Briefly Noted

Parallel natural language processing

Geert Adriaens and Udo Hahn (editors)
(University of Leuren and University of Freiburg)

Parallel Natural Language Processing is an application-oriented text addressing parallel implementations of aspects of natural language understanding. Although the focus is not on any particular cognitive theory, attention is paid to human performance data throughout. The implementation issues and difficulties considered provide a grounding and a checkpoint to balance the existing body of more theoretical work.

The book is a collection of papers, each chapter standing on its own yet each also tied in with several other chapters in the topics addressed and the aspects examined. In general, for each language aspect included, at least two papers cover the topic from slightly differing viewpoints. Additionally, most topics are balanced by a contrasting point of view; for example, formal context-free grammars versus parsing of free speech with mistakes, and complete story understanding versus single noun-phrase shades of meaning. Each chapter is focused and thus manageable, enabling the authors to make some useful observations about their topic. The book as a whole thus serves as a look at the entirety of natural language understanding and the editors have done a good job indeed of encasing and covering this topic.

The first, rather lengthy, chapter provides a good review and overview of the topic and issues at hand. Starting with a look at psycholinguistic and cognitive arguments both for and against autonomous components versus interactive models of natural language processing, the editors conclude that parallel implementations of language processing should be further explored. This is followed by a discussion of parallelism from a computer science perspective: computing models, architectures, operating systems, and programming languages. A review of parallelism in natural language processing follows, with an eye toward the issues raised in the computer science review. A 24-page bibliography provides a thorough reference for the plethora of topics covered in this chapter.

The remaining 12 chapters of the book are individually authored papers covering parallel natural language processing. The first series of papers focuses on context-free parsing, starting with a theoretical account of the advantages obtained through parallelism, and results of implementations on parallel hardware. Further proposed implementations are presented using object-oriented systems and connectionist paradigms. Finally, parsing is considered as a constraint satisfaction and energy minimization problem.

The next series of papers refocuses on the interaction perspective, considering various methodologies: connectionism, concurrent processes, frame-based actors and bulletin boards, and object-oriented functional programming. Finally, the book closes with an examination of language generation using connectionist and parallel unification models. The interested reader may refer to Chapter 1, Section 5.3 for a thorough overview of individual chapter contents.

Due to the implementation focus, the reader may need a strong computer science background to get the most from this book, which assumes more than passing familiarity with a wide range of topics such as object-oriented and functional programming, deadlock avoidance techniques, and efficient implementations of Cocke-Younger-Kasami parsing. However, in general each concept is briefly introduced so the attentive reader can follow the specific topic even without a formal understanding.—*Jeanne Milostan, University of California, San Diego*

Natural language understanding (second edition)

James Allen
(University of Rochester)

In the latest ACL *Survey of Computational Linguistics Courses* (Dorr 1994), the first edition (1987) of James Allen's textbook *Natural language understanding* was by far the most frequently cited reference, being mentioned

more than twice as often as any other. The publication of the second edition of this book should consequently be welcome news for faculty who teach courses using it as a required or recommended text or for selected readings.

The book has been substantially revised and expanded, with new chapters added, and outdated material removed. The overall organization remains intact, however, with major divisions into three parts covering syntactic processing, semantic interpretation, and context and world knowledge. The original Part IV, 'Response generation,' has been omitted, and the material in its two chapters on question answering and generation has been integrated into other chapters. There is a new appendix on speech recognition and spoken language.

A major change in Part I, 'Syntactic Processing,' is the switch from augmented transition networks to feature-based context-free grammars as the primary formalism. In addition, a new chapter on statistical methods complements the existing material by introducing basic probability theory, part-of-speech tagging, and probabilistic context-free grammars. The focus in Part II, 'Semantic interpretation,' is now more on underlying principles and issues, and less on specific computational techniques. Allen has extended his logical-form language to handle more-complex sentences, and a chapter on ambiguity resolution has been added that covers statistical word-sense disambiguation and semantic preferences as well as rule-based methods. Part III, 'Context and world knowledge,' has been generally updated to reflect work that has been published since the first edition appeared.

The software that accompanies this edition features a bottom-up chart parser and a set of grammars and lexicons that are keyed to the examples in specific sections of Parts I and II. Unix, Macintosh, and DOS versions in standard Common Lisp are available by anonymous *ftp*.

In the past twelve years I have taught an introductory graduate course in computational linguistics five times, most recently in early 1995, using five different textbooks: successively, Winograd (1983), Grishman (1986), Allen (1987), Smith (1991), and the book reviewed here. Offered this year at the CUNY Graduate Center, the class attracted students from computer science, linguistics, and psychology, whose backgrounds in linguistics and computation ranged from practically none to solid. Allen's new book was supplemented with recommended readings from Grosz, Sparck Jones, and Webber (1986) and suggested readings from Pereira and Grosz (1994). Part I was covered in its entirety along with selected chapters in Parts II and III. The laboratory component of the course required the students to use existing software. Assignments included a morphological processor, a tagger, and several parsers, both those provided by Allen and others accessed by e-mail.

