Understanding is Meaning Constitution.

Perception-based processing of natural language texts in procedural models of SCIP Systems.

Burghard B. Rieger
Computational Linguistics, University of Trier,
D-54286 TRIER, Germany

Keynote Lecture: 7th *Joint Conf. on Information Science (*JCIS-03), Durham, NC, USA, Sept. 26–30, 2003

1. Introduction

The alliance of logics and linguistics as mediated by (language) philosophy and (discrete) mathematics has long been (and partly still is) dominating the way in what (notational) terms natural languages structures and their functions are to be explicated and how cognitive processes of *understanding* should be modeled. As it is common practice in cognitive modeling and formal semantics[1] to identify real world entities with the structures that represent them, this identification is rather more hiding instead of revealing what enables a sign structure to represent and stand for (or symbolize) something else. Some of the problems such models encounter are due to the declarative *formats* employed (symbolic, compositional, propositional) and the *procedures* chosen (rule-based, modular, deterministic) in depicting and manipulating the entities (elements, structures, relations, functions, and processes) which are to represent the (world, linguistic, situational) knowledge considered conditional for the explicative comprehension of how natural languages serve the purposes they do.

Other than these declarative models of cognitive processes which operate on symbol representations and essentially static knowledge bases, procedural approaches strive to come to grips with the dynamics of cognition as a multi-layered process [8] that allows to cope with the variability and vagueness, adaptivity and learning, emergence and plasticity of *knowledge* and *understanding* [9]. Procedural modeling employs (numerical or sub-symbolic, distributed, non-propositional) formats whose (parallel, pattern-based, quantitative) computation may result in (the *emergence* of) entities which are the outcome rather than presuppositions of processing, and whose modeling is a form of *realization* rather than *simulation* [4].

2. Semiotic Cognitive Information Processing

Modeling *Semiotic Cognitive Information Processing* (SCIP) [6], [7] is inspired by *information systems theory*. It concentrates on (natural and/or artificial) systems' embeddedness in their respective environments (*situatedness*) whose knowledge-based processing of *information* makes them *cognitive*, and whose sign and symbol generation, manipulation, and understanding capabilities render them *semiotic*. SCIP systems' capability to perform cycles of cognitive processes and to represent their results in increments of emerging structures allows to model the dynamics of learning and development. Activation of earlier representational results from prior processing and selection of relevant portions from these dependency structured representations (*dispositions*) which are modified according to changing conditions, relevancies, and states of evolving system-environment adaptedness, is what makes

[1] *Situation theory* [1], [2] excepted.

this form of complex, multi-resolutional information processing be tied to (or even identified with) the faculty of language *understanding*. Whenever cognitive processes are modeled as being based upon structures whose representational status is not a presupposition to but a result from such algorithmic processing, then these algorithms – being able to instantiate and modify the structures they are operating on – may qualify as *semiotic* and part of *computational semiotics*.

3. Dynamic Image Generating Semantics

The *perception based approach* of SCIP systems to discourse understanding is – like vision [3] – part of a *dynamic image generating semantics* (DIGS) which complements the symbolic (de)composition of propositional structures in traditional formal approaches to the semantics of natural language [11], [12]. Grounded in system-environment situations, DIGS 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 natural) language material processed.

3.1 In order to demonstrate the SCIP systems' potential of discourse *understanding*, it can be evaluated against the certainty of formally defined language *descriptions* [13]. For this purpose a particular test scenario was chosen, confining the discourse material to language descriptions of *real world situations* (not to symbolic structures representing them) on the one hand, and the processing to *well defined formalism* with algorithmized and implemented numerical pattern detection, measuring, and/or mapping procedures (not to formal definitions of rule based symbol manipulation functions) on the other.

3.2 For the *description* process an algorithmic language production approach was implemented based on a formally specified *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 For the process of *understanding*, some well defined, semiotic algorithms were implemented for the detection and recursive computation of combinatorial constraints in texts as well as their multi-layered, multi-resolutional representation in (patterns of) distributions of (observable and emergent) entities. In all, they realize a procedural notion of *semioticity*, formally defined as a system of *morphisms* which specify PEIRCE's conception of *semiosis*[2] 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 (patterns of) material language entities in NL discourse, the processing results or states of the system's *semantic space* structure can be visualized. These image representations resemble the over-all real world scenario as described by the natural language texts processed which is tantamount to the realized constitution of meaning or the understanding of discourse and what it purports to communicate.

