

Book Reviews

Ontological Semantics

Sergei Nirenburg and Victor Raskin

(University of Maryland, Baltimore County, and Purdue University)

Cambridge, MA: The MIT Press, 2004,

xxi+420 pp; hardbound, ISBN

0-262-14086-1, \$50.00, £32.95

Reviewed by

John F. Sowa

In this book, Nirenburg and Raskin present an important body of work in computational linguistics that they and their colleagues have been developing over the past 20 years. For a unifying perspective, they organize their assumptions, theories, and techniques around the theme of **ontological semantics**. Along the way, they critique many alternative views of semantics, which they distinguish from their own. Their analyses contribute to a much-needed debate about the history and future of computational linguistics, but to preserve some balance, teachers and students should keep a few of the alternatives on their reference shelf.

The book is divided into two parts: a philosophical part I and a practical part II. The first part consists of an introductory chapter 1 and four chapters that survey important but controversial issues about linguistics, both theoretical and computational. In those chapters, the authors make a good case for their version of ontological semantics, but the alternatives are not treated in detail. In part II, the authors present their text-meaning representation (TMR) and demonstrate how it is used in language analysis. Any discussion of technical material must use some notation, and TMR is sufficiently flexible to illustrate a wide range of semantic-based methods that could be adapted to many other formalisms. For most readers, part II would be the more important.

Chapter 1 is a good 25-page overview of computational linguistics with an emphasis on semantics. Students and novices, however, need examples, and none are given until chapter 6. The authors suggest that “a well-prepared and/or uninterested reader” skip the remainder of part I and go straight to chapter 6, which begins with an excellent five-page example. The authors follow that advice when they teach courses from this text.

In Chapter 2, the authors present their “Prolegomena to the Philosophy of Linguistics.” Their ideas are well taken, and some are as old as Socrates: Examine the assumptions, challenge conventional wisdom, and test conclusions against experience. The basis of their approach is what they call the four components of a scientific theory:

1. The *purview* of a theory is “the set of phenomena for which the theory holds itself—and is held—accountable.”
2. The *premises* are belief statements “taken for granted by the theory and not addressed in its body.”

3. The *body* of a theory “is a set of its statements, variously referred to as laws, propositions, regularities, theorems, or rules.”
4. *Justification* “is the component of a theory that deals with considerations about the quality of descriptions and about the choices a theory makes in its premises, purview, and body.”

Under various names and with varying definitions, similar components are present in most theories about theories. The authors’ claims of novelty in proposing them “surprisingly, for the first time in the philosophy of science” are overstated.

An important point that the authors fail to mention is the purpose of a project and the nature of the subject matter: Theories in mathematics, engineering, and the empirical sciences are very different in kind and methods of justification. Since computational linguistics is primarily an engineering discipline, it uses theories from mathematics and the sciences, and it helps test and develop them. But the primary justification for an engineering project is the ability to solve a problem within the limits of budgets and deadlines. The authors spend too much time arguing against engineering goals that are different from their own. For some applications, such as machine translation, an analysis of truth conditions may be unnecessary. For other applications, such as translating an English question into a database query, truth conditions are the focus of the task. Instead of recognizing that different engineers have different goals, they have tried to banish truth conditions from linguistics. The following passage indicates a serious misunderstanding:

—*First, we maintain that reference is relevant for the study of coreference and anaphora . . . relations in text. Second, while we agree that truth plays no role in the speaker’s processing of meaning, we are also aware of the need to “anchor” language in extralinguistic reality. Formal semanticists use truth values for this purpose. We believe that this task requires a tool with much more content, and that an ontology can and should serve as such a tool.* (page 109)

First, logicians do not use truth values to anchor language in reality; they use references, which are resolved to entities (objects, properties, and events) in some situation. Second, truth values are not primary, but derived from the mapping of linguistic references to actual entities; a sentence is true if and only if the linguistic configuration of references and relations conforms to the extralinguistic configuration of entities. Third, every logician from Aristotle to the present has insisted that an ontology of every general term is essential to determine the correct mapping from language to reality. Aristotle himself never used the word *ontology*, even though he created the subject; logicians are more likely to use the words *theory*, *axiomatization*, and *conceptualization* as synonyms for what Nirenburg and Raskin call an ontology.

