The Logical Structure of English: Computing Semantic Content

Allan Ramsay
(University College Dublin)

London: Pitman, 1990, iv + 209 pp.
Paperbound ISBN 0-273-03287-9, £19.99

Reviewed by
Martin Volk
University of Koblenz-Landau

How do you judge whether to read a book? You look at the front and back cover, the name of the author, the table of contents, the bibliography, and some formal aspects, such as layout. If you have the time, you read some pages to get a taste. Doing all this with Allan Ramsay's book will give you a good basis for your decision. From the back cover you will learn that "recent work in formal semantics has provided analyses of a number of isolated phenomena. The present work shows what it is like to integrate these fragments into a single unified treatment. The book presents a set of syntactic-semantic rules which build up a description of the logical content of English sentences on the basis of the logical contents of their components. These rules, which are the heart of the book, give an account of a substantial fragment of English."

Allan Ramsay is a professor of Artificial Intelligence at University College Dublin and in the past has done work in parsing (Ramsay 1985), formal methods in AI (Ramsay 1988), and in feature logic (Ramsay 1990).

From the table of contents we learn that the book is organized into six chapters and comes with an additional preface, an epilogue, a bibliography, an appendix listing all the "Rules and Frames" introduced in the book, and an index. The bibliography consists of 63 very diverse entries, including the landmarks in the respective fields: from Koskenniemi on morphology to Searle on speech acts. It also contains a good number of entries on different aspects of logic, which gives you a warning that this book focuses on the formal aspects of language.

The layout of the book is pleasant. The examples are numbered in a clear manner. Syntactic and semantic structure listings are all framed in boxes and thus contrasted from the text. Spelling and syntactical errors are few in number. When browsing through the book, one is immediately struck by the amount of structure displays. On 200 pages Ramsay has 75 Frames and 20 Rule structures and countless more auxiliary structures. This results in bothersome reading, since the notation is not comprehensible on first sight. Especially reading Chapter 4 reminded me of reading sparsely annotated C code rather than narrative text.

So what is in the book? Chapter 1 is an easy-to-read introduction to the task at hand: formalizing a fragment of English. It briefly explains the prerequisites of compositionality and computational representation that underlie the work. Some examples illustrate the difficulties of formalizing natural language in an intuitive way. Chapters 2 through 5 are the heart of the book. They are of about equal length. Chapters 2 and 3 introduce the semantic and the syntactic frameworks used for the formalization. Chapter 4 tries to integrate semantic and syntactic aspects into a whole, whereas Chapter 5 points to the limits of this approach and to an extension using thematic roles. Chapter 6 is a brief tribute to discourse and speech act considerations.

Chapter 2, "Semantic Framework," elaborates the language for formalizing semantic aspects of natural language. The semantic aspects consist of propositions and properties. The language chosen by Ramsay is an extension of first-order predicate logic and is called *property theory*. In particular, he introduces an operator for abstraction or, as he calls it, objectification. Property theory is supplied with a "Game-theoretic semantics." "The basic objects that underlie our semantics are going to be not sets of sentences, not sets of possible worlds, but games. ... It is a game in which you argue with yourself about some proposition. You win if you show that you can defend the proposition in question against all possible attacks" (p. 31). This is a less known approach to semantics, which, according to Ramsay, dates back to Lorenz (1961) and Lorenzen (1959). He chose this approach because it incorporates a treatment of knowledge and belief that is supposedly superior to possible worlds semantics. I found this choice interesting and thought-provoking, but I could not find out how this choice influenced the formalization later in the book. Chapter 2 ends with discussions about correctness (soundness and completeness) of the proposed language, as well as some details of how to apply this language to the phenomena of mutual knowledge, intensionality, commitments, discourse entities, and temporal relations. The latter two strike me as being treated very briefly and superficially. It is a problem throughout the book that, because of the variety of issues, many have been glossed over.

Chapter 3, "Syntactic Framework," elaborates the syntactic formalism. This formalism is in the tradition of unification-based theories. It borrows most notably from HPSG. The Head Feature Convention and Foot Feature Principle are taken from GPSG. The chapter starts with a formal definition of classification (of words), substitution, unification, and rules. Ramsay then demonstrates how phrasal constituents such as sentences, verb phrases, and noun phrases are represented. The notation is very similar to HPSG, though macros are used to abbreviate frequently needed structures, while the *phon*-feature from HPSG is missing here. Next, words and word classes are discussed, with some detail on how suffixes contribute to the lexical information attached to each word and word class. That is, a lexical lookup consists of unifying the frames for the word class, the word, and the suffix. Finally, rules are introduced. They are illustrated with examples for sentences and for complex noun phrases. In particular, relative clauses are discussed at length, distinguishing five different types: those in which the relative pronoun is the subject and those in which it is an object; those without a relative pronoun; those without a relative pronoun and without the auxiliary *be*; and appositive NPs or VPs that resemble relative clauses.

