On Distributed Representation in Word Semantics*

Burghard B. Rieger
Department of Computational Linguistics – University of Trier
P.O.Box 3825 – D-5500 TRIER – Germany

Abstract

The dualism of the rationalistic tradition of thought is sketched in view of the *semiotic problem* of meaning constitution. Being a process of cognition which is based upon communicative interaction by signs, their usages (in linear order and selective combination) constitute language structures. Other than *symbolic* representational formats employed sofar in natural language processing by machine, it is argued here, that *distributional* representations correspond directly to the way word meanings are constituted and understood (as fuzzy structures of world knowledge) by (natural and artificial) information processing systems. Based upon such systems' theoretical performance in general and the pragmatics of communicative interaction by real language users in particular, the notions of *situation* and *language game* as introduced by Barwise/Perry and Wittgenstein respectively are combined to allow for a numerical reconstruction of processes that simulate the constitution of meaning and the interpretation of signs. This is achieved by modelling the linear or *syntagmatic* and selective or *paradigmatic* constraints which natural language structure imposes on the formation of (strings of) linguistic entities. A formalism, a related algorithm, and test results of its implementation are given in order to substantiate the claim for an artificial *cognitive information processing system (CIPS)* that operates in a linguistic environment as some meaning acquisition and understanding device.

1 Why should it be aimed at? or the semiotic problem.

Although our understanding of the bunch of complex intellectual activities subsumed under the notion of *cognition* is still very limited, particularly in how knowledge is acquired from texts and what processes are responsible for it, recent achievements in wordsemantics, conceptual structuring, and knowledge representation within the

*Some preliminary ideas for this paper were presented in a number of talks delivered on various occasions among which the 7th Workshop on Parallel Processing, Logic, Organization, and Technology (WOPPLOT 89), at Wildbad Kreuth, Germany, and the Joint Annual Meeting (Spring) of *The Institute for Management Science* and the *Operations Research Society of America* (TIMS/ORSA 1990), Las Vegas, Nevada, USA, deserve mentioning because of subsequent, very stimulating discussions which I owe a lot. — The completion of this paper was made possible by a grant from *The German Marshall Fund of the United States* during my Sabbatical stay as a visiting scholar to the ICSI.

intersection of cognitive psychology, artificial intelligence and computational linguistics appear to offer promising results. Their seminal combination is likely to gain momentum in the future in a wide range of disciplines and applications concerned with *natural language understanding* by machine, opening up new vistas to overcome the traditional duality of mind and matter in models of meaning.

1.1 The dualism of the rationalistic tradition of thought—as exemplified in its notions of some independent (objective) reality and the (subjective) conception of it—has been, and still appears to be, the common ground and widely accepted frame for modelling the semantics of natural language. According to this view, the meaning of a language term is conceived as something static which is somehow related to (and partly derivable from) certain other entities, called signs, any term is composed of. As signs and their meanings are related by some function, called *interpretation*, language *terms* composed of such *signs*, and their associated *meanings* are conceived as forming structured sets of entities which—by virtue of their being signs—at the same time do belong to the (objective) reality and its (subjectively) interpretable representation of it.
According to this conception, these very sets of related entities (or parts thereof) will picture reality (or its recognized structuredness) to the extent the signs employed are interpretable. Therefore, some (linguistic and world) knowledge has to be presupposed and accessible in order to let signs and their meanings be identified and their understanding be derived via interpretation. Hence, *understanding* of language expressions could basically be identified with a process of matching some input strings with supposedly predefined configurations of word meanings and/or world structures whose representations had to be accessible to the understanding system (natural or artificial) as provided by its particular (though limited) *knowledge*. The so-called *cognitive paradigm* [1] of advanced structural and particularly procedural linguistics can easily be traced back to stem from this fundamental duality, according to which natural language understanding can be modelled as the knowledge-based processing of information.

1.2 Subscribing to this notion of understanding, however, tends to be tantamount to accepting certain presuppositions of theoretical linguistics (and particularly some of its model-theoretical semantics). They may be exemplified by the representational means developed and used so far in cognive psychology (*CP*), artificial intelligence (*AI*), and computational linguistics (*CL*).
These approaches employ mostly graph-theoretical formats of trees and nets (*Fig.* 1). Their nodes/vertices and the arcs/edges between them are meant to depict en-

[1]"In adopting a mentalist, individual-oriented stance, the *cognitive paradigm* sets itself apart both, from approaches concentrated on the analysis of observable language use (*performance*), and from those that consider social interaction to be primary for communication. In hypothesizing that the relevant aspects of knowledge (*competence*) can be characterized in formal structures, the *cognitive paradigm* is in disagreement with views such as *phenomenology* which argue that there is an ultimate limitation in the power of formalization and that the most important aspects of language lie outside its limits." (Winograd 1983, pp.20–21)

tities of variant ontological status, like e.g. objects, properties, relations, processes, meanings, etc., or classes thereof, like e.g. concepts, types, variables, slots, etc., to form larger representational structures, like e.g. frames, scenes, scripts, etc. which are to be specified by the kind of labels attached (and/or functions related) to them.

Figure 1: Graph-theoretical formats for *categorial*-type representations of knowledge and meaning.