Chapter 3 is a brief, 11-page history of semantics, but it is distorted by the fact that the word *semantics* was not coined until the end of the 19th century. The subject matter, however, was established by Aristotle in the books *Categories*, *On Interpretation*, *Analytics*, *Rhetoric*, and *Poetics*. Under the name of logic or theory of signs, the subject was thoroughly developed by the Hellenistic and medieval philosophers. Most books on logic before the 20th century devoted at least half their

text to conceptual analysis and ontology. The truncated view of history ignores 2,000 years of research:

- Frege is credited with the distinction between extension and intension, but those words are Hamilton's translation of the 17th century *étendue* and *compréhension*, which were just new names for a distinction that had been analyzed in detail by Aristotle and the medieval scholastics.
- The citations for the "dawn of metalanguage" are to the 1950s. But Aristotle's theory of definition in terms of *genus* and *differentiae* is metalanguage, and so is his theory of syllogisms for analyzing the components of meaning. The scholastics introduced the terms *first intentions* for language about physical objects (e.g., *Homo est animal*) and *second intentions* for language about language (e.g., *Homo est species*).
- Katz and Fodor are given well-deserved credit for being the first in the Chomskyan school to integrate a componential analysis of word meaning with a compositional analysis of sentence meaning. Ockham (1323), however, combined both componential and compositional analyses to determine the truth conditions for Latin sentences. He not only anticipated Frege's compositionality and Tarski's model theory, he went beyond them by applying the techniques to a natural language instead of an artificial one.
- Kamp did not "add to the agenda of formal semantics a treatment of coreference and anaphora." The "donkey sentences" that Kamp analyzed were English translations of examples used by the scholastics for analyzing similar phenomena.
- The modern history is just as flawed. The authors claim that first-order predicate calculus (FOPC) has failed "to have made a historical impact." In fact, FOPC is the foundation for the SQL databases that run the world economy. That is certainly an impact.

A major omission is the early semantic work in AI and MT. One of the pioneers in ontological semantics for MT was Margaret Masterman (1961), a former student of Wittgenstein's. She organized her ontology as a lattice defined in terms of 100 primitive concepts, which Wilks adopted as a basis for preference semantics. Hutchins (1986) showed that her MT system did a better job of word selection than purely syntactic systems of that time. Appropriately, her first publication on the subject was in the *Proceedings of the Aristotelian Society*. Another pioneer was Silvio Ceccato (1961), who based his **correlational nets** on a selection of 56 relation types, which included case relations, type-subtype, type-instance, part-whole, and miscellaneous logical, numerical, causal, spatial, and temporal relations. In parsing, Ceccato built dependency trees, which he "correlated" with predefined nets to resolve ambiguities; in generation, he used the nets to guide word selection. The single most influential collection of the early work in these two fields, edited by Minsky (1968), included classic papers by McCarthy, Quillian, Bobrow, and Raphael, among others.

Chapter 4 summarizes the goals and issues of lexical semantics with numerous citations of authors who contributed to the field. Unfortunately, it has very few

examples comparing the ways different authors would analyze similar phenomena. Nirenburg and Raskin cite seven authors in the Russian meaning–text school but don’t give a single example to show how a meaning–text analysis would differ from their own text-meaning analysis. Throughout the chapter, they discuss Pustejovsky’s generative lexicon but never illustrate the arguments with examples. They consider Pustejovsky “as a representationalist, antiformalist ally,” but they never explain why they consider lexical semantics incompatible with formal semantics. That is especially odd, since the next chapter positions “ontological semantics within the field of formal ontology.”

