Book Reviews

Semantic Relations Between Nominals

Vivi Nastase[1], Preslav Nakov[2], Diarmuid Ó Séaghdha[3], and Stan Szpakowicz[4]
([1]FBK, Trento; [2]QCRI, Qatar Foundation; [3]University of Cambridge; [4]University of Ottawa)

Reviewed by
Lucy Vanderwende
Microsoft Research

Understanding noun compounds is the challenge that drew me to study computational linguistics. Think about how just two words, side by side, evoke a whole story: *cacao seeds* evokes the tree on which the cacao seeds grow, and to understand *cacao powder* we need to also imagine the seeds of the cacao tree that are crushed to powder. What conjures up these concepts of *tree* and *grow*, and *seeds* and *crush*, which are not explicitly present in the written word but are essential for our complete understanding of the compounds? The mechanisms by which we make sense of noun compounds can illuminate how we understand language more generally. And because the human mind is so wily as to provide interpretations even when we do not ask it to, I have always found it useful to study these phenomena of language on the computer, because the computer surely does not (yet) have the type of knowledge that must be brought to bear on the problem. If you find these phenomena equally intriguing and puzzling, then you will find this book by Nastase, Nakov, Ó Séaghdga, and Szpakowicz a wonderful summary of past research efforts and a good introduction to the current methods for analyzing semantic relations.

To be clear, this book is not only about noun compounds, but explores all types of relations that can hold between what is expressed linguistically as nominal. Such nominals include entities (e.g., *Godiva, Belgium*) as well as nominals that refer to events (*cultivation, roasting*) and nominals with complex structure (*delicious milk chocolate*).[1] In doing so, describing the different semantic relations between *chocolate in the 20th century* and *chocolate in Belgium* is within the scope of this book. This is a wise choice as there are then some linguistic cues that will help define and narrow the types of semantic relations (e.g., the prepositions above). Noun compounds are degenerate in the sense that there are few if any overt linguistic cues as to the semantic relations between the nominals.

The book has three main chapters: an introduction to the various relation schemas that have been used in the past to describe the nominal relations, an overview of methods for extracting semantic relations with supervision, and an overview of methods for extracting semantic relations with little or no supervision. The preface promises a

---

1 Examples are drawn from the Introduction of *Semantic Relations Between Nominals*. The examples with reference to cacao and chocolate make for a sweet introduction.

very brief chapter summing up the lessons learned and makes good on that promise as Chapter 5 is very brief indeed.

It was a pleasure to be reminded of de Saussure's characterization of the two foundational types of relations: syntagmatic relations, which hold between words present in the text, and paradigmatic relations, the associations of the text words to the broader context (i.e., to words not in the text). The syntagmatic relations are then introduced to us as the predicates in logic, taking one or more nominal arguments, or as the labeled arcs that connect concepts, making the notion of semantic relations accessible to people with differing backgrounds in computer science. What follows is "a menagerie of relation schemata," Section 2.2, where approximately eight different relation sets are inventoried, an admirably comprehensive overview of the past decades of research in this area.[2] Whereas the authors acknowledge that from the NLP perspective, "[the aim] is to select the most useful representation for a particular application" (page 12), no examples are provided that demonstrate the impact of selecting one representation over another. Thus, although the overview is of great value to a reader wishing to familiarize themselves with the topic, the advanced reader will be left with the impression that the authors do not know how to select from among the relation sets either.

When data labeled with semantic relations is available, supervised models can be used. This scenario is explored in detail in Chapter 3. An overview of the data resources available is followed by a brief discussion of the types of features often employed in learning. This, and the next chapter, are clearly intended for readers with some familiarity with machine learning who may not be as familiar with the full sweep of models available. In particular, the organization of the material provides good entry points where the reader can find plentiful references if they are curious about the approach. The summary sections are very good reviews and would suffice to get a general idea of how the task of extracting semantic relations can be approached. In particular, in Section 3.5, the authors provide useful points to consider for navigating the myriad possibilities of models and resources.

When no labeled data is available, a wider set of paradigms for extracting semantic relations needs to be explored (in Chapter 4). These range from manually authored patterns (Hearst 1992, inter alia) for extracting predetermined relation types, to identifying novel relations with open relation extraction. OpenIE learns relations as expressed by verbs or prepositions using sophisticated algorithms to determine when the verbs or prepositions express the same or different relation type (Lin and Pantel 2001, inter alia). These paradigms are well described in this chapter and, once again, the authors provide plentiful information that invites the reader to go into greater depth wherever their interests lie, with the odd exception that the authors steer the reader away from any approach that uses parsing, claiming that "deeper processing (e.g., syntactic parsing) is altogether infeasible on a Web scale" (page 79, passim). We know that Google is parsing the Web already (Petrov and McDonald 2012) and parsing is becoming orders-of-magnitude faster (Canny et al., 2013), so the reader would be well advised to stay open to the possibilities. Parsing was demonstrated to be crucial for extracting higher-order semantic relations such as Location and Purpose (Montemagni and Vanderwende 1992).[3]

---

2 Readers may also wish to know about the project NomBank (Meyers et al., 2004), which is not included in this chapter.

3 Coincidentally, this paper was presented in the same session as Marti Hearst's paper at COLING 1992; note the page numbers in the References. At the time, however, very few institutions had access to a broad-coverage parser, but I suspect that parsing and features derived from parsing will soon become commonplace.

In the end, the authors note that "relation extraction is not an end task, but its purpose is to build resources to be used by other NLP and AI applications" (page 80). And the application is what will determine which set of relations is appropriate, which will in turn determine the best approach for extracting such semantic relations, whether through supervised or unsupervised means. This book provides a wonderfully comprehensive overview of the choices that a practitioner makes today. I wish there was a last chapter cataloguing NLP and AI applications that use the semantic relations, but perhaps I can look forward to that in the next edition.

References

Canny, J., D. Hall, and D. Klein. 2013. A multi-teraflop constituency parser using GPUs. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1,898–1,907, Seattle, WA.

Hearst, Marti 1992. Automatic acquisition of hyponyms from large text corpora. In *Proceedings of the 15th International Conference on Computational Linguistics (COLING)*, pages 539–545, Nantes.

Lin, Dekang and Patrick Pantel. 2001. Discovery of inference rules for question-answering. *Natural Language Engineering*, 7(4):343–360.

Meyers, A., R. Reeves, C. Macleod, R. Szekely, V. Zielinska, B. Young, and R. Grishman. 2004. The NomBank Project: An interim report. In *Proceedings of the HLT-NAACL 2004 Workshop: Frontiers in Corpus Annotation*.

Montemagni, Simonetta and Lucy Vanderwende. 1992. Structural patterns vs. string patterns for extracting semantic information from dictionaries. In *Proceedings of the Fourteenth International Conference on Computational Linguistics (COLING)*, pages 546–552, Nantes.

Petrov, Slav and Ryan McDonald. 2012. Overview of the 2012 Shared Task on Parsing the Web. In *Notes of the First Workshop on Syntactic Analysis of Non-Canonical Language (SANCL)*, Montreal.

Lucy Vanderwende is a Senior Researcher at Microsoft Corporation. Her thesis is titled "The Analysis of Noun Sequences using Semantic Information extracted from On-Line Dictionaries" (1996, Georgetown University). The research was continued in the MindNet project (Richardson, Dolan and Vanderwende 1998). She continues to be fascinated by semantic relations and works in the area of information extraction from biomedical text and on probabilistic frame-induction (Cheung, Poon and Vanderwende 2013). Vanderwende's address is: One Microsoft Way, Redmond, WA 98052 USA; e-mail: `lucyv@microsoft.com`.