

here, but I am dubious of the effectiveness of trying to incorporate a crash course in linguistic fundamentals into a course in CL. To be fair, I also think that logic at a level that Allen presupposes (and that linguists do not automatically acquire as a part of their training) ought to be required, too, plus some formal language and automata theory.

At a level of fine detail, there are a number of glitches of various kinds, ranging from solecisms like confusion of term and referent (e.g.: "For example, the NP *the fish* is ambiguous between being a single fish and being many fish" (p.11)) to badly chosen examples. As an instance of the latter I offer the sentence *It's impossible for you to have seen the moon in our car, because it wouldn't fit*, given on p.256 as a case of an inherently anomalous NP (*the moon in our car*) losing its anomaly in context. I would have thought that the sentence in question shows how world knowledge would enable one to know that the sequence *the moon in our car* was not an NP in this context, or at least not one in which the prepositional phrase functions as a restrictive modifier of the head (in which case the anomaly is not there to begin with). I would add, less as a criticism than as a suggestion for amplification in a later edition (which may well come about), that the discussion on non-intersective adjectives on p. 291f take into account the extensive treatment of the semantics of such items to be found in Keenan and Faltz (1985). Finally, I should mention that numerous pages of the copy I received had large black smudges in the margins.

The foregoing reservations aside, I think that this is a good effort at a systematic and comprehensive introduction to the field of CL.

#### REFERENCES

- Grishman, R. 1986 *Computational Linguistics*. Cambridge University Press.
- Harris, M.D. 1985 *Introduction to Natural Language Processing*. Reston Publishing Company, Reston, VA.
- Heidorn, G.; Jensen, K.; Miller, L.A.; and Chodorow, M.S. 1982 The EPISTLE Text-Critiquing System. *IBM Systems Journal* 21:305-326.
- Keenan, E. and Faltz, L.M. 1985 *Boolean Semantics for Natural Language*. Reidel, Dordrecht, Netherlands.
- Pullum, G.K. 1984 If it's Tuesday, This Must be Glossematics. *Natural Language and Linguistic Theory* 2:151-156.
- Tennant, Harry 1981 *Natural Language Processing*. Petrocelli Books, New York, NY and Princeton, NJ.
- Winograd, Terry 1983 *Language as a Cognitive Process, Vol.1: Syntax*. Addison-Wesley, Reading, MA.

#### NOTES

1. I do, however, disagree with the following remark: "It is important to realize that you do not store features and make agreement checks just so that you can detect illegal sentences. Making judgements on grammar is not a goal of computational linguistics" p. 10. Systems like CRITIQUE, né EPISTLE (Heidorn et al. 1982) are intended to do precisely that (among other things).
2. I have a stylistic quibble with Allen's reference on p. 178 to Mitch [sic] Marcus; full names should be used in formal exposition.

3. I use quotes here since, strictly speaking, *production* would be a better term. The word *generate* in its technical sense really means nothing more than "define", and its use as a synonym for *produce* merely perpetuates ancient conceptual confusions.

*Michael Kac* is professor of linguistics at the University of Minnesota, where he teaches an introductory course in computational linguistics. His current research interests include parsing and applications of formal language theory in linguistics. His address is: Department of Linguistics, 142 Klaeber Court, University of Minnesota, Minneapolis, MN 55455. E-mail: kac@umn-cs.arpa.

---

#### A NATURAL LANGUAGE INTERFACE FOR COMPUTER-AIDED DESIGN

Tariq Samad  
(Carnegie-Mellon University)

(The Kluwer international series in engineering and computer science; natural language processing and machine translation)

Boston: Kluwer Academic Publishers, 1986, ix+188 pp.

ISBN 0-89838-222-X; \$38.95 (hb)

Reviewed by  
Bruce Ballard  
AT&T Bell Laboratories

This short book describes an implemented natural language system called *Cleopatra*, which, in its 7,000 lines of Franz Lisp, seeks to provide an interface for use in a CAD system. Since the existing application involves circuit design, the work is reminiscent of the SOPHIE system (Brown and Burton 1975). An example sentence that illustrates the level of sophistication toward which *Cleopatra* strives is, "What is the voltage at n13 after the voltage at n1 reaches 2 v and before the voltage at n23 reaches 3 v?" Most of the book discusses the system's treatment of syntax, in accordance with the finding that "the pragmatics and the semantics required for as limited a sub-domain as ours are very small subsets of the pragmatics and semantics required generally".

*Cleopatra* is predicated on the belief that "the design of a natural language interface for CAD cannot be based on prior assumptions about which English constructions will be useful" and the further suggestion that "the complexity and diversity of the CAD domain . . . render previous approaches to natural language interfaces of little practical utility for our purposes." From the evidence supplied in the book, however, I remain to be convinced that the subset of the CAD domain addressed to date is in fact more complex or more diverse than what has been studied and reported on in the literature. Neither do I believe the techniques presented, though interesting in their own right, render the chosen domain more approachable than a thoughtful attempt at using existing techniques would have.

