

Turtle Geometry: The Computer as a Medium for Exploring Mathematics by Harold Abelson and Andrea diSessa, 1981

From Images to Surfaces: A Computational Study of the Human Early Visual System by William Eric Leifur Grimson, 1981

Computational Models of Discourse Edited by Michael Brady and Robert C. Berwick, 1983

Robot Motion: Planning and Control by Michael Brady, John Hollerbach, Tomas Lozano-Perez, Matthew Mason, and Timothy Johnson, 1983

COMPUTATIONAL MODELS OF DISCOURSE

Edited by Michael Brady and Robert C. Berwick

Contributors:

James Allen
Robert C. Berwick
Jerrold Kaplan
David McDonald
Candace L. Sidner
Bonnie Lynn Webber

The MIT Press

Cambridge, Massachusetts
London, England

PUBLISHER'S NOTE

This format is intended to reduce the cost of publishing certain works in book form and to shorten the gap between editorial preparation and final publication. The time and expense of detailed editing and composition in print have been avoided by photographing the text of this book directly from the author's computer printout.

Printed in the United States of America.

Library of Congress Cataloging in Publication Data

Main entry under title:

Computational Models of Discourse

(The MIT Press series in artificial intelligence)
Bibliography: p.
Includes index.
1. Artificial Intelligence. 2. Linguistics--Data processing. 3. Speech processing systems. I.Brady, Michael, 1945- . II. Berwick, Robert C. III. Allen, James. IV. Series.
Q335.C56 1983 001.53'5 82-20402
ISBN 0-262-02183-8

CONTENTS

THE AUTHORS

James Allen
Assistant Professor
Department of Computer Science
University of Rochester
Rochester, NY 14627

Robert C. Berwick
Assistant Professor
Department of Elec. Eng. and Computer Science
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge MA 02139

J. Michael Brady
Senior Research Scientist
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, MA 02139

David Israel
Research Scientist
Bolt, Beranek, and Newman Inc.
50 Moulton Street
Cambridge, MA 02139

S. Jerrold Kaplan
Vice President, Business Development
Teknowledge Inc
525 University Avenue
Palo Alto, CA 94301

David McDonald
Assistant Professor
Computer and Information Science
University of Massachusetts
Amherst, MA 01003

Candace L. Sidner
Research Scientist
Bolt, Beranek, and Newman Inc.
50 Moulton Street
Cambridge, MA 02139

Bonnie Lynn Webber
Associate Professor
Department of Computer and Information Science
Moore School of Electrical Engineering
University of Pennsylvania
Philadelphia, PA 19104

FOREWORD

Michael Brady

It should be noted at the outset that my personal research interests in artificial intelligence are in vision and robotics, not in linguistics. Two years ago, however, at the time that I was joining MIT, my general reading of the natural language literature in artificial intelligence suggested to me an undercurrent of change in computational linguistics analogous to that which has taken place, for example, in vision [Brady 1981]. This book on computational models of discourse stems from a series of conversations between myself and the other editor, Robert C. Berwick, that explored that change and the closeness of the analogy with vision.

What is the nature of the change? In a nutshell, it seems to me that artificial intelligence is crystallizing into more or less independent subdisciplines as an inevitable side-effect of maturing. Specialist journals have appeared, catering to specialized subdisciplines from linguistics to robotics. Even the field's traditional non-specialist journal *Artificial Intelligence* recently devoted a special issue to vision [Brady 1981]. It recognized that artificial intelligence researchers outside vision increasingly feel out of touch with work in that area. On the other hand, increasing numbers of perceptual psychologists want to become familiar with the computational approach to vision. Many artificial intelligence researchers feel similarly out of touch with other subdisciplines, notably robotics, search, and automated deduction.

Before exploring the symptoms of change, let us consider the state of natural language understanding research, say up to the time of the Schank and Colby's collection [Schank and Colby 1973], or the influential workshop on *Theoretical issues in natural language processing.* The vast expenditure of time and money on the machine translation projects of the 1950's and early 1960's, and the detailed formal mathematical study of parsing in the 1960's, provided little real insight into the problems of how natural language could be understood by, or through the use of, computers. The first wave of efforts in what is nowadays called computational linguistics had a considerable impact both on artificial intelligence and on a limited number of conventional linguists. The work of Schank, Winograd, Wilks, Woods, and their associates, suggested that it was possible to build a (huge) computer program capable of interesting linguistic behavior.

Although the emphases of different research efforts were in different aspects of language, they shared the characteristic that they resulted in entire (mostly) working systems. All the systems were required to deal with a wide range of linguistic challenges. They included a parser, a lexicon, and embodied proposals about the representation and role of semantic and pragmatic information in understanding and responding to an English sentence. The detailed interaction between the various subsystems was a central concern. In Winograd's system, foi example, the fact that the sequence of processing could not be specified in advance was claimed as a major feature. The program SHRDLU contributed to the theory that sophisticated process interaction is central to modeling the flexibility of human thought, perception, and language understanding. In retrospect, however, these approaches to computational linguistics seem to be overly concerned with mechanism.