Depending on the theoretical framework provided by the respective disciplines' epistomological basis or performative goals, these formats converge in being essentially *symbolic* representations of a *categorial*-type [2] format. Roughly, these can be characterized as consisting of sets of related entities of some well-defined, pre-structured kind (the *world*-structure) which are (to be) associated with sets of some other kind of entities (the *sign*-labels) or aggregates thereof. This is achieved by the well-established, pre-defined meanings that the signs supposedly have which—in turn—may be understood not only to relate them to the entities (by way of *referring*) but also to relate the entities (by way of *interpreting*) as modelled by the *symbolic* representations. Accordingly, *word meanings* and/or *world knowledge* are uniformly represented as (more or less complex) labelled graphs (*Fig.* 2) with the (tacid) understanding that associating vertices and edges with labels from some interpreted system of sign-entity-relationships (like e.g. that of natural language terms and their meanings) will render these graph-theoretical configurations also an interpreted model of the structures of either the sign-system that provided the graphs' labels or the system of entities that was to be depicted.
Although from a rationalistic point-of-view there seems no other way to describe and discuss any of the semantic characteristics and properties of meaning outside of and independent from its *symbolic representation* and *declarative predication*, the mere application of these techniques in semantic modelling will only repeat the process of ascribing some properties to some entities on another propositional level, it will not, however, provide the *semiotic* answer to how signs may function as symbols the way the do for cognitive system (natural or artificial), and why a predicate can be declared of anything and be interpreted and understood the way it is (or is not). Obviously, such representational formats do not model the *processes* and the emergence of *structures* that constitute word meaning and/or world knowledge, but

[2]"Behind all the theories of linguistic structure that have been presented in the twentieth century, there is a common set of assumptions about the nature of the structural units. This set of assumptions can be called 'categorial view'. It includes the implicid assertion that all linguistic units are categories which are discrete, invariant, qualitatively distinct, conjunctively defined, [and] composed of atomic primes." (Labov 1973, p.342)

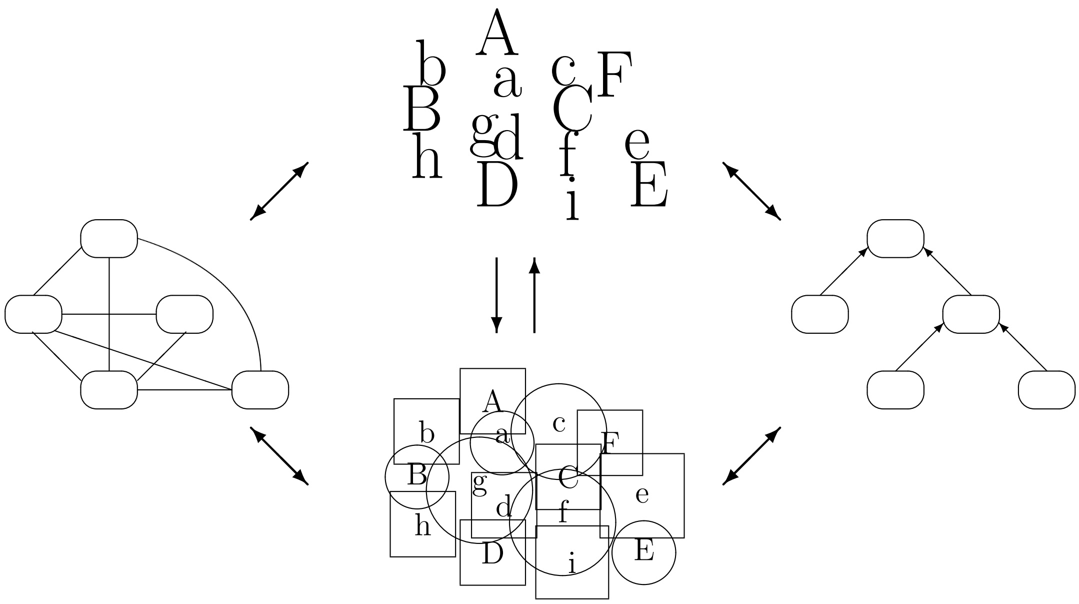

Figure 4: *Language games* employ, modify and/or create signs and entities by establishing entity-sign-relationships via recurrent types of correlated items and structures, by way of *semiosis*.

merely make use of them [3]. As it is agreed that cognition is (among other commitments) responsible for, if not even identifiable with, the processes of how a previously unstructured surrounding for a cognitive system may be devided into some identifiable portions the structural results of which are open to permanent revision according to the system's capabilities, there are still considerable difficulties to understand how by such a hypothetical structuring of the unstructured an (at least) twofold purpose can be served namely

- ▷ to let such identifiable portions—as a sort of prerequisit to entity formation—acquire *situational significance* (*Fig.* 3) for a system, and
- ▷ to let some of these entities by way of their particular situational significance be recognized as *signs* whose interpretations may vary according to the *language games* (*Fig.* 4) these entities are employed to be elements of.