[2]"By *semiosis* I mean [...] an action, or influence, which is, or involves, a coöperation of *three* subjects, such as *sign*, its *object*, and its *interpretant*, this tri-relative influence not being in any way resolvable into actions between pairs." [5, p.282]

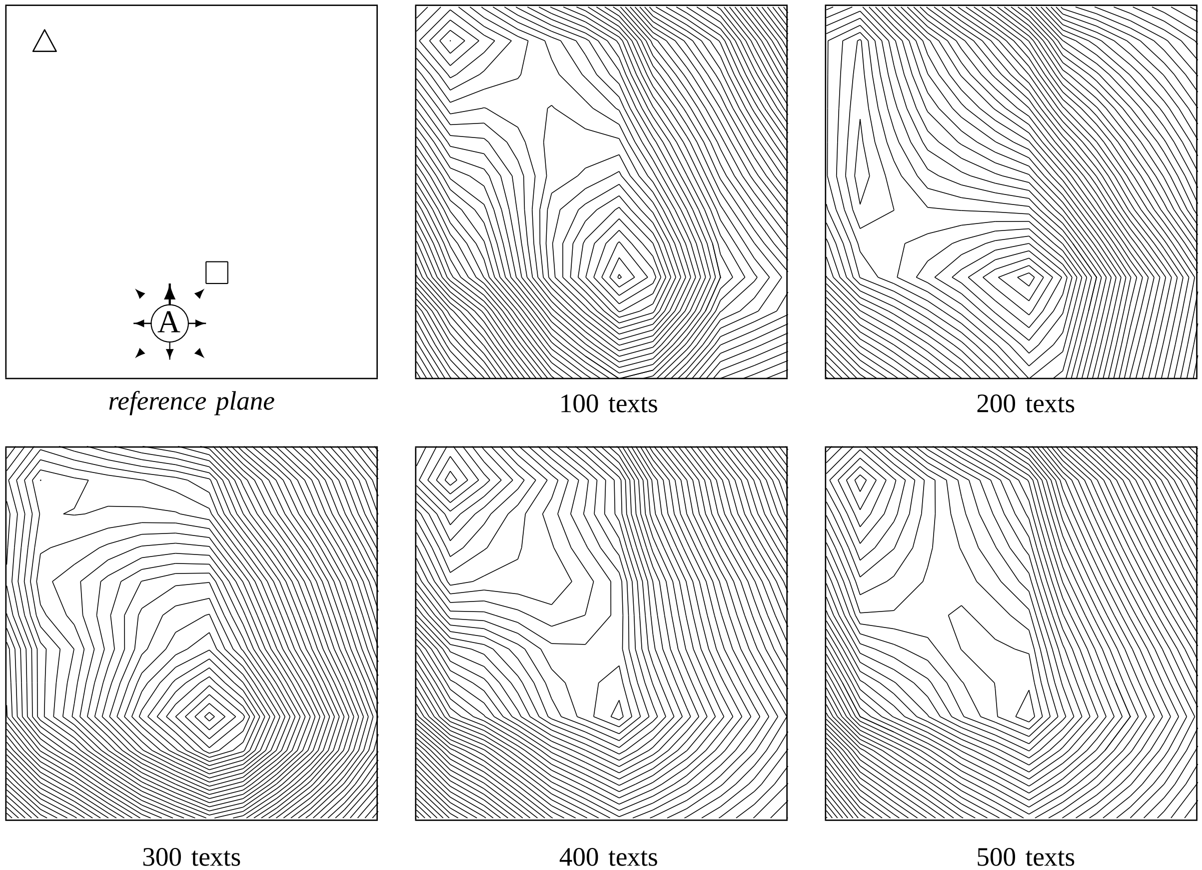

Fig. 1. Reference plane (2-dim *reality*) with two stationary objects △ and □ and a mobile agent A, oriented North. The agent's random walk produces changing system-positions relative to the object-locations (SPOL relations) whose descriptions in simple, declarative sentences (*propositions*)[3] are automatically generated employing four *core predicates* (*left, right, front, behind*) modified by five *hedge predicates* (first order: *near, far*; second order: *extremely, very, rather*) as specified by a formal grammar (with *syntax* and *semantics*). These define and control the semantic contents of the language *descriptions* generated, not however the way these descriptions are processed by the SCIP model. The processing results are visualized as 2-dim images of potential object locations (*isoreferentials*) depicting the SCIP system's *understanding* as states of incremental *meaning constitution* based on sub-symbolic, perceptual processing of textual constraints in an increasing number of (100 to 500) texts. Thus, the SCIP language understanding performance can be tested against the real world situations which these texts describe and refer to.