Inevitably, in order to achieve a total, working, natural language understanding program, early systems were forced to ignore, gloss over, or otherwise compromise on, many aspects of language understanding. The treatment of time, mood, purpose, theme, the determination of focus and reference, as well as many other issues, were accorded only a preliminary treatment. Some of the main insights which informed the construction of these systems were computational, and some of the main lessons which the authors claimed could be learned from their research, were about computational issues. the architecture of process interactions and the design of representations for semantic processing. For this reason, the detailed operation of the early natural language understanding systems and the issues they addressed were largely accessible to the general artificial intelligence community. At the interdisciplinary workshop on *Theoretical issues in natural language processing* [Nash-Webber and Schank 1975], six of the eight sessions concerned memory and knowledge representation. Most of the papers in those sessions could have just as easily been presented at a workshop on vision or reasoning (many were). Interestingly, only one of the forty papers presented at that workshop is referenced in this volume.[1]

What are the symptoms of change in computational linguistics that I referred to earlier? First, like vision, robotics, search, and formal reasoning, it has become

---

1. To be fair, some of the systems described at TINLAP-1 were refined and re-presented at TINLAP-2. Several references are made to TINLAP-2.

increasingly technical. In vision and robotics technical typically refers to sophisticated mathematical analysis. Although linguistics is, for the most part, (currently) less demanding mathematically than vision, there are, as always, exceptions. Berwick's grammatical analysis (and many similar analyses) of the Marcus parser and the McDonald generation program (chapter 4) is a model of precision. Similarly, Webber's (chapter 6) representation of the possible meanings of noun phrases fully exploits the power of the predicate calculus. Mathematics aside, technical refers to the precision and close attention to detail that is a feature of much current work in computational linguistics.

Second, though related to the first point, the problem with which any individual researcher is concerned in any particular project, seems to have narrowed considerably, with a corresponding increase in depth of analysis. The individual chapters in this volume restrict their attention to response generation, the determination of reference and focus, and the scope of quantifiers. In vision, people work on such problems as directional selectivity, the shape of subjective contours, and binocular stereo. The subject matter is primarily limited to what might be considered modules in the human's visual or linguistic system rather than being limited by a domain of application that potentially requires the deployment of the full panoply of linguistic or visual abilities. This does not imply that the epistemological base of a piece of work can always be completely unrestricted, as the paper by Allen in this volume illustrates. Domain restrictions are mostly a reflection of the lack of muscle and inappropriate architecture of today's computers. As the subject matter of individual research papers has become more restricted and specialized, so they in turn have become of less interest to artificial intelligence researchers whose primary interests are in other subdisciplines.

Recent papers in computational linguistics are also more demanding to read. One reason for this is immediately apparent to the casual artificial intelligence reader. Considerably more than lip service is paid to *non-* computational linguistics and vision. A more specialized background is assumed of the reader. Although the architects of early systems referred to the linguistics literature, they typically did so in general terms. Indeed it was even suggested by some authors that linguistics researchers simply had not uncovered ideas which were precise enough to constrain the detailed construction of a program or against which to evaluate its detailed behavior. There seems to have been an implicit assumption that the new conceptualizations introduced by computation were so radically different from anything that had been used previously in modeling language understanding that it was appropriate to make a completely fresh start.

In many recent papers, including those in this volume, extensive reference is made to detailed psychological and linguistic data. Results cited are used as evidence in support of the importance or appropriateness of a piece of work, as a source of constraint, or to justify some constraint or design decision, for example of a representation. Mitchell Marcus, for example, claims to have designed his parsing system PARSIFAL [Marcus 1980] to embody the constraints on human parsing which Chomsky has uncovered. The constraints are extremely detailed, concerning technical issues such as subject raising and embedded complements. Moreover, Marcus makes a number of detailed claims about program organization which he claims are implied by Chomsky's linguistics findings. The flow is not one way however. Computation contributes powerful ideas about representation and process, even if they are not omnipotent. While, the coupling of artificial intelligence with linguistics and psychology is not new, it has perhaps become more earnest and detailed.

One consequence of this change in computational linguistics is the tendency, remarked earlier, to publish in specialized journals, and to have the work recognized by the linguistics, philosophy, and psychology communities. This requires that computational linguists be able to discuss their ideas at least in part in terms that regular linguists are familiar with. A growing number of computational linguists have a formal qualification in linguistics, philosophy, or psychology.

The editors chose the contributions to this volume in part to illustrate the issues raised above, and in part to form a coherent whole. Jerrold Kaplan and James Allen provide rather different views of what is involved in understanding the meaning of a question in a discourse. An issue of deep and immediate concern for both of them is the need to generate responses, to participate in the discourse. David McDonald's contribution can be critically evaluated in the context of their work, and conversely. Also, they need to determine reference and focus, for, as both of them point out, much is left unstated in a discourse, and a participant has to exploit his or her knowledge to the full to keep talking. These are the questions studied by Candace Sidner and Bonnie Webber. Robert Berwick's introduction explores these interactions in more depth, and provides an overall context and critical review of the other five contributions. The overlap between the separate contributions is apparent even from the references cited; accordingly we have provided a unified bibliography.