Sergei Nirenburg and Victor Raskin

Ontological Semantics

The MIT Press 2004
Cambridge, Massachusetts, London, England
ISBN 0-262-14086-1
x + 420 pp.

Reviewed by
Petr Němec

The book under review presents one particular approach to natural language understanding with emphasis on automatic semantic analysis. "Ontological Semantics" is in fact a name for the "ideal form" of the system the authors propose together with the related theories and assumptions.

The monograph is divided into two major parts: a theoretical Part I and a practical Part II. Being more a computational linguist than a formal semanticist I would like to focus my attention mainly on the second part describing the layout of a complete natural language processing system.

Part I may be viewed as a (polemic) introduction to computational linguistics, semantics, formal ontology, partly even to the theory of science. The practical motivation of the authors is apparent in their argumention against the efforts to base the entire natural language understanding system on a single formalism which, being suitable for one task, makes it unnecessarily hard or impossible to deal with another. Instead, they argue for a concept of "microtheories" ? separable and replaceable modules corresponding to the respective tasks of natural language analysis (morphological and syntactic analysis, coreference, temporality, word sense desambiguation etc.) and communicating with each other in a well-defined manner.

Part II actually describes such a system. First, the central notion of "text meaning representation" (TMR) is introduced. TMR is basically a set of propositions connected through text-level discourse relations. Additionally, there is an arbitrary number of modalities and references together with a single TMR time specification (contains the time specifications for the respective propositions) and a single style specification. These basic categories contain graph structures of complex values and pointers. In spite of the complexity of the concrete attributes this approach is in general very intuitive ? the relevant information is captured as a graph centered around the respective propositions augmenting any information necessary with attribute values and binary relations.

The system relies heavily on the use of external knowledge, which is divided into four components: the ontology, the fact repository, the lexicon, and the onomasticon. The ontology represents a concept (type) hierarchy. The actual instances of a type (e.g. London of type City) are present in the fact repository. The language expressions (e.g. the word "London") representing a proper name are listed in the onomasticon and linked to the entities of the fact repository. The lexicon contains all other language material and is linked to the ontology. A separate chapter describes ways of aquisition of this material.

With these components at hand the authors describe at various levels of detail the functionality of the subsystems corresponding to the respective microtheories: disambiguation of valency frames via matching of selectional restrictions, restriction relaxation in certain contexts, ellipsis resolution, modality, event phase, iteration, temporality and establishment of coreference relations. Discourse processing at the suprapropositional level is also discussed.

It is important to realize that the described system and the presented examples are more a vision of an ideal system rather than a description of any actual implementation (CAMBIO/CREST, Microcosmos).

Unfortunately, the authors do not make this distinction explicit (sometimes they even remark on the slightest details of a particular implementation) thus tacitly leaving the reader with the impression that the presented - very difficult - examples are handled by some existing system systematically, robustly and with high coverage. No present existing system would be able to match such performance.

In summary, I recommend this book mainly because it provides a promising view of the design of a realistic natural language understanding system. It also contains some insightful information on the respective microtheories.