It should therefore be tried to reconstruct both, the *significance* of entities and the meanings of *signs* as a function of a first and second order semiotic embedding relation of *situations* (or contexts) and of *language games* (or cotexts), a cognitive information processing system (CIPS) is not only a part but the procedural constituent of the entire process (*Fig.* 5). There is some chance for doing so because human beings as CIPSs with symbol manipulation and understanding capabilities of highest performance have *language* at their disposal whose structuredness may serve as guidelines. This is a powerful cognitive means not only to represent entities

[3]For illustrative examples and a detailed discussion see Rieger 1985b; 1988; 1989b, Chapter 5: pp.103–132.

and their very complex relations but also to experiment with and to test hypothetical structures and models of entities by way of natural language texts. As their ontological status is again that of very complex structured entities whose *first order* situational significance appears to be identical with their being signs, aggregates, and structures thereof, their *second order* situational significance which allows for their semantic interpretatibility is constituted by their being an instatiation of some language game. Therefore, word meaning may well be reconstructable through the analyses of those elastic constraints which the two levels of semiotic embedding impose on natural language texts constituting language games.
It has long been overlooked that relating arc-and-node structures with sign-and-term labels in symbolic knowledge representation formats is but another illustration of the traditional *mind-matter*-duality. In presupposing a realm of *meanings* and their relations independent of but very much like the objects and structures in the *real world*, this duality does neither allows to explain where the structures come from nor how the signs and labels come to signify anything at all. Their emergence, therefore, never occurred to be in need of some explanatory modelling because the existence of *objects, signs* and *meanings* seemed to be out of all scrutiny and hence was accepted unquestioned. Under this presupposition, fundamental *semiotic* questions of *semantics*—simply did not come up, they have hardly been asked yet [4], and are still far from being solved.

1.3 As long as meanings were conceived as some independent, pre-existing and stabel entities, very much like objects in a presupposed real world, these meanings could be represented accordingly, i.e. as entries to a knowledge base built up of structured sets of elements whose semantics were signalled symbolically by linguistic labels attached to them. However, the fundamental question of how a label may be associated with a node in order to let this node be understood to stand for the entitity (meaning or object) it is meant to represent in a knowledge base, has to be realized, explored, and eventually answered:

- ▷ it has to be *realized* that there are certain entities in the world which are (or become) signs and have (or acquire) interpretable meaning in the sense of signifying something else they stand for, beyond their own physical existence (whereas other entities do not).
- ▷ it has to be *explored* how these (*semiotic*) entities may be constituted and how the meaning relation be established on the basis of which regularities of observables (*uniformities*), controlled by what constraints, and under which boundary conditions of *pragmatic* configuration of communicative interactions like *situations*.
- ▷ it has to be *answered* why some entities may signify others by serving as labels for them (or rather by the *meanings* these labels purport), instead of being signified semiotically by way of positions, load values and/or states distributed over a system of *semiotic/non-semiotic* entities which allows for the distinction

[4] see however Rieger (1977)

of different *distributional* patterns being made, not however, for representing these by different (*symbolic*) labels.

In doing so, a *semiotic paradigm* will have to be followed which hopefully may allow to avoid (if not to solve) a number of spin-off problems, which originate in the traditional distinction and/or the methodological separation of the meaning of a language's term from the way it is employed in discourse. It appears that failing to mediate between these two sides of natural language semantics, phenomena like *creativity, dynamism, efficiency, vagueness,* and *variability* of meaning—to name only the most salient—have fallen in between, stayed (or be kept) out of the focus of interest, or were being overlooked altogether, sofar. Moreover, the classical approach informal theory of semantics which is confined to the sentence boundary of propositional constructions, is badly in want of operational tools to bridge the gap between formal theory of language description (*competence*) and empirical analysis of language usage (*performance*) that is increasingly felt to be responsible for the unwarranted abstractions of fundamental properties of natural languages.

2 What is it based upon? or the situational setting.

The enthusiasm which the advent of the '*electronic brains*' had triggered during the 1950s and early 1960s was met by promising learning machines of which the pattern-recognizing perceptron-type[5] was widely discussed. Processing of numerical vector values representing features loadings of a described entity consisted in cycles of systematic change of weights of features according to the actual input data. Starting from a random set of values such systems—under certain boundary conditions—would converge in a finite number of cycles to the desired set of feature weightings whose *distribution*, instead of a single *symbol*, would represent the entity.
Due to the apparantly essential incapabilities these architectures were criticized for[6], neural networking went out of fashion with the early 1970s. Justified or not, mainstream research turned to *symbolic* instead of *distributed* representational formats of knowledge and information processing with the investigation in decision-making and problem-solving tasks gaining importance over knowledge acquisition and learning which became second rate problems.
Meanwhile, the hardware situation has changed, microelectronic circuitry is available to allow parallel computing devices to then unforeseeable extent, and a revival of the early connectionist approaches can be witnessed. The reasonable attraction, however, which advances in parallel distributed processing (*PDP*)[7] have gained recently in both, cognitive and computer sciences appear to be unwarranted in respect to some of the underlying presuppositions that these models share with more traditional, declarative and predicative formats of word meaning and/or world knowledge representation.

