

Briefly Noted

Natural Language Computing: An English Generative Grammar in Prolog

Ray C. Dougherty
(New York University)

Hillsdale, NJ: Lawrence Erlbaum Associates, 1994, xlvii + 349 pp. and 3.5-inch diskette for IBM PC or compatible.
Hardbound, ISBN 0-8058-1525-2, \$69.95.
Paperbound, ISBN 0-8058-1526-0, \$36.00

A notorious problem with computational linguistics is that there is no path into it. The usual CL course presumes a strong background in both linguistics and computer programming—an unlikely combination, especially for students who came to linguistics from the study of languages and literature.

This textbook presupposes no background in either subject. It can be used to give a taste of Prolog programming and natural language parsing to students with no computer experience. Prolog is introduced through simple logic puzzles, and then applied to bibliographic databases, morphological paradigms, and simple parsers. The author takes care to warn students about typing errors, endless loops, and other computer pitfalls that they might not be primed to recognize. An accompanying diskette includes a freeware Prolog interpreter for the IBM PC, together with program source code usable in any Prolog implementation.

The disadvantage of this slow, gentle approach is that even by the end of the book, neither Prolog nor parsing has been covered in depth. Despite the book's title, only a few grammar rules are presented. Basic parsing techniques such as DCGs and difference lists are not introduced; instead, a typical grammar rule is:

$vp(X) :- append(A,B,X), v(A), np(B).$

Feature structures or their equivalent are not used, nor is any mechanism for semantic interpretation indicated. Nor is it clear how the *Syntactic Structures*-style phrase-structure rules relate to the ultramodern Chomskyan minimalism that the author expounds.

Thus, the student who completes this book will still need a Prolog course as well as a linguistics course before doing serious NLP. Nonetheless, Dougherty has performed a useful service by making CL accessible to even the most technophobic of undergrad-

uates. He is using the book himself at New York University, and it evidently works well there. It could even be used with talented high school students.—*Michael A. Covington, University of Georgia*

Structures and Categories for the Representation of Meaning

Timothy C. Potts
(University of Leeds)

Cambridge University Press, 1994, xv + 308 pp.
Hardbound, ISBN 0-521-42481-5, \$54.95

This work addresses the *representation* problem . . . how best to represent the *meanings* of linguistic expressions . . . One might call this the 'semantic form' of expressions (p. xi, italics in the original).

The book begins with "three chapters in which I survey the contributions made by linguistics, logic and computer science respectively to the representation of meaning" (p. xii). These three chapters are not easy to understand, principally because of Potts's obtuse style, an example of which is that instead of saying "'either *P* or *Q*' is false if '*P*' and '*Q*' are both false; otherwise, it is true," he says, "we lay down that a proposition having the structure represented by 'either *P* or *Q*' is to be accounted false if a false proposition is substituted for '*P*' and a false proposition for '*Q*', but is otherwise to be accounted true" (p. 53). These chapters are also outdated. In particular, the chapter on computer science, discussing the work of researchers whose goals are the closest to Potts's own stated goals, is mainly a review of work as of the seventies. There are citations to several of the papers in Findler (1979), but only three to more recent research publications: Hayes (1980), Sowa (1984), and Hobbs and Shieber (1987). Perhaps the most valuable aspect of these three chapters is Potts's criticisms of some of the work he surveys. Of course, some of the problems noted have been corrected in literature that Potts hasn't yet got around to reading.

By the end of the three survey chapters, Potts has introduced two techniques that he

then develops into his own representation—categorical grammars and graphs as representation formalisms. He takes the categorial analysis to be the prior of the two, with his graphs, which he calls *categorial graphs*, being the clearer representation of sentence meaning. Unfortunately, “formalism” and “clearer” must be taken with a grain of salt. Potts never formally defines his categorial graphs, let alone gives a formal semantics for them. Although I have had extensive experience reading, interpreting, and devising graphical representations of meaning, I could not understand the details of Potts’s graphs. But then, neither, apparently, can he: “The relationship between semantic and syntactic structures has not been spelled out, so that it is not fully determinate what our semantic representations represent at the syntactic level” (p. 168).