Chapter 5 is a survey of formal ontology, an ancient subject that has become the latest hope for conferring interoperability on incompatible systems. Most of that work, however, has not been adapted to natural language processing. Work on lexical resources, such as WordNet, is only loosely connected to work on formal ontology. Section 5.3 discusses “the difficult and underexplored part of formal ontology, namely, the relations between ontology and natural language.” The most difficult problem, which the proponents of formal ontology fail to address, is the nature of ambiguities in natural languages. A good parser can enumerate syntactic ambiguities, and selectional constraints are usually sufficient to resolve most of them. The most serious ambiguities are subtle variations in word senses (sometimes called **microsenses**), which change over time with variations in word usage or in the subject matter to which the words are applied. Such variations inevitably occur among independently developed systems and Web sites, and attempts to legislate a single definition will not stop the growth and shift of meaning. From their long experience with NL processing, Nirenburg and Raskin probably have a deeper understanding of the nature of ambiguity than the proponents of the Semantic Web. Section 5.4 is a wish list of features from formal ontology that NL processors would need. Providing them is still a major research problem.

After all the preliminaries, chapter 6 introduces text-meaning representation with examples that illustrate the mapping from English to TMR. Section 6.1 begins with the sample sentence *Dresser Industries said it expects that major capital expenditure for expansion of U.S. manufacturing capacity will reduce imports from Japan*. The next five pages carry out an informal analysis of that sentence without introducing any special notation, not even TMR. Then section 6.2 introduces TMR and shows how the results of the analysis in section 6.1 are mapped into it. The remaining sections of chapter 6 discuss the fine points of using TMR and compare them to other computational and theoretical techniques.

TMR is essentially a network of frames, each of which has a head and a list of binary relations that link the head to a frame, a pointer to another frame, a simple value, or a more complex combination for defaults, semantic types, relaxable types, etc. Each TMR is a set of six kinds of frames: one or more **propositions**, zero or more **discourse relations**, zero or more **modalities**, one **style**, zero or more **references**, and one **TMR time**. The kinds of frames are illustrated with numerous examples discussed throughout chapters 6, 7, 8, and 9. Unfortunately, there is no appendix or other reference section that gives a complete grammar or table of all the options for a well-formed TMR. From the examples, one can surmise that the head of each proposition frame is a **concept instance** that represents a state or event, which is linked by case roles to the participants. In the middle of chapter 7 is a table of nine case roles; at the end of chapter 8 is a list of five types of discourse relations, each of which may have several subtypes. The authors acknowledge that TMR has been evolving over the years, but a complete list of options for one version would be appreciated.

Chapter 7 presents the four **static knowledge sources**: ontology, fact database, lexicon, and onomasticon. The subdivision into four sections is uneven: The fact database is described in three pages, and the onomasticon in half a page, but the ontology and lexicon sections take 36 pages and 15 pages, respectively. The discussion of inheritance (a description logic with defaults) should be in a separate section, and some material belongs in an appendix: the table of case roles, the list of 34 axioms that define constraints on the Mikrokosmos ontology, and the description of the software for browsing the knowledge sources. The question of what information to put in the onomasticon, the lexicon, or the ontology raises some troublesome issues: *Toyota*, for example, is in the onomasticon because it is the name of an instance of type corporation, but *Toyota Corolla* is in the ontology because it is a type of car, which can have many instances.

Chapter 8, which at 62 pages is the longest in the book, shows how TMR is used in text analysis. Section 8.1 presents the stages of tokenization, morphology, lexical lookup, and syntactic analysis. Section 8.2 covers the construction of dependency structures for propositions, which includes matching selectional restrictions and relaxing them for sentences such as *The gorilla makes tools*. Sections 8.3 and 8.4 cover problems of ambiguity, nonliteral language, and the inevitable exceptions. Section 8.5 treats time, aspect, and modality. Section 8.6 handles discourse: reference and coreference, discourse relations, and the temporal ordering of the propositions. This is a good chapter, but one might like to see some discussion of alternative methods of parsing and semantic interpretation. It would also be interesting to see a step-by-step processing of the sample sentence that was analyzed by hand in section 6.1.