Chapter 4, "Basic Meaning," tries to bridge the gap between the semantic and the syntactic structures. Towards this goal every (syntactic) structure is annotated with a feature "semantics," which again is divided into presupposition and content. The notion of presupposition is very simplistic: "The presupposition is some proposition which the speaker believes that the hearer already believes" (p. 20). Thus for a sentence such as *Mary sold the cows*, the presuppositions given (p. 144) are (roughly):

- there exists a "salient" person named Mary, and
- there exists a "salient" subset of all cows.

However, we don't learn how one finds out about what is salient: "Characterizing the circumstances in which some item is indeed salient... is outside the scope of the present work" (p. 112). Starting off with the semantic formalization of simple sentences Ramsay then focuses on the formalization of auxiliaries, temporal aspects, complex NPs, and coordination.

Chapter 5, "Roles and Structures," advances the formalism by introducing states and events. This is done by borrowing from Fillmore's notion of thematic roles that distinguishes between surface subjects and semantic subjects. Thematic roles are applied to the phenomena of agentive and agentless verbs (e.g., *kill* vs. *die*) as well as to passives, generics (bare plurals and habituals), sentential and verbal complements, intensional objects, and free modifiers (prepositional and adverbial phrases).

Chapter 6, "Linguistic Action," gives a brief overview of the relevance of taking general behavior into account while formalizing natural language understanding. This is done by showing parallels between AI planning theory and speech act theory to the task at hand. The main idea is that (linguistic) actions consist of preconditions and effects. Ramsay illustrates how this could be incorporated into his formalism by showing how the actions of *informing*, *nagging*, and *declaring* differ with respect to their preconditions and effects. The chapter concludes with a call for increasing the formalization of mutual knowledge in future NLP systems. None of this is actually implemented in Ramsay's system. His formalization stops at the single-sentence level, and leaves contextual and extratextual considerations for further studies.

The above summary shows that this book tackles a wide variety of phenomena. This results in the superficial treatment of most details as, for example, temporal aspects (pp. 51, 127). In addition, when reading phrases such as "This looks as though it is an empirical claim. ... As far as we are concerned it is simply a technical device" (p. 64), or "We simply assume...without further worrying about..." (p. 127), I got the uneasy feeling that pragmatic considerations of getting the job done were more important than scientific rigor.

The main question left unanswered is "Whom is this book written for?" It is certainly not a textbook. It does not contain any exercises. It is also not meant to present new results in the formalization of English. Nor is the implementation meant to be applied in any particular NLP system. The book is simply a report on the formalization of a variety of phenomena of English. As such, it may serve as a repository of methods and results on how to treat these phenomena in a computational framework. Unfortunately the obscure notation prevents a reader from using it as a lookup source for individual phenomena. He or she must work through a substantial part of the book to get used to the representation.

Ramsay offers a disk that contains the implementation of the rules introduced in the book. The program is written in Quintus Prolog, and I had no problems running the demo sentences on my Sun. The output of the program for these sentences is even harder to read than the structure displays in the book. Unfortunately there is neither program documentation nor any manual for the software.

References

Lorenz, K. (1961). *Arithmetik und Logik als Spiele*. Doctoral dissertation, University of Kiel.

Lorenzen, P. (1959). "Ein dialogisches Konstruktivitätskriterium." *Proceedings of the Symposium on the Foundations of Mathematics*, Warsaw.

Ramsay, A. (1990). "Disjunction without tears." *Computational Linguistics*, **16**(3), 171–174.

Ramsay, A. (1988). *Formal Methods in Artificial Intelligence*. Cambridge: Cambridge University Press.

Ramsay, A. (1985). "Effective parsing of GPSG." *Proceedings, 2nd European Conference on Computational Linguistics*, Geneva.

Martin Volk is a teaching and research assistant in the Department of Computational Linguistics at the University of Koblenz-Landau. His background is in Computer Science, Linguistics and

Psychology. He holds a Master's degree in Artificial Intelligence from the University of Georgia. His current research interests are in tools for syntax analysis and in the cognitive aspects of parsing. Volk's address is: Universität Koblenz-Landau, Institut für Computerlinguistik, Rheinau 3-4, 5400 Koblenz, Germany.