The four substantive chapters are useful for the linguistic issues that they address, even if they are not useful for the representation scheme that they develop. These issues, which must eventually be faced by all knowledge representation formalisms that aspire to complete coverage of natural language include: quantifier scope; pronouns; relative clauses; count nouns, substance nouns, and proper names; generic propositions; deictic terms; plurals; identity; and adverbs.

Appropriately, the book does not end on a note of claimed accomplishment, but on a note of work yet to do: “The purpose of a philosophical book is to stimulate thought, not to put it to rest with solutions to every problem . . . It is still premature to formulate a graph grammar for semantic representation of everyday language . . . The representation problem is commonly not accorded the respect which it deserves” (p. 288). Many people agree, and have, accordingly, produced a vast literature that Potts is apparently not familiar with. (Some relevant collections are Cercone and McCalla 1987, Sowa 1991, and Lehmann 1992.) Nevertheless, Potts is still correct when he suggests that there is much work left to do.—Stuart C. Shapiro, *State University of New York at Buffalo*

References

Cercone, Nick and McCalla, Gordon (editors) (1987). *The Knowledge Frontier: Essays in the Representation of Knowledge*. Springer-Verlag.

Findler, Nicholas V. (editor) (1979).

Associative Networks: The Representation and Use of Knowledge in Computers. Academic Press.

Hayes, Patrick J. (1980). “The logic of frames.” In *Frame Conceptions and Text Understanding*, edited by Dieter Metzger, 46–61. de Gruyter, 1980. Also in *Readings in Knowledge Representation*, edited by Ronald J. Brachman and Hector J. Levesque, 287–295. Morgan Kaufmann. 1985.

Hobbs, Jerry R., and Shieber, Stuart M. (1987). “An algorithm for generating quantifier scopings.” *Computational Linguistics*, 13(1–2), 47–63.

Lehmann, Fritz (editor) (1992). *Semantic Networks in Artificial Intelligence*. Pergamon Press.

Sowa, John F. (1984). *Conceptual Structures*. Addison-Wesley.

Sowa, John F. (editor) (1991). *Principles of Semantic Networks: Explorations in the Representation of Knowledge*. Morgan Kaufmann.

Dictionnaire encyclopédique de pragmatique

Jacques Moeschler and Anne Reboul (Université de Genève and CNRS Centre de Recherche en Informatique de Nancy)

Editions du Seuil, 1994, 579 pp.
Hardbound, ISBN 2-02-013042-4, FF 380

In 1972, Ducrot and Todorov produced the *Dictionnaire encyclopédique des sciences du langage* with the same publisher, but the present work on pragmatics is thicker and discusses its topics at greater length, as in a course book. It is organized thematically, with a good index and a glossary of some 80 items. There are many discussions of domain definition; in particular, it is steadfastly argued that pragmatics and linguistics do not overlap (despite a nice slip in presupposition in a heading on page 500 referring to “Pragmatics and the rest of linguistics” [“La pragmatique et le reste de la linguistique”]). I regretted the almost complete absence of references or discussion on semantics (except for Fauconnier), semiotics, social psychology and cognitive psychology (except for the classical grounding in sociolinguistics), literary studies, and recent cultural critical theories. Likewise, power, gender, and cross-cultural differences

(Wierzbicka 1991) are not mentioned. If the view of pragmatics is to exclude the study of language-in-use from the reach of linguistics, then these companion areas should be at least considered in a book of this sort.

The eighteen chapters cover a variety of topics: truth-conditionality, context and reference, mental spaces, logical and nonlogical connectors, discourse laws and conversational postulates, presupposition, implicature, argumentation, Ducrot's theory of enunciation, deixis, vague concepts, metaphor, narration, coherence, and discourse analysis and conversation analysis (Moeschler's favorite topic: 1985, 1989). The work mentions many European authors, while strangely apologizing for mentioning many Americans. Some references to the latter deserve updating (e.g., Putnam has changed, and Fodor has explored new directions). Moeschler and Reboul are close to Ducrot on enunciation, and especially to Sperber and Wilson for their theory of relevance (1986), an approach on which they have just published a book (Moeschler, Reboul, Luscher, and Jayez 1994). Grice, Fodor, and Searle are extensively discussed.

