# Perception Based Processing of NL Texts.
## Modeling discourse understanding as visualized learning in SCIP Systems.

**Burghard B. Rieger**
FB II: Computational Linguistics, University of Trier
D-54286 TRIER, Germany
e-mail: rieger@uni-trier.de

**Keywords:** Natural language processing, discourse understanding, text analyzing algorithms, procedural semantics, fuzzy meaning, dynamic representation, semantic vector space, imaging and visualization.

**Extended Abstract**

**1.** It is common practice in theoretical linguistics, formal semantics and cognitive modeling to identify real world entities with the (symbolic) structures that represent them. Some of the problems in logics and linguistics that these models encounter, are due to the (crisp) declarative *formats* of (symbolic, compositional, propositional) *representations* employed, and the (rule-based, modular, deterministic) *procedures* chosen in processing language entities (elements, structures, relations, functions, and processes) whose meanings are specified model dependent via truth conditions. In order to understand how natural languages (NL) serve the purposes they do, it has to be investigated what makes a sign stand for (or symbolize) something else. In doing so, procedural and fuzzy approaches to modeling NL understanding have devised some means to come to grips with the dynamics of cognition as a multi-layered process of structure identification [7] that allows to cope with the variability and vagueness, adaptivity and learning, emergence and plasticity of *knowledge* and *understanding* [8]. Fuzzy modeling techniques allow for (numerical, sub-symbolic, distributed, non-propositional) formats whose (parallel, pattern-based, quantitative) computation result in (the *emergence* of) meanings which are the outcome rather than the presuppositions of processing, and whose modeling is a form of *realization* rather than *simulation*[3].

**2.** *Semiotic Cognitive Information Processing* (SCIP) models [5, 6] are inspired by *information systems theory* and concentrate on (natural or artificial) system-environment situations whose knowledge-based processing of *information* makes them *cognitive*, and whose sign and symbol generation, manipulation, and understanding capabilities render them *semiotic*. SCIP systems' ability comprises their performance in knowledge-based information processing and representing its results, organizing these representations by activating others from prior processing, planning acts by selecting from such organized and represented dispositions, and modifying them according to changing conditions, results, and states of evolving system-environment adaptedness. Based on NL structures, SCIP performance is a form of complex, multi-resolutional information processing tied to (and even be identified with) language *understanding*. Whenever such cognitive processes are modeled as being based upon structures whose representational status is not a presupposition to but a result from an algorithmic processing, then these algorithms – being

able to initiate and modify the structures they are operating on – may qualify as *semiotic* and part of *computational semiotics*.

**3.** The *perception based approach* of SCIP systems to NL text processing for discourse understanding is – like vison [2] – part of an image generating semantics (BIGS for *Bild gebende Semantik*) which complements the symbolic (de)composition of propositional structures in traditional NL semantics. Grounded in system-environment situations, BIGS represents meanings as structured sets of perspectival relations (*dispositional dependencies*) among new entities (*meaning points*) which emerge in multi-layered vector space mappings (*corpus space, semantic space*) from computation of (patterns of *syntagmatic* and *paradigmatic*) combinatorial constraints in (not necessarily NL) material processed.

**3.1** In order to demonstrate the SCIP systems' potential of discourse *understanding*, it is evaluated against the real world *situations* whose *descriptions* are processed. For this purpose a test scenario was chosen, confining the discourse material to (syntactically correct, semantically true) *language* descriptions of *real world* situations (not to symbolic representations of them) on the one hand, and delegating the processing to well defined *formalism* implemented as algorithms of mapping and/or measuring procedures (not to formal or symbolic definitions of abstract relations or functions) on the other hand.

**3.2** The process of *description* is algorithmically controlled by a formal grammar (*syntax* and *semantics*) as provided by computational linguistics. These define a notion of *correctness* and *truth* for the *dynamic generation* of propositional structures which describe changing real world situations in a formally controlled way. Assembled to collections of increasing size, this language material forms a PHT-corpus (of pragmatically homogeneous texts) whose semantic contents are the described situations these texts refer to.

**3.3** The process of *understanding* is controlled by implemented semiotic algorithms for the recursive computation of combinatorial constraints in texts and their multi-layered, multi-resolutional representation in (patterns of) distributions of (observable and emergent) entities. These realize a procedural notion of *semioticity*, formally defined as a system of *morphisms* which allow to specify Peirce's [4] triadic conception of *semiosis* for empirical application in a SCIP setting.

**3.4** As SCIP is defined to work sub-symbolically – without any (presupposed knowledge of) syntax or semantics – on the basis of perceiving (structures of) material language entities in NL discourse, coming up with a pictorial representation of these structures that resembles the real world scenario described by that discourse, is tantamount to the realized constitution of meaning or the understanding of discourse and what it purports to communicate.