[5]Rosenblatt (1962)
[6]Minsky/Papert (1969)
[7]Rumelhart/McClelland (1986)

2.1 From a computational point-of-view, the so-called *local* representation of entities appears to be the most natural: in a given network of computing elements each of them will be identified with one of the entities to be represented, so that the properties and the relations of the original's elements are mirrored by the structure of the network representing them. Alternatively,

> given a parallel network, items can be represented [. . .] by a pattern of activity in a large set of units with each unit encoding a microfeature of the item. Distributed representations are efficient whenever there are underlying regularities which can be captured by interactions among microfeatures. By encoding each piece of knowledge as a large set of interactions, it is possible to achieve useful properties like content-addressable memory and automatic generalization, and new items can be created without having to create new connections at the hardware level. In the domain of continuously varying spatial features it is relatively easy to provide a mathematical analysis of the advantages and drawbacks of using distributed representations.[8]

Only very recently, however, the underlying presuppositions have been addressed critically to set forth some fundamental questioning[9]. To let a given network of *distributed* representations perform the way it does will necessitate the—mostly implicid— introduction of foils and filters, at least during the learning phase. In these models of automatic generalizing or learning, initial or underlying structures have to be presupposed in order to combine constraints of different level to match specified patterns or parts of it, instead of inducing them[10]
Approaching the problem from a *cognitive* point-of-view, it can still be conceded that any identification and interpretation of external structures has to be conceived as some form of *information processing* which (natural/artificial) systems—due to their own structuredness—are (or ought to be) able to perform. These processes or the structures underlying them, however, ought to be derivable from rather than presupposed to procedural models of meaning. Other than in those approaches to cognitive tasks and natural language understanding available sofar in information processing systems that *AI* or *CL* have advanced, it is argued here that *meaning* need not be introduced as a presupposition of *semantics* but may instead be derived as a result of semiotic modelling. It will be based upon a phenomenological reinterpretation of the formal concept of *situation* and the analytical notion of *language game*. The combination of both lends itself easily to operational extensions

[8]Hinton/McClelland/Rumelhart 1986, p.108

[9]"Learning in structured connectionist systems has been studied directly. A major problem in this formulation is 'recruiting' the compact representation for new concepts. It is all very well to show the advantages of representational schemes [. . . of networks of distributed structures], but how could they arise? This question is far from settled, but there are some encouraging results. The central question is how a system that grows essentially no new connections could recruit compact groups of units to capture new concepts and relations." (Feldman 1989, p.40)

[10]"In brief, there are more problems than solutions. Although it is true that one may view Connectionism as a new *research programm*, expecting it to solve the difficult problems of language without wiring in more traditional symbolic theories by hand is a form of day-dreaming." (Braspenning 1989, p.173)

in empirical analysis and procedural simulation of associative meaning constitution which may grasp essential parts of what PEIRCE named *semiosis*[11].

2.2 Revising some fundamental assumptions in model theory, BARWISE/PERRY have presented a new approach to formal semantics which, essentially, can be considered a mapping of the traditional duality, mediated though by their notion of *situation*. According to their view as expressed in *Situation Semantics*[12] , any language expression is tied to reality in two ways: by the *discourse situation* allowing an expression's meaning being *interpreted* and by the *described situation* allowing its interpretation being *evaluated* truth-functionally. Within this relational model of semantics, *meaning* appears to be the derivative of information processing which (natural/artificial) systems—due to their own structuredness—perform by recognizing similarities or invariants between situations that structure their surrounding realities (or fragments thereof).
By recognizing these invariants and by mapping them as *uniformities* across *situations*, cognitive systems properly *attuned* to them are able to identify and understand those bits of information which appear to be essential to form these systems' particular view of reality: a flow of *types of situations* related by *uniformities* like individuals, relations, and time-space-locations which constrain an external "world teaming with meaning"[13] to become fragments of persistent *courses of events* whose expectability renders them interpretable.
In semiotic sign systems like natural languages, such uniformities appear to be signalled by *word-types* whose employment as *word-tokens* in texts exhibit a special form of *structurally conditioned* constraints. Not only allows their use the speakers/hearers to convey/understand meanings differently in different discourse situations (*efficiency*), but at the same time the discourses' total vocabulary and word usages also provide an empirically accessible basis for the analysis of *structural* (as opposed to *referencial*) aspects of *event-types* and how these are related by virtue of word-uniformities accross phrases, sentences, and texts uttered. Thus, as a means for the *intensional* (as opposed to the *extensional*) description of (abstract, real, and actual) *situations*, the regularities of word-usages may serve as an access to and a representational format for those elastic constraints which underly and condition any word's *linguistic meaning*, the *interpretations* it allows within possible contexts of use, and the *information* its actual employment on a particular occasion may convey.
Owing to BARWISE/PERRYs situational approach to semantics—and notwithstanding its (mis)conception as a duality (i.e. the *independent-sign-meaning* view) of an information-processing system on the one hand which is confronted on the other hand with an external reality whose accessible fragments are to be recognized as its environment—the notion of *situation* proves to be seminal. Not only can it be

[11]"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." (Peirce 1906, p.282)
[12]Barwise/Perry (1983)
[13]Barwise/Perry (1983), p.16

employed to devise a procedural model for the situational embeddedness of cognitive systems as their primary means of mutual accessability[14], but also does it allow to capture and specify the semiotic unity of the notion of *language games* (i.e. the *contextual-usage-meaning* view) as introduced by WITTGENSTEIN:

> And here you have a case of the *use of words*. I shall in the future again and again draw your attention to what I shall call *language games*. There are ways of using signs simpler than those in which we use the signs of our highly complicated everyday language. Language games are the forms of language with which a child begins to make use of words. The study of language games is the study of primitive forms of language or primitive languages. If we want to study the problems of truth and falsehood, of the agreement and disagreement of propositions with reality, of the nature of assertion, assumption, and question, we shall with great advantage look at primitive forms of language in which these forms of thinking appear without the confusing background of highly complicated processes of thought. [. . .] We are not, however, regarding the *language games* which we describe as incomplete parts of a language, but as languages complete in themselves, as complete systems of human communication[15].

2.3 Trying to model *language game* performance along traditional lines of cybernetics by way of, say, an information processing *subject*, a set of objects surrounding it to provide the informatory *environment*, and some positive and/or negative *feedback* relations between them, would hardly be able to capture the cognitive dynamism that self-organizing systems of knowledge acquisition and meaning understanding are capable of[16].

It is this dynamism of cognitive processing in natural systems which renders the so-called *cognitive paradigm* of information processing of current artificial systems so unsatisfactory. Modelling the meaning of an expression along reference-theoretical lines has to presuppose the structured sets of entities to serve as range of a denotational function which will provide the expression's interpretation. Instead, it appears feasible to have this very range be constituted as a result of exactly those cognitive procedures by way of which understanding is produced. It will be modelled as a multi-level dynamic description which reconstructs the possible structural connections of an expression towards cognitive systems (that may both intend/produce and realize/understand it) and in respect to their *situational* settings, being specified by the expressions' pragmatics.

[14] Rieger/Thiopoulos 1989

[15] Wittgenstein (1958), pp.17 and 81; *my italics*

[16] "[. . .] feedback is a method of controlling a system by reinserting into it the results of its past performance. If these results are merely used as numerical data for the criticism of the system and its regulations, we have the simple feedback of control engineers. If, however, the information which proceeds backward from the performance is able to change the general method and pattern of perfomance, we have a process which may well be called learning." (Wiener 1958, p.60)

In phenomenological terms, the set of structural constraints defines any cognitive (natural or artificial) system's possible range in constituting its schemata whose instantiations will determine the system's actual interpretations of what it perceives. As such, these cannot be characterized as a domain of objective entities, external to and standing in contrast with a system's internal, subjective domain; instead, the links between these two domains are to be thought of as *ontologically fundamental*[17] or pre-theoretical. They constitute—from a *semiotic* point-of-view—a system's primary means of access to and interpretation of what may be called its "world" as the system's particular apprehension of its environment. Being fundamental to any cognitive activity, this basal identification appears to provide the grounding framework which underlies the duality of categorial-type rationalistic mind-world or subject-object separation.

In order to get an idea of what is meant by the pre-theoretical *proto-duality* of *semiosis*, any two of the feedback-related operational components separated in system-and-environment, in subject-and-object, or in mind-and-matter distinctions are to be thought of as being merged to form an indecomposable model which bears the characteristics of a self-regulating, *autopoietic* system

> organized (defined as a unity) as a network of processes of production, transformation, and destruction of components that produces the components which: (i) through their interactions and transformations regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it as a concrete unity in the space in which they exist by specifying the topological domain of its realization as such a network.[18]

Together, these approaches may allow for the development of a process-oriented system modelling cognitive experience and semiotic structuring procedurally. Implemented, this system will eventually lead to something like machine-simulated cognition, as an intelligent, dynamic perception of reality by an information processing system and its textual surroundings, accessible through and structured by world-revealing (linguistic) elements of communicative language use. For natural language semantics this is tantamount to (re)present a term's meaning potential by a *distributional pattern* of a modelled system's state changes rather than a *single symbol* whose structural relations are to represent the system's interpretation of its environment. Whereas the latter has to *exclude*, the former will automatically *include* the (linguistically) structured, pragmatic components which the system will both, embody and employ as its (linguistic) import to identify and to interpret its environmental structures by means of its own structuredness.

Thus, the notion of *situation* allows for the formal identification of both, the (*internal*) structure of the cognitive subject with the (*external*) structure of its environment. Perceived as a situational fragment of the objective world, a n d exhibited as systematic *constraints* of those systems that are properly attuned, the common

[17] Heidegger (1927)

[18] Maturana/Varela (1980), p.135

persistency of courses-of-events will be the means to *understand* a linguistically presented reality.
Based upon the fundamentals of *semiotics*, the philosophical concept of communicative *language games* as specified by the formal notion of *situations*, and tied to the observables of actual language performance, allows for an empirical approach to word semantics. What can formally been analyzed as *uniformities* in BARWISEian *discourse situations* may be specified by word-type regularities as determined by co-occurring word-tokens in pragmatically homogeneous samples of natural language texts. Going back to the fundamentals of structuralistic descriptions of regularities of *syntagmatic* linearity and *paradigmatic* selectivity of language items, the correlational analyses of discourse will allow for a two-level word meaning and world knowledge representation whose dynamism is a direct function of elastic constraints established and/or modified in communicative interaction by use of linguistic signs in language performance.