I resented the absence of 'computational pragmatics'—if it can be called that—as in human-machine communication. One would have expected a discussion of the contributions of, say, Bunt (e.g., 1994), Cohen (e.g., Cohen, Morgan, and Pollack 1990), Hovy (e.g., 1993), McKeown (e.g., 1985), Wilks and Ballim (1991), and many others. Instead the book argues that "technicians of informatics have been designing their own linguistics, which amounts to a technique applied to language rather than a science of language" ["les techniciens de l'informatique [ont élaboré] leur propre linguistique, qui se résume à une technique appliquée au langage plus qu'à une science du langage" (p. 497)]. The fact that models, when original, are implemented, certainly does not detract from their quality; quite the contrary.

In sum, the book has to be acknowledged as a comprehensive one on "core" pragmatics, and a well-written and well-argued initiator, navigator, and reference work. It compares favorably with Levinson's (1983) textbook.—*Jean-François Delannoy, University of Ottawa*

References

- Bunt, Harry (1994). "Context and dialogue control." *Think*, May 1994, 19–31.
- Cohen, Philip R.; Morgan, Jerry; and Pollack, Martha (editors) (1990). *Intentions in Communication*. The MIT Press.
- Ducrot, Oswald, and Todorov, Tzvetan (1972). *Dictionnaire encyclopédique des sciences du langage*. Éditions du Seuil. Available in English as: *Encyclopedic Dictionary of the Sciences of Language*. Johns Hopkins University Press, 1979.
- Hovy, Eduard (1993). "Automated discourse generation using discourse structure relations." *Artificial Intelligence*, 63(1–2), 341–385.
- Levinson, Stephen C. (1983). *Pragmatics*. Cambridge University Press.
- McKeown, Kathleen R. (1985). *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge University Press.
- Moeschler, Jacques (1985). *Argumentation et conversation: Éléments pour une analyse pragmatique du discours*. Hatier.
- Moeschler, Jacques (1989). *Modélisation du dialogue: Représentation de l'inférence argumentative*. Hermès.
- Moeschler, Jacques; Reboul, Anne; Luscher, Jean-Marc; and Jayez, Jacques (1994). *Langage et pertinence: Référence temporelle, anaphore, connecteurs et métaphore*. Presses Universitaires de Nancy.
- Sperber, Dan and Wilson, Deirdre (1986). *Relevance: Communication and Cognition*. Blackwell.
- Wierzbicka, Anna (1991). *Cross-Cultural Pragmatics: The Semantics of Human Interaction*. Mouton de Gruyter.
- Wilks, Yorick, and Ballim, Afzal (1991). *Artificial Believers: The Ascription of Belief*. Lawrence Erlbaum Associates.

Coherence, Continuity, and Cohesion: Theoretical Foundations for Document Design

Kim Sydow Campbell

(Air Force Institute of Technology)

Hillsdale, NJ: Lawrence Erlbaum Associates (LEA's Communication Series, edited by Jennings Bryant and Dolf Zillmann), 1995, ix + 113 pp.

Hardbound, ISBN 0-8058-1301-2, \$29.95.
Paperbound, ISBN 0-8058-1703-4, \$17.95

Campbell's goal is to explain how we sense

coherence, or "contextual unity," in a text, and how cohesive (textual) devices help put it there in the first place. She considers not only the coherence of the wordstream of a text but also the effects that layout and document design have on the perception of textual coherence. Her basic point is that sensing textual coherence is a matter of gestalt perception, much the same as our perception of unity and structure in visual and auditory stimuli. For example, the perception of relevance, with bridging inferences where necessary, is like gestalt vision, where we fill in the missing pieces. In the case of coherence, the knowledge that is used includes linguistic, pragmatic, social, and idiosyncratic world knowledge.

