Meaning Extraction from a Peircean Perspective

J.J. Sarbo, J.I. Farkas, F. Grootjen,
P. v. Bommel and T. v.d. Weide

University of Nijmegen, Toernooiveld 1
6525 ED Nijmegen, The Netherlands
e-mail: janos@cs.kun.nl

Keywords language analysis, C.S. Peirce, semiotic, meaning extraction.

Meaning extraction is a form of information management. It is closely related to summarising ([Spa93]), since an extract may alternatively be viewed as a substitute for the source document. However an extract is more than a summary in as much as it also reflects the hierarchical structure of the meaning of the input text.

Meaning extraction assumes some representation for the source document. This, in turn, requires an adequate modelling of language, in particular, of its hierarchic (e.g. syntactic) structure. We argue that traditional models of natural language are not satisfactory in this respect, as it is witnessed by the limited success of such models in natural language processing. Traditional language modelling takes as its starting point that hierarchical structure is somehow given. However, this assumption is sometimes too rigid, and cannot fit the high flexibility of language use.

The approach proposed in this paper is based on Peirce's semiotic ([Pei31]) which provides us with a deeper foundation of language. In this approach, which is monostratal, hierarchical structure arises, via the interaction of language symbols, as a result of linguistic semiosis. Because interactions are events, language may be considered a set of symbol-events, a process.

We apply the process view of language to syntactic (and morphologic) structure ([AA82]), and by using English as an example we illustrate the potential of the Peircean framework in parsing ([Sar99], [FS99]).

The Peircean model of syntactic structure provides us with a framework for meaning representation for sentences. In as much as sentences are symbols, the above model can also be applied to texts. It will be argued that the hierarchial structure of the meaning of a text arises as a result of the interaction of sentences, and precisely by adopting the same approach and representation as in the syntactic analysis of sentences.

We exemplify our approach by the meaning extraction of a sample English text. A comparison of the Peircean model of language with the 'generative' one, from a philosophical point of view, can be found in [DFS99].

References

[AA82] F. Aarts and J. Aarts. *English Syntatic Structures.* Pergamon Press, Oxford, 1982.

[DFS99] G.Y. Debrock, J.I. Farkas, and J.J. Sarbo. Syntax from a Peircean Perspective. In *5th International Congress on Terminology and Knowledge Engineering*, Innsbruck (Austria), 1999.

[FS99] J.I. Farkas and J.J. Sarbo. A Peircean Framework of Syntactic Structure. In *Seventh International Conference on Conceptual Structures (ICCS'99)*, Springer-Verlag, 1999.

[Pei31] C.S. Peirce. *Collected Papers of Charles Sanders Peirce.* Harvard University Press, Cambridge, 1931.

[Sar99] J.J. Sarbo. Formal Conceptual Structure in Language. In D. M. Dubois, editor, *AIP Conference Proceedings 465*, Woodbury, New York. 1999.

[Spa93] K. Sparck Jones. What Might Be a Summary? In G. Knorz, J. Krause, and C. Womser-Hacker, editors, *Information Retrieval'93: von der Modellierung zur Anwendung.* Universitatsverlag Konstanz. 1999.