

Computational lexical semantics

Patrick Saint-Dizier and Evelyne Viegas (editors)

(Institut de Recherche en Informatique de Toulouse, CNRS and Brandeis University)

Cambridge University Press (Studies in natural language processing, edited by Branimir K. Boguraev), 1995, ix + 447 pp.

Hardbound, ISBN 0-521-44410-1, \$69.95

Reviewed by

Paul Deane

Dataware Technologies

1. Overview

The last decade has seen a striking expansion of interest in the lexicon and in lexical semantics. Within theoretical linguistics, this trend can be measured by the increasing interest generative grammarians have displayed toward such issues as lexical conceptual structure and argument structure and by the increased appeal of cognitive semantics. Within computational linguistics, the same period has seen a burgeoning of interest in the construction of semantically realistic lexicons and their integration with larger Natural Language Processing (NLP) systems. *Computational lexical semantics* supplies a fascinating snapshot of the state-of-the-art in the early 1990s. The articles date from late 1991 or early 1992, reflecting the immediate impact of Pustejovsky's theory of the generative lexicon (1991), but lacking references to George Miller's WordNet system (1990). The book is organized into six sections:

- I. Psycholinguistics for lexical semantics;
- II. Foundational issues in lexical semantics;
- III. Lexical databases;
- IV. Lexical semantics and artificial intelligence;
- V. Applications;
- VI. Computer models for lexical semantics.

The book provides wide coverage. A variety of problems are addressed, though the focus is upon polysemy, disambiguation, and co-composition—that is, processes by which word meanings are dynamically derived and contextually modulated. There are gaps in coverage, however, which appear to reflect the early state of work in the field. Polysemy generally falls into one of three categories: metonymy (associations between concepts in the same domain), metaphor (mappings across conceptual domains), and various patterns of semantic variation that are often lumped together and labeled as prototype effects. But this book (and, it would appear, computational approaches generally) concentrates almost exclusively on explicating metonymy (which

is, in fact, the least problematic of the three and thus the easiest to formalize). Similarly, the articles in the book are strongest when dealing with logical aspects of meaning. Very little attention is devoted to dealing with less easily formalized aspects of lexical knowledge, such as the role of attention and frequency of use in disambiguation processes. These, of course, are not criticisms of the book, but a reflection of its focus on problems that seem tractable using well-understood tools and techniques.

2. Contents of the book

The articles may be divided under headings somewhat different from those used by the editors. While the editors organized the information by field and task, the actual research is heavily weighted toward a few key issues, in particular polysemy and lexical disambiguation.

General articles. 'An introduction to lexical semantics from a linguistic and a psycholinguistic perspective', by Patrick Saint-Dizier and Evelyne Viegas; 'Polysemy and related phenomena from a cognitive linguistic viewpoint', by D. A. Cruse; 'Mental lexicon and machine lexicon: Which properties are shared by machine and mental word representations? Which are not?' by Jean-François Le Ny.

Saint-Dizier and Viegas provide a brief review of lexical semantics covering classic concepts in several key frameworks, including Jackendoff's lexical conceptual structure, Pustejovsky's generative lexicon, and Mel'čuk's Explanatory Combinatory Dictionary (1988). Cruse focuses on providing a linguistic explication of the complexities that make polysemy difficult to handle, including its partial productivity, lack of clear boundaries, and context-sensitivity. Le Ny provides a useful checklist of the properties of the mental lexicon that are and are not paralleled in current NLP lexicons. Critically, Le Ny notes, such basic psycholinguistic properties as activation (and the related properties of activability and pre-activation) are not built into most NLP models of the lexicon. The concerns raised by Le Ny and Cruse are very important, particularly since they do not seem yet to have been fully addressed in current NLP work.

Polysemy in general. 'Word meaning between lexical and conceptual structure', by Peter Gerstl; 'Lexical semantics and terminological knowledge representation', by Gerrit Burkert; 'A preliminary lexical and conceptual analysis of *break*: A computational perspective', by Martha Palmer and Alain Polguère. Of closely related interest: 'Inheriting polysemy', by Adam Kilgarriff.