Overall, Allen's second edition deserves decent marks. The updated material, particularly the emphasis on feature-based CFGs and the chapters on statistical methods, was badly needed to bring the coverage in line with current work in the field. The print quality is also a significant improvement over the first edition. Asked to evaluate the text at the end of the course, my students gave it an average 3.3 rating on a five-point scale with 5 at the top.

Two areas of remaining weakness are the notational system and the software. The notational system that Allen develops for describing grammars and logical form, while used consistently throughout the book, is not identical to any other in the literature (a perennial problem). Its nuances are not always explained adequately, and even the best students found it obscure and hard to follow at times.

For my class, the Unix version of Allen's software was installed on a network of IBM RS-6000s running AIX and CLISP, a freeware Common Lisp. The parsers worked fine, but the grammars and lexicons were primitive. In some cases, the rules and vocabulary for more advanced sections didn't include all of the features of earlier sections, making it difficult to run a suite of test sentences repeatedly. Only the most dedicated hackers were successful in enhancing the system by modifying the grammars and lexicons. One student managed to install the package on a Macintosh, but it required an inordinate effort because the Unix format had not been uniformly replaced in the Macintosh version. The software package has been upgraded frequently since it was tested early in 1995, and one can hope that these deficiencies have

been remedied.—*Virginia Teller, Hunter College and the Graduate School, The City University of New York*

References

Allen, James (1987). *Natural language understanding*. Benjamin/Cummings.

Dorr, Bonnie J. (1994). *Survey of computational linguistics courses*, second edition. Association for Computational Linguistics.

Grishman, Ralph (1986). *Computational linguistics: An introduction*. Cambridge University Press.

Grosz, Barbara; Sparck Jones, Karen; and Webber, Bonnie (editors) (1986). *Readings in natural language processing*. Morgan Kaufmann.

Pereira, Fernando and Grosz, Barbara (editors) (1994). *Natural language processing*. The MIT Press.

Smith, George (1991). *Computers and human language*. Oxford University Press.

Winograd, Terry (1983). *Language as a cognitive process. Volume I: Syntax*. Addison-Wesley.

Semantics and the lexicon

James Pustejovsky (editor)
(Brandeis University)

Dordrecht: Kluwer Academic Publishers (Studies in linguistics and philosophy, edited by Gennaro Chierchia, Pauline Jacobson, and Francis J. Pelletier, volume 49), 1993, vii + 416 pp.
Hardbound, ISBN 0-7923-1963-X, $140.00, £92.00, Dfl 225.–

"The goal of this book is to integrate the research being carried out in the field of lexical semantics in linguistics with the work on knowledge representation and lexicon design in computational linguistics. Rarely do these two camps meet and discuss the demands and concerns of each other's fields. Therefore, this book is interesting in that it provides a stimulating and unique discussion between the computational perspective of lexical meaning and the concerns of the linguist for the semantic description of lexical items in the context of syntactic descriptions. This book grew out of the papers presented at a workshop held at Brandeis University in April, 1988."—*From the preface*

The contents of the volume are as follows:

Part I: Fundamentals of lexical structure
"X-bar semantics," by Ray Jackendoff
"The syntax of metaphorical semantic roles," by George Lakoff
"Levels of lexical representation," by Malka Rappaport, Mary Laughren, and Beth Levin
"Case marking and the semantics of mental verbs," by William Croft
"Type coercion and lexical selection," by James Pustejovsky

Part II: Mapping from lexical semantics to syntax
"Nominalization and predicative prepositional phrases," by Jane Grimshaw and Edwin Williams
"Adjectives, nominals, and the status of arguments," by Robert J.P. Ingria and Leland M. George
"Unaccusativity in Dutch: Integrating syntax and lexical semantics," by Annie Zaenen
"Verbs in depictives and resultatives," by T.R. Rapoport
"Explicit syntax in the lexicon: The representation of nominalizations," by Tom Roeper

Part III: Computational models of lexical knowledge
"Lexical structure and conceptual structures," by John F. Sowa
"Lexical semantic constraints," by Dan Fass
"Lexical and conceptual structures for knowledge based translation," by Sergei Nirenburg and Christine Defrise
"Models for lexical knowledge bases," by Branimir Boguraev and Beth Levin
"Providing machine tractable dictionary tools," by Yorick Wilks, Dan Fass, Cheng-Ming Guo, James McDonald, Tony Plate, and Brian Slator