4. Experiments and Tests

The 2-dim scenario of the *real world* (Fig.1 upper left) is a reference plane with two stationary objects (environment), and an oriented mobile SCIP agent (system) which are structurally coupled [10] by a corpus of *situated* (true and correct) NL expressions[3] 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 to compute its over-all structure as an image (Fig.1) of regions of potential object locations by profile lines of common likelihood (*isoreferentials*).

A prototype SCIP implementation will be presented realizing the formally controlled *description* of changing real world situations, and the SCIP system's subsequent *understanding* of these descriptions in a multi-level process of constraint detection and representation whose visualizations allow for *ad oculos* tests of the system's understanding capabilities. The

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

demonstration of these processes cover variable system-environment situations to illustrate the real-time performance of a perception based, procedural approach to the dynamics of semiotically grounded *meaning constitution* as a base model for (natural language) *understanding*.

REFERENCES

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

[2] K. Devlin: *Logic and Information*, Cambridge (Cambridge UP), 1991.

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

[4] H. Pattee: Simulations, Realizations, and Theories of Life. In: Langton (ed): *Artificial Life*, [Santa Fé Institute Studies in the Science of Complexity VI], Reading, MA (Addison Wesley), 1989, pp. 63–77.

[5] 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.

[6] B. B. Rieger: Meaning Acquisition by SCIPS. In: Ayyub (ed): *ISUMA-NAFIPS-95*, Los Alamitos, CA (IEEE Computer Society Press), 1995, pp. 390–395.

[7] B. B. Rieger: Situations, Language Games, and SCIPS. Modeling semiotic cognitive information processing systems. In: Meystel/Nerode (eds): *Architectures for Semiotic Modeling and Situation Analysis in Large Complex Systems*, Bala Cynwyd, PA (AdRem), 1995, pp. 130–138.

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

[9] B. B. Rieger: Semiotic Cognitive Information Processing: Learning to Understand Discourse. A Systemic Model of Meaning Constitution. In: Kühn/ et al. (eds): *Perspectives on Adaptation and Learning*, Berlin/ Heidelberg/ New York (Springer), 2003, pp. 347–403.

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

[11] L. Wittgenstein: *The Blue and Brown Books*. R. Rhees' edn., Oxford (Blackwell), 1958.

[12] L. Wittgenstein: *Über Gewißheit – On Certainty.*. G.E.M. Anscombe's and G.E. von Wright's edn., New York/ San Francisco/ London (Harper & Row), 1969.

[13] L. A. Zadeh: Quantitative Fuzzy Semantics. *Information Science*, 3: 159–176, 1971.

Burghard B. Rieger, Professor em. of Computational Linguistics and Head of Department of Linguistic Computing at the University of Trier, Germany, has been a researcher and academic teacher for more than three decades. His interdisciplinary work is on topics ranging from German language and literature to linguistics and cognitive science with an early affinity to quantitative and computational approaches. Most of his research is in computational semantics and knowledge representation with special focus on vagueness and fuzzy modeling. His recent work and current interest is in computational semiotics as the study and implementation of dynamic systems of meaning acquisition and language understanding by man and machine. — He received his PhD and Dr. habil. in Linguistics from the Technical University (RWTH) Aachen and held various university positions as researcher, lecturer, and visiting professor (Nottingham, Aachen, Amsterdam, Essen, Trier) before he was appointed Professor *ordinarius* (Chair of Computational Linguistics) at the University of Trier (1986). He wrote two books on quantitative text analysis and on fuzzy computational semantics and is the author of more than 80 articles. He is the editor of several collections and conference proceedings on topics in Empirical Semantics, Computational Linguistics, and Linguistic Computing. He was president of the German Society for Linguistic Computing GLDV (1989–93) and vice-president of the International Society for Terminology and Knowledge Engineering TKE (1990–94), served as Dean and Vice-Dean of his Faculty (1997–2001), and is now Professor *emeritus* of Trier University.