Computational lexical semantics implies the construction of very large lexical databases that go beyond traditional NLP lexicons (or most AI knowledge bases for that matter) by supporting highly flexible, dynamic interpretative processing. Gerstl's article explores what information is required to support dynamic lexical interpretation. After a detailed literature review, he argues programmatically for an analysis in which word meanings are composed of a set of interacting 'factors', some of which constrain and others of which expand the potential for interpretation. Burkert's (equally programmatic) article focuses on aspects of lexical meaning amenable to traditional knowledge representation formalisms (i.e., term subsumption languages). By contrast, Palmer and Polguère's article is extremely data-oriented. They focus upon the word *break*, arguing that it is best analyzed as a hierarchy of sense components, each of which entails specific constraints on the word's overall syntactic and semantic structure. Adam Kilgarriff's article is extremely interesting because it shows that at least some types of variations in word meaning can be stored in a semantic net and inherited. It should be noted, however, that Kilgarriff's article focuses on metonymy:

other types of polysemy are left largely unaddressed. This emphasis on accounting for metonymy is typical of the current state of NLP work on polysemy, as may be seen in the articles discussed below.

Metonymy, the generative lexicon and related issues. ‘Linguistic constraints on type coercion’, by James Pustejovsky; ‘From lexical semantics to text analysis’, by Sabine Bergler; ‘Lexical functions, generative lexicons, and the world’, by Dirk Heylen; ‘Semantic features in a generic lexicon’, by Gabriel G. Bès and Alain LeComte. A closely related application of generative lexicon theory: ‘The representation of group denoting nouns in a lexical knowledge base’, by Ann Copestake. Of related interest: ‘A lexical semantic solution to the divergence problem in machine translation’, by Bonnie J. Dorr.

Much of the semantic variability of natural language is due to the influence of logical metonymy—where word meanings shift (following conceptually natural associations) during the process of semantic composition. Pustejovsky’s theory of the generative lexicon integrates a linguistically motivated view of metonymy with a rigorous formal semantics through the mechanism of type coercion to make the lexicon into a generative mechanism capable of deriving contextually appropriate meanings ‘on the fly’. Pustejovsky’s article outlines the basic theory; Sabine Bergler discusses its applications to text comprehension. Dirk Heylen draws parallels between Pustejovsky’s generative lexicon and Mel’čuk’s Explanatory Combinatory Dictionary, pointing out close parallels between the two systems. Bès and LeComte seek to define a metalanguage for describing proposals about word meaning, proposing mechanisms parallel to Pustejovsky’s, though purposely more generic.

Copestake’s article is quite interesting, as it has a strong theoretical base but addresses many of the practical problems involved in constructing a large lexical semantic database. Copestake’s system implements a version of Pustejovsky’s generative lexicon in a unification-based lexical knowledge representation. She develops a type-coercion analysis of the polysemy and related syntactic properties of group nouns, and illustrates how the relevant lexical entries can be automatically acquired from a machine-readable dictionary. But, as might be expected, there are significant difficulties, including problems in defining default inheritance relationships and the absence of fully consistent cues for group noun membership in the underlying dictionary data.

Dorr’s article focuses on machine translation but shares with the articles listed above a deep concern for providing an adequate linguistic underpinning for NLP work. She argues that use of Jackendoff’s lexical conceptual structures provides a useful solution to the problem of divergence in machine translation. The argument is simple: since lexical conceptual structures (*a*) are closely tied to syntactic structure, and (*b*) are nonetheless deep semantic representations, they are well suited to provide an interlingua. Independently needed, language-specific mapping mechanisms can then be exploited to account for different divergent syntactic expressions of the common interlingua.

Theoretically motivated work using Jackendoff’s conceptual semantics and/or Pustejovsky’s generative lexicon forms the cutting edge of current NLP work. The articles listed above provide a good picture of early advances along this line of research. It may be too early, however, to judge how far the generative lexicon approach can be taken, until large-scale lexicons based upon its principles have been constructed. But there can be no doubt that these approaches mark a significant advance over earlier computational models of the lexicon.

Disambiguation, defaults, and logical approaches to NLP. ‘Large neural networks for the resolution of lexical ambiguity’, by Jean Veronis and Nancy Ide; ‘Blocking’, by Ted Briscoe,

Ann Copestake, and Alex Lascarides; 'A non-monotonic approach to lexical semantics' by Daniel Kayser and Hocine Abir. Of related interest: 'Introducing Lexlog', by Jacques Jayez; 'Constraint propagation techniques for lexical semantics descriptions', by Patrick Saint-Dizier.