3 How could it be achieved? or the linguistic solution.

The *representation* of knowledge, the *understanding* of meanings, and the *analysis* of texts, have become focal areas of mutual interest of various disciplines in cognitive science. In linguistic semantics, cognitive psychology, and knowledge representation most of the necessary data concerning lexical, semantic and external world information is still provided introspectively. Researchers are exploring (or make test-persons explore) their own linguistic or cognitive capacities and memory structures to depict their findings (or to let hypotheses about them be tested) in various representational formats. By definition, these approaches can map only what is already known to the analysts, not, however, what of the world's fragments under investigation might be conveyed in texts unknown to them. Being *interpretative* and unable of auto-modification, such knowledge representations will not only be restricted to predicative and propositional structures which can be mapped in well established (concept-hierarchical, logically deductive) formats, but they will also lack the flexibility and dynamics of more *re-constructive* model structures adapted to automatic meaning analysis and representation from input texts. These have meanwhile been recognized to be essential[19] for any simulative model capable to set up and modify a system's own knowledge structure, however shallow and vague such knowledge may appear compared to human understanding.

3.1 Other than introspective data acquisition and in contrast to classical formalisms for knowledge representation which have been conceived as depicting some of the (inter)subjective reflections of entities which an external, objective world and reality would provide, the present approach focusses on the *semiotic* structuredness which the communicative use of language in discourse by speakers/hearers will both, constitute and modify as a paradigm of cognition and a model of *semiosis*. It has been based on the algorithmic analysis of discourse that real speakers/writers

[19] Winograd (1986)

produce in actual situations of performed or intended communication on a certain subject domain. Under the notion of *lexical relevance* and *semantic disposition*[20], a conceptual meaning representation system has operationally been defined which may empirically be reconstructed from natural language texts.

Operationalizing the WITTGENSTEINian notion of *language games* and drawing on his assumption that a great number of texts analysed for the terms' usage *regularities* will reveal essential parts of the concepts and hence the meanings conveyed[21], such a description turns out to be identical with a analytical procedure. Starting from the sets of possible combinations of language units, it captures and reformulates their *syntagmatic* and *paradigmatic* regularities (providing the units function as *signs*) via two consecutive processes of abstraction based upon constraints that can empirically be ascertained.

In terms of *autopoietic* systems, it is a mere presupposition of propositional level approaches to natural language semantics that linguistic entities which may be combined to form language expressions must also have independent meanings which are to be identified first in order to let their composite meanings in discourse be interpreted. This presupposition leads to the faulty assumption that word meanings are somewhat static entities instead of variable results of processes constituted via *semiotically* different levels of abstraction. Although structural linguistics offers some hints[22] towards how language items come about to be employed the way they are, these obviously have not been fully exploited yet for the reconstructive modelling of such abstractions which will have to be executed on different levels of description and analysis too.

Thus, complementing the *independent-sign-meaning* view of information processing and the propositional approach in *situation semantics*, the *contextual-usage-meaning* view in word semantics may open up new vistas in natural language processing and its semantic models[23].

3.2 Within the formal framework of situation semantics, lexical items (as *word-types*) appear to render basic uniformities (as *word-tokens*) in any discourse whose *syntagmatic* or *linear* a n d *paradigmatic* or *associative*[24] relatedness can not only be formalized in analogy to *topos theoretical* constructions[25] but also allows for the empirical analyses of these structures and their possible restrictions in order to devise mechanisms to model operational constraints.

---

[20]Rieger 1985a

[21]Wittgenstein (1969)

[22]In subscribing to the systems-view of natural languages, the distinction of *langue/parole* and *competence/performance* in modern linguistics allowes for different levels of language description. Being able to *segment* strings of language discourse and to *categorize* types of linguistic entities is to make analytical use of the *structural coupling* represented by natural languages as semiotic systems.

[23]Rieger (1989b)

[24]According to the terminology of early linguistic structuralism as well as recent connectionistic models in cognitive networking.

[25]For the mathematical concept of *topoi* see Goldblatt (1984); for its application to natural language semantics see Rieger/Thiopoulos (1989)

These constraints may be formalized as a set of *fuzzy subsets*[26] of the vocabulary. Represented as a set-theoretical system of meaning points, they will depict the distributional character of word *meanings*. Being composed of a number of operationally defined elements whose varying contributions can be identified with values of the respective membership functions, these can be derived from and specified by the differing usage regularities that the corresponding lexical items have produced in discourse. This translates the WITTGENSTEINian notion of *meaning* into an operation that may be applied empirically to any corpus of *pragmatically homogeneous* texts constituting a *language game*.
Based upon the distinction of the *syntagmatic* and *paradigmatic* structuredness of language items in discourse, the core of the representational formalism can be captured by a two-level process of abstraction (called α- and δ-abstraction) providing the set of *usage regularities* and the set of *meaning points* of those word-types which are being instantiated by word-tokens as employed in natural language texts. The resultant structure of these constraints render the set of potential *interpretations* which are to be modelled in the sequel as the *semantic hyperspace structure* (SHS).
It has been shown elsewhere[27] that in a sufficiently large sample of *pragmatically homogeneous* texts produced in sufficiently similar *situational* contexts, only a restricted vocabulary, i.e. a limited number of lexical items, will be used by the interlocutors, however comprehensive their personal vocabularies in general might be. Consequently, the words employed to convey information on a certain subject domain under consideration in the discourse concerned will be distributed according to their conventionalized communicative properties, constituting *usage regularities* which may be detected empirically from texts. These consist of structured sets of strings of linguistic elements which, however, are not considered as sentences but primarily as sequences of non-function words (*lexemes*) that make up these strings (*texts*).