To provide some details, Cleopatra's parsing is "lexically driven" in that "each word in the input sentence invokes procedures that direct the parsing process". Samad observes that "we can invoke any arbitrary function when evaluating a constraint". Similarly, the integration procedures that build meaning structures may also involve "arbitrary Lisp functions".

Cleopatra attacks ambiguity by use of "confidence levels" related to such phenomena as "the relative frequencies of occurrence of different senses of a word, the likelihood of particular structures, and the correspondence of conjuncts". Parsing is viewed as a parallel process, but the author clearly and honestly states that the existing program is breadth-first.

While the author claims to have achieved a "sharp contrast to the limited domains of previous natural language interfaces", I cannot concur. Although he does handle certain issues in greater detail than have most previous systems (for example, time references and some types of conjunction), there are also, as he so often and honestly admits, many typical distinctions that have been largely ignored (apparently, without suffering a great loss). For example, the system "is not very sophisticated about auxiliaries or adverbs" and does not yet "know about" person or number.

Samad suggests that the "final determinant" of a system with the practical aspirations of his is the evaluation it receives from its intended users. But his claim that "Cleopatra is more than a vehicle to demonstrate the feasibility of our ultimate goal. It is a useful CAD tool" is simply not substantiated. Although "we are confident that experimental studies will confirm the utility of Cleopatra", there's no indication that such investigations have been conducted.

In conclusion, the author is to be commended for his interest in building a complete system that can be used for some meaningful purpose. His frank and honest discussion of the details and limitations of his work are also to be praised, and the presence of some inherently interesting example sentences in the domain under study should be mentioned. However, I find little in the book that helps clarify any problems of language processing, nor do I suspect the techniques presented can provide any "value added" over what's available from existing literature. I am also disappointed by the fact that most of the book concerns implementational issues discussed at the level of data structures. Although I cannot recommend the book as a text, it could certainly be found useful as a case study.

#### REFERENCE

Brown, J. S. and Burton, R. R. 1975 Multiple Representations of Knowledge for Tutorial Reasoning. In Bobrow, Daniel G. and Collins, Allan (eds.). *Representation and Understanding*. Academic Press, New York, NY: 311-349.

*Bruce Ballard* is a Member of Technical Staff at AT&T Bell Laboratories in Murray Hill, NJ. He holds a Ph.D. from Duke University and has been working in the field of computational

linguistics since 1977. He has been on the faculty at Ohio State University and Duke University and currently teaches natural language processing at Rutgers University. Ballard's address is 3C-440A, AT&T Bell Labs, 600 Mountain Avenue, Murray Hill, N.J. 07974 USA. E-mail: allegra!bwb.att.com

#### THE FORMAL COMPLEXITY OF NATURAL LANGUAGE

Walter J. Savitch; Emmon Bach; William Marsh; and Gila Safran-Naveh (eds.)

(University of California, San Diego; University of Massachusetts, Amherst; Xerox Palo Alto Research Center; and University of Cincinnati)

(Studies in linguistics and philosophy 33)

Dordrecht: D. Reidel, 1987, xviii+451 pp.

ISBN 1-55608-046-8, \$69.00, Dfl 145.00, £ 44.95 (hb)

*Reviewed by*

*Alexis Manaster-Ramer*

*IBM T.J. Watson Research Center*

This book anthologizes a number of papers dealing with mathematical models of, and mathematical claims about, human languages. The collection begins with a stage-setting paper by Stanley Peters, "What is mathematical linguistics?" and gives the last word to Gerald Gazdar and Geoffrey K. Pullum in their "Computationally relevant properties of natural languages and their grammars". The papers in between are grouped into three sets:

*Early nontransformational grammar:*

Janet Dean Fodor, "Formal linguistics and formal logic".

Emmon Bach and William Marsh, "An elementary proof of the Peters-Ritchie theorem".

Thomas Wasow, "On constraining the class of transformational languages".

Gilbert H. Harman, "Generative grammars without transformation rules: A defense of phrase structure".

P. T. Geach, "A program for syntax".

*Modern context-free-like models:*

Geoffrey K. Pullum and Gerald Gazdar, "Natural languages and context-free languages".

Gerald Gazdar, "Unbounded dependency and coordinate structure".

Hans Uszkoreit and Stanley Peters, "On some formal properties of metarules".

Emmon Bach, "Some generalizations of categorial grammars".

*More than context-free and less than transformational grammar:*

Joan Bresnan et al., "Cross-serial dependencies in Dutch".

Stuart M. Shieber, "Evidence against the context-freeness of natural language".

James Higginbotham, "English is not a context-free language".

Christopher Culy, "The complexity of the vocabulary of Bambara".