Word-sense disambiguation is a fundamental and potentially intractable problem. Veronis and Ide argue that it can be accomplished by a neural network without prior linguistic analysis. They construct a network in which words are associated with the words in their (machine-readable dictionary) definition. In the data sets upon which they report, they achieve excellent disambiguation, but their method would seem to have severe limitations: it is not at all clear that it would scale up well, both for computational reasons (the size of the required neural network) and theoretical reasons (the potential for interference from crosstalk among definitions as the number of words in the network increases.)

Default logic is an obvious alternative to connectionist techniques of lexical disambiguation. Briscoe et al.'s article deals with the important issue of lexical blocking—the prevention of one form or meaning from occurring because a competing element already exists. In principle, the resolution principles that Briscoe et al. employ apply equally well to formal and semantic ambiguity. Kayser and Abir argue that disambiguation can be formally modeled by using default logic to set preferences for one meaning over another. The articles by Jayez and Saint-Dizier concern implementations of logic programming-based NLP systems; while not directly concerned with default logic, many of the same technical issues arise, such as the propagation of inherited information and the resolution of multiple constraints in a complex knowledge base.

These articles address one of the thorniest problem types in lexical semantics. However, default logic has an apparent weakness when applied to language: its failure to account for the effects of analogy, habituation, and other essentially cognitive factors. For example, Kayser and Abir are forced to postulate a 'strength' factor (in effect, an impressionistic measure of relative psychological dominance, or capacity to attract attention in a neutral context) to induce the default logic to choose a single (likely) word meaning from the choices that remain after clearly inappropriate meanings have been eliminated. It remains to be seen whether an account can be developed that maintains the strengths of a formal logic while incorporating realistic theories of dominance, attention, and other extra-logical aspects of human cognition.

Structural analysis. 'Lexical semantics: Dictionary or encyclopedia', by Marc Cavazza and Pierre Zweigenbaum; 'Lexical functions of the Explanatory Combinatorial Dictionary for lexicalization in text generation', by Margarita Alonso Ramos, Agnès Tutin, and Guy LaPalme.

Cavazza and Zweigenbaum describe a classically structuralist technique for analyzing word meaning into contrasting semantic components. Since they apply the technique to medical texts, where the key words tend to be highly terminologized, the results are quite good. Keeping a strict separation between domain prototypes and word definitions, they focus on demonstrating how an NLP system can infer prototypes from partial, phrasal descriptions. Ramos et al. describe the use of Mel'čuk's lexical functions to generate anaphor matches, paraphrases, and the like in a speech generation system.

3. Conclusions

One of the striking facts about this book is the way in which it illustrates a general trend: the growing recognition (throughout linguistics and its allied fields) of the im-

portance of the lexicon. *Computational lexical semantics* provides a valuable overview of NLP work instantiating this trend. The articles are generally of very high quality, and provide a fairly balanced view of NLP approaches, with approximately equal representation of research originating in North America and Europe.

References

- Miller, George A.; Beckwith, Richard; Fellbaum, Christiane; Gross, Derek; and Miller, Katherine (1990). "Introduction to WordNet: An on-line lexical database." *International Journal of Lexicography*, 3(4), 235–244.
- Mel'čuk, Igor and Zholkovsky, Alexander (1988). "The Explanatory Combinatorial Dictionary." In Evens, Martha W. (editor), *Relational Models of the Lexicon*, 41–74. Cambridge: Cambridge University Press.
- Pustejovsky, James (1991). "The generative lexicon." *Computational Linguistics*, 17(4), 409–441.

Paul Deane wrote his Ph.D. thesis at Chicago on *Semantic Theory and the Problem of Polysemy* (1987). His research focuses on lexical semantics, polysemy, and metaphor, with articles appearing in such journals as *Lingua*, *Cognitive Linguistics*, *Metaphor and Symbolic Activity*, and *Journal of Pragmatics*. In addition, he has published a monograph on cognitive syntax, *Grammar in mind and brain* (Mouton de Gruyter, 1993). He is currently employed as linguist and lexical semanticist for a project in cognitive modalities in information retrieval. Deane's address is: Dataware Technologies, One Antares Drive, Suite 200, Nepean, Ontario, Canada K2E 8C4; e-mail: PDeane@dataware.com.