According to Campbell, the cohesive elements by which the producer of a text establishes coherence and continuity in a text are governed by two main principles: *similarity* of structure and elements, and *proximity*. Campbell shows how these relate to the gestalt perception of structure, through meta-principles of *reinforcement* and *conflict*. For example, similarity can create a background against which dissimilar things can be foregrounded. Thus, discontinuity is a crucial factor in indicating structure by creating a transition.

Campbell addresses her book to technical writers and communication theorists, not computational linguists. But as research in natural language generation becomes more interested in the automatic production of whole documents, not just sentences or short paragraphs, NLG systems are starting to become technical writers, and their builders must therefore become communication theorists. In such systems, the problem of maintaining cohesion and coherence in the text produced becomes much larger than merely ensuring that coreference links are felicitous. Campbell's book is valuable for its provision of a clear way of looking at the problem, a vocabulary for discussing it, and intriguing hints as to how the research should proceed.—*Graeme Hirst, University of Toronto*

The CHILDES Project: Tools for Analyzing Talk, Second Edition

Brian MacWhinney
(Carnegie Mellon University)

Hillsdale, NJ: Lawrence Erlbaum Associates, 1995, xii + 458 pp.
Hardbound, ISBN 0-8058-2027-2, \$35.00

The Child Language Data Exchange System, CHILDES, was established in 1984 by Brian MacWhinney, Catherine Snow, and colleagues as a means for sharing transcripts of child language among researchers in developmental psycholinguistics. It stands as one of the early successes in the construction of computerized corpora and tools for the study of language, and it continues to evolve. The present book serves as an introduction and user's guide to the system and its many components.

Prime Time Freeware for AI, Issue 1-1

Mark Kantrowitz (editor)
(Carnegie Mellon University)

Sunnyvale, CA: Prime Time Freeware, 1994, 220 pp. and two CD-ROMs.
ISBN 1-881957-11-X, \$60.00

Prime Time Freeware is a large package of free AI software chosen from the AI Repository at Carnegie Mellon University. It covers all areas of AI, AI programming languages, and utilities. The computational linguistics section includes parsers, corpora, some specialized bibliographies, and code from some textbooks. The entire package amounts to five gigabytes, compressed onto two CD-ROMs, along with a book that explains how to use the package and lists and indexes all the software in the package. The publisher plans to issue a new edition annually.

Computerized Grammars for Analysis and Machine Translation

Bengt Sigurd (editor)
(University of Lund)

Lund: Lund University Press (Travaux de l'Institut de Linguistique de Lund 29), 1994, 148 pp.
Also distributed by Chartwell-Bratt, Bromley, Kent, U.K.
Paperbound, ISBN 91-7966-304-4 (Lund) and 0-86238-383-8 (Chartwell-Bratt), \$30.00

This book presents some grammars developed within the Swedish machine translation

project SWETRA. The grammars are original contributions but are based on different features of grammatical theories, above all Diderichsen's Field Grammar, TG, GPSG, and GB. The grammars are implemented in Prolog (DCG), and the rules and their functions are explained in great detail. Printouts show how the grammars function in the analysis, generation, and translation of sentences. The basic languages treated are Swedish and English, but Japanese, Russian, Indonesian, and Mapudungu are also touched upon."—*From the publisher's announcement*

The contents of the volume are as follows:

- "Swetra Referent Grammar," by Bengt Sigurd
- "Generalized Word Order Grammar (GWOG)," by Bengt Sigurd
- "Prolog Implementations of English and Swedish GB-Grammars," by Bengt Sigurd and Mats Eeg-Olofsson
- "Appending X-Bar Grammar (AXG) for a Fragment of Swedish," by Bengt Sigurd
- "Appending X-Bar Grammar (AXG) for a Fragment of Japanese," by Birgitta Lastow
- "Nexus Grammar (NEXG) for Swedish and English," by Bengt Sigurd and Barbara Gawrońska
- "Predicate Grammar (PredG)," by Barbara Gawrońska