**4.** The 2-dim scenario of the *real world* (Fig.1) is a reference plane with two stationary objects (environment), and an oriented mobile SCIP agent (system) which are structurally coupled [9] by a PHT corpus of (true and correct) natural language (NL) expressions[1] of possible system-position/object-location (SPOL) relations. The perception-based, non-symbolic processing of these descriptions for vectorial meaning points' representation in *semantic space* allows its over-all structure be computed – which will be derived – as an image (Fig.2) of regions of potential object locations by profile lines of common likelihood (*isoreferentials*).

**5.** A prototype SCIP implementation as testbed for the *description* and *understanding* processes covering variable system-environment situations is presented to illustrate the performance of a perception based, procedural approach to the dynamics of semiotically

[1] *e.g."Triangle is very far in front, very near left. Square is very near in front, extremely near right.* etc."

grounded (natural language) meaning constitution.

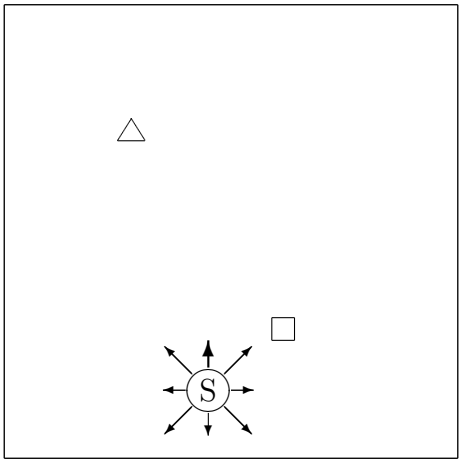


Figure 1: Reference plane with locations of stationary objects △ and □ , and a mobile agent S whose system-positions relative to object-locations determine propositional expressions of SPOL relations.

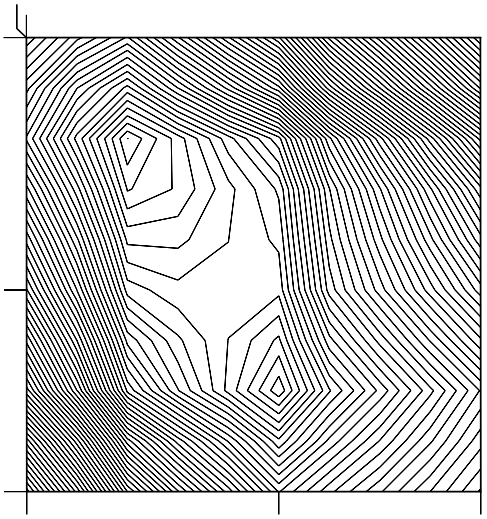

Figure 2: 2-dim image of potential object locations (*isoreferentials*) depicting sub-symbolic *understanding* by perception based, non-propositional processing of NL descriptions of SPOL-relations.

## References


[1] J. Barwise/J. Perry: *Situations and Attitudes*. [Bradford Books], Cambridge, MA (MIT Press), 1983.

[2] D. Marr: *Vision*, SanFrancisco (Freeman), 1982.

[3] H. Pattee: Simulations, Realizations, and Theories of Life. In: Langton, C.G.(ed): *Artificial Life*. [Santa Fe Institute Studies in Complexity], Reading, MA (Addison Wesley), 1989, pp. 63–77.

[4] C. S. Peirce: Pragmatism in Retrospect: A Last Formulation. In: Buchler (ed): *The Philosophical Writings of Peirce*, (CP 5.11–5.13), New York (Dover), 1906, pp. 269–289.

[5] B. B. Rieger: Meaning Acquisition by SCIPS. In: Ayyub, B.M. (ed): *ISUMA-NAFIPS-95*, IEEE-Transactions: Joint Int. Conf. on Uncertainty Modeling and Analysis, Los Alamitos, CA (IEEE Computer Society Press), 1995, pp. 390–395.

[6] B. B. Rieger: Semiotics and Computational Linguistics. On Semiotic Cognitive Information Processing. In: Zadeh, L.A./ Kacprzyk, J.(eds): *Computing with Words in Information/Intelligent Systems*, Heidelberg (Physica), 1999, pp. 93–118.

[7] B. B. Rieger: Computing Granular Word Meanings. A fuzzy linguistic approach in Computational Semiotics. In: Wang, P.P.(ed): *Computing with Words*. [Wiley Intelligent Systems Series 3], New York, NY (Wiley & Sons), 2001, pp. 147–208.

[8] B. B. Rieger: Semiotic Cognitive Information Processing: Learning to Understand Discourse. A systemic model of meaning constitution. In: Stamatescu, I.O. (ed): *Adaptation and Learning*, Berlin/ Heidelberg/ NewYork (Springer), 2002 [in print].

[9] F. Varela/E. Thompson/E. Rosch: *The Embodied Mind. Cognitive Science and Human Experience*, Cambridge, MA (MIT Press), 1991.