Chapter 9 addresses knowledge acquisition: the problem of constructing the four knowledge sources discussed in chapter 7. This is a universal problem that everybody involved with NL processing has to face, and nobody working in the field is completely satisfied with the available resources. In this chapter, the authors focus on the methods they have used in developing the Mikrokosmos ontology and associated lexicon, but they discuss issues involved in adopting and adapting resources such as WordNet and machine-readable dictionaries. They try to take an ideal scientific stance toward the subject, but most readers are likely to adopt a mixed strategy of adapting whatever resources they are given or are likely to find on the Internet. As a fact database, many readers are likely to be given, in advance, a conventional relational database, and the authors should discuss the issues of incorporating such resources.

Chapter 10 is a three-page conclusion in which the authors apologize for the lack of detail on applications and processing. Earlier in the book they say that the kinds of applications for which TMR has been used “include machine translation, information extraction (IE), question answering (QA), general human-computer dialog systems, text summarization, and specialized applications combining some or all of the above.” A couple of more chapters on language generation and reasoning would have been more useful than most of the five chapters of part I. For students, a glossary would be especially welcome, since the authors frequently mention a word such as *defeasible* and follow it with a parenthetical list of citations instead of a definition.

An embarrassing lapse shows that even people who design semantic processors are forced to use less sophisticated tools for routine chores. The index contains five references to I. J. Good, who had not been working on computational linguistics, but only one reference leads to Good’s publications and four lead to capitalized occurrences of the word *good*. This lapse is even more embarrassing for Good

(1965), who predicted "It is more probable than not that, within the twentieth century, an ultraintelligent machine will be built and that it will be the last invention that man need make."

Despite the historical and philosophical inaccuracies, this is a valuable textbook on computational linguistics. Its greatest strength is its engineering contribution, and its greatest weakness is the constant bickering with linguists and logicians who study different aspects of the rich and complex subject of language. Humans and machines require both logical and lexical processing for language understanding, and the authors could better inform students by showing what their approach does best than by trying to limit the range of topics linguists are allowed to explore.

References

- Ceccato, Silvio. 1961. *Linguistic Analysis and Programming for Mechanical Translation*. Gordon and Breach, New York.
- Good, Irving John. 1965. Speculations concerning the first ultraintelligent machine. In F. L. Alt and M. Rubinfoff, eds., *Advances in Computers*, volume 6. Academic Press, New York, 31–88.
- Hutchins, W. John. 1986. *Machine Translation: Past, Present, Future*. Ellis Horwood, Chichester, England. Available at <http://ourworld.compuserve.com/homepages/WJHutchins/PPF-TOC.htm>.
- Masterman, Margaret. 1961. Translation. *Proceedings of the Aristotelian Society*, 61:169–216.
- Minsky, Marvin, ed. 1968. *Semantic Information Processing*. MIT Press, Cambridge, MA.
- Ockham, William of. 1323. *Summa Logicae*. Part I translated as *Ockham's Theory of Terms* by M. J. Loux, University of Notre Dame Press, Notre Dame, IN, 1974. Part II translated as *Ockham's Theory of Propositions* by A. J. Freddoso and H. Schuurman, University of Notre Dame Press, Notre Dame, IN, 1980.

John F. Sowa worked for 30 years on research and development projects at IBM, and he is a cofounder of a startup company, VivoMind Intelligence, Inc. He has published and edited several books and numerous articles on knowledge representation, computational linguistics, and related areas of artificial intelligence. He is a Fellow of the AAAI and is best known for his work on the theory of conceptual graphs and their application to natural language semantics. Sowa's e-mail address is sowa@bestweb.net.