Predicative Forms in Natural Language and in Lexical Knowledge Bases

Patrick Saint-Dizier (editor)
(IRIT-CNRS, Toulouse)

Dordrecht: Kluwer Academic
Publishers (Text, speech and language
technology, edited by Nancy Ide and
Jean Véronis, volume 6), 1999,
viii+377 pp; hardbound, ISBN
0-7923-5499-0, $132.00, £77.00,
Dfl 220.00

Reviewed by
Manfred Stede
Technical University of Berlin

When (computational) linguists speak of "lexical semantics," quite often they are actually referring to the more narrow area of verb semantics. Similarly, this book chooses the fairly abstract title "predicative forms" to join 12 contributions, most of which are concerned with verbs, their meaning, and the relationships between meaning and syntax. The book is the follow-up publication of a workshop held under the same title in Toulouse in August 1996. Taken together, the collection provides a good overview of contemporary approaches to lexical (verb) semantics, and thanks to a 50-page introductory chapter written by the editor, it is also useful reading for graduate students who are considering immersing themselves in lexico-semantic issues. This introduction offers brief explanations of some central themes and theories that are referred to in the later chapters (argument structure, thematic roles, lexical relations, polysemy, verb alternations and corresponding verb classes, WordNet, lexical-conceptual structure, and the generative lexicon), always with references to further reading.

The additional introductory chapter is one of three elements that distinguish the book from plain workshop proceedings. The second element is an index, but it is unfortunately a rather sparse one. The third add-on is a chapter that summarizes the contributions to a panel session held at the workshop—see the end of this review. In contrast to these bonuses, one mildly annoying feature of many workshop proceedings can be found in this book too: Some parts of it display an unfortunate abundance of typographical and other errors.

Let us begin our glance at the individual papers with the two that are mainly concerned with words other than verbs. One is by Federica Busa, who works in the framework of the generative lexicon (Pustejovsky 1995) and proposes a classification of agentive nominals along with representations that can explain some interesting aspects of their syntactic behavior. The other study deals with connectives: Jacques Jayez and Corinne Rossari use generalized quantifier theory to offer a classification of some connectives that involve inferences in recognizing the relation holding between the conjuncts.

The most application-oriented chapter is by Evelyne Viegas et al. The authors first describe the architecture of the Mikrokosmos parser and the format of the lexical representations used therein, and then discuss the practical issues of semiautomatically

acquiring the knowledge resources for such a large system, focusing in particular on how to balance the information between language-neutral ontology and language-specific lexicons.

Among the papers devoted to verbs, three are centered on the notion of argument structure. The most "theoretical" paper of the book is by Charles Jones, who examines the borderline between morphology and syntax in the Government-and-Binding framework, and argues that the rules operating in the two realms are of different kinds. Approaching the subject from an applied perspective, Dimitrios Kokkinakis is interested in automatically acquiring the argument structure of Swedish verbs from large corpora. He starts with a handbuilt valency database and then uses a tagger, lemmatizer, and chunk parser to validate this information in the corpus. Working with a comprehensive set of French verbs, Patrick Saint-Dizier builds a catalogue of 200 possible verbal "contexts" (generalized argument structures) and then forms verb classes according to which configurations a verb can appear in. The resulting 953 classes reflect the alternation behavior of verbs and are thus similar in nature to those proposed by Levin (1993). But Saint-Dizier points out the difference that Levin's classes are primarily based on shared semantic features, and thus exceptions in the alternation behavior within a class are allowed; his own classification is, by contrast, solely motivated by shared syntactic behavior. Accordingly, he arrives at a large number of classes, most of which are very small.

From a different perspective, Levin's verb classes are compared to the classes of WordNet (Miller and Fellbaum 1991) in the paper by Christiane Fellbaum, who first describes the treatment of verbs in WordNet and then argues that the alternation behavior of verbs is in part reflected by the WordNet classification, as some nonterminal WordNet classes roughly correlate with some alternations. This is remarkable, because according to Fellbaum the WordNet classes were meant as purely semantic; syntax did not play a role in setting them up. While Saint-Dizier in his work on alternations deliberately avoids semantic issues, Palmer et al. present a detailed study of the meaning of verbs involving motion and paths, as well as some of their alternations. By adding appropriate semantic features to a grammar in the formalism of feature-based LTAGs, the authors argue that their representations can be used in transfer-based MT to handle certain divergences that are commonly seen to favor using interlingual systems.

Similar to Palmer et al., María Ángeles Zarco uses lexical-conceptual structure (LCS) (Jackendoff 1990) to represent the meaning of verbs. Here, LCS serves as an interlingua for Spanish, Italian, and English support-verb constructions and their corresponding lexicalized forms. And LCS is also the semantic framework of choice for Achim Stein, who analyses polysemy in Italian verbs and suggests a type hierarchy along with LCS-based representations. The idea is to shift the polysemy to the level of the LCS primitives, which can thus assume multiple readings.

As this overview shows, the range of topics and theoretical frameworks underlying the various chapters of the book is quite broad. Only a subset of the papers have clear interconnections or provide some interestingly opposing views of a common theme—which is the syntax–semantics interface of verbs. The book covers the major current (especially computationally inspired) approaches to this topic, and this survey function I would consider its main strength. One chapter in particular is devoted to contrasting divergent conceptions: Summarizing their contributions to a panel session held at the workshop, five authors (Busa, Dubois, Fellbaum, Saint-Dizier, and Viegas) present their respective views on representing the transfer verbs *buy* and *sell*, often with comments on the other contributions, which serves to partially re-create the flavor of a panel.

References

Jackendoff, Ray. 1990. *Semantic Structures*. The MIT Press, Cambridge, MA.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. The University of Chicago Press.

Miller, George and Christiane Fellbaum. 1991. Semantic networks of English. *Cognition*, 41, 197–229.

Pustejovsky, James. 1995. *The Generative Lexicon*. The MIT Press, Cambridge, MA.

Manfred Stede is a researcher in the Department of Computer Science at TU Berlin. One of his main research interests is *Lexical semantics and knowledge representation in multilingual text generation*, and he has published a book under this title in 1999. At present, he works in the context evaluation group of the German Verbmobil speech-to-speech MT project. Stede's address is: TU Berlin, FB Informatik, Sekr. FR 6-10, Franklinstr. 28/29, 10587 Berlin, Germany; e-mail: stede@cs.tu-berlin.de.