3.3 The statistics used so far for the analysis of *syntagmatic* and *paradigmatic* relations on the level of *words* in discourse, is basically descriptive. Developed from and centred around a correlational measure to specify intensities of co-occurring lexical items, these analysing algorithms allow for the systematic modelling of a fragment of the lexical structure constituted by the vocabulary employed in the texts as part of the concomitantly conveyed world knowledge.
A modified correlation coefficient has been used as a *first* mapping function α. It allows to compute the relational interdependence of any two lexical items from their textual frequencies. For a text corpus

$$K = \{k_t\}, t = 1, \ldots, T \tag{1}$$

of pragmatically homogeneous discourse, having an overall length

$$L = \sum_{t=1}^{T} l_t; 1 \leq l_t \leq L \tag{2}$$

[26]Zadeh (1965)
[27]Rieger (1981)

measured by the number of word-tokens per text, and a vocabulary

$$V = \{x_n\}; n = 1, \ldots, i, j, \ldots, N \tag{3}$$

of n word-types of different identity i, j whose frequencies are denoted by

$$H_i = \sum_{t=1}^{T} h_{it}; 0 \leq h_{it} \leq H_i \tag{4}$$

the modified correlation-coefficient $\alpha_{i,j}$ allows to express pairwise relatedness of word-types $(x_i, x_j) \in V \times V$ in numerical values ranging from -1 to $+1$ by calculating co-occurring word-token frequencies in the following way

$$\alpha(x_i, x_j) = \frac{\sum_{t=1}^{T}(h_{it} - h_{it}^*)(h_{jt} - h_{jt}^*)}{\left(\sum_{t=1}^{T}(h_{it} - h_{it}^*)^2 \sum_{t=1}^{T}(h_{jt} - h_{jt}^*)^2\right)^{\frac{1}{2}}}; \quad -1 \leq \alpha(x_i, x_j) \leq +1 \tag{5}$$

where $h_{it}^* = \frac{H_i}{L} l_t$ and $h_{jt}^* = \frac{H_j}{L} l_t$.

Evidently, pairs of items which frequently either co-occur in, or are both absent from, a number of texts will positively be correlated and hence called *affined*, those of which only one (and not the other) frequently occurs in a number of texts will negatively be correlated and hence called *repugnant* of varying intensities or α-values.
As a fuzzy binary relation,

$$\tilde{\alpha} : V \times V \to I \tag{6}$$

can be conditioned on $x_n \in V$ which yields a crisp mapping

$$\tilde{\alpha} \mid x_n \; : \; V \to C; \; \{y_n\} =: C \tag{7}$$

where the tupels $\langle (x_{n,1}, \tilde{\alpha}(n, 1)), \ldots, (x_{n,N}, \tilde{\alpha}(n, N)) \rangle$ represent the numerically specified, *syntagmatic* usage-regularities that have been observed for each word-type x_i against all other $x_n \in V$ and can therefore be abstracted over one of the components in each ordered pair, thus, by α-*abstraction* defining an element

$$x_i(\tilde{\alpha}(i, 1), \ldots, \tilde{\alpha}(i, N)) =: y_i \in C \tag{8}$$

Hence, the regularities of usage of any lexical item will be determined by the tupel of its *affinity/repugnancy*-values towards each other item of the vocabulary which—interpreted as coordinates— can be represented by points in a vector space C spanned by the number of axes each of which corresponds to an entry in the vocabulary.

3.4 Considering C as representational structure of abstract entities constituted by *syntagmatic* regularities of word-token occurrences in *pragmatically homogeneous* discourse, then the similarities and/or dissimilarities between these abstract entities will capture the *paradigmatic* regularities of the correspondent word-types. These can be modelled by the δ-*abstraction* which is based on a numerically specified evaluation of differences between any two of such points $y_i, y_j \in C$ They will be the more adjacent to each other, the less the usages (*tokens*) of their corresponding lexical items $x_i, x_j \in V$ (*types*) differ. These differences may be calculated by a distance measure δ of, say, EUCLEDian metric.

$$\delta(y_i, y_j) \;=\; \left(\sum_{n=1}^{N}(\alpha(x_i, x_n) - \alpha(x_j, x_n))^2\right)^{\frac{1}{2}}; \qquad 0 \leq \delta(y_i, y_j) \leq 2\sqrt{N} \tag{9}$$

Thus, δ may serve as a *second* mapping function to represent any item's differences of usage regularities measured against those of all other items. As a fuzzy binary relation, also

$$\tilde{\delta} : C \times C \rightarrow I \tag{10}$$

can be conditioned on $y_n \in C$ which again yields a crisp mapping

$$\tilde{\delta} \mid y_n \;:\; C \rightarrow S;\; \{z_n\} =: S \tag{11}$$

where the tupels $\langle(y_{n,1}, \tilde{\delta}(n, 1)), \ldots, (y_{n,N}\tilde{\delta}(n, N))\rangle$ represents the numerically specified *paradigmatic* structure that has been derived for each abstract *syntagmatic* usage-regularity y_j against all other $y_n \in C$. The distance values can therefore be abstracted again as in (7), this time, however, over the other of the components in each ordered pair, thus defining an element $z_j \in S$ called *meaning point* by

$$y_j(\tilde{\delta}(j, 1), \ldots, \tilde{\delta}(j, N)) =: z_j \in S \tag{12}$$

By identifying $z_n \in S$ with the numerically specified elements of potential paradigms, the set of possible combinations $S \times S$ may structurally be constrained and evaluated without (direct or indirect) recourse to any pre-existent external world. Introducing a EUCLEDIAN metric

$$\partial \;:\; S \times S \rightarrow I \tag{13}$$

the hyperstructure $\langle S, \partial\rangle$ or *semantic space* (*SHS*) is constituted providing the *meaning points* according to which the *stereotypes* of associated lexical items may be generated as part of the semantic paradigms concerned.
As a result of the two *consecutive* mappings (*Tab.* 1), any meaning point's position in SHS is determined by all the differences (δ- or distance-values) of all regularities of usage (α- or correlation-values) each lexical item shows against all others in the discourse analysed. Thus, it is the basic analyzing algorithm which—by processing natural language texts—provides the processing system with the ability to recognize and represent and to employ and modify the structural information available to the system's performance constituting its *understanding*.

$$
\begin{array}{ccccc}
 & \alpha\text{-abstraction} & & \delta\text{-abstraction} & \\
V \times V & \Downarrow & C \times C & \Downarrow & S \times S \\
\begin{array}{c|ccc}
\tilde{\alpha} & x_1 & \ldots & x_N \\ \hline
x_1 & \alpha_{11} & \ldots & \alpha_{1N} \\
\vdots & \vdots & \ddots & \vdots \\
x_N & \alpha_{N1} & \ldots & \alpha_{NN}
\end{array}
& \xrightarrow{\tilde{\alpha}|x_i} &
\begin{array}{c|ccc}
\tilde{\delta} & y_1 & \ldots & y_N \\ \hline
y_1 & \delta_{11} & \ldots & \delta_{1N} \\
\vdots & \vdots & \ddots & \vdots \\
y_N & \delta_{N1} & \ldots & \delta_{NN}
\end{array}
& \xrightarrow{\tilde{\delta}|y_j} &
\begin{array}{c|ccc}
\partial & z_1 & \ldots & z_N \\ \hline
z_1 & \partial_{11} & \ldots & \partial_{1N} \\
\vdots & \vdots & \ddots & \vdots \\
z_N & \partial_{N1} & \ldots & \partial_{NN}
\end{array} \\
 & \Uparrow & & \Uparrow & \\
 & \textit{Syntagmatic} & & \textit{Paradigmatic} & \\
 & & \textit{C o n s t r a i n t s} & &
\end{array}
$$

Table 1: Formalizing (*syntagmatic/paradigmatic*) constraints by consecutive (α- and δ-) abstractions over usage regularities of items x_i, y_j respectively.

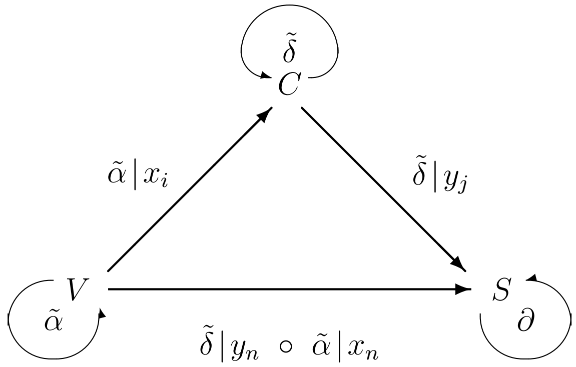

Figure 6: Fuzzy mapping relations $\tilde{\alpha}$ and $\tilde{\delta}$ between the structured sets of vocabulary items $x_n \in V$, of corpus points $y_n \in C$, and of meaning points $z_n \in S$.

This answers the question where the labels in our representation come from: put into a discourse environment, the system's text analyzing algorithm provides the means how the topological position of any metrically specified meaning point $z \in \langle S, \partial \rangle$ is identified and labeled by a vocabulary item $x \in V$ according to the two consecutive mappings which can formally be stated as a composition of the two restricted relations $\tilde{\delta} \mid y$ and $\tilde{\alpha} \mid x$ (*Fig.* 6). It is achieved without recurring to any investigator's or his test-persons' word or world knowledge (*semantic competence*), but solely on the basis of usage regularities of lexical items in discourse which are produced by real speakers/hearers in actual or intended acts of communication (*communicative performance*).

3.5 Sofar the system of word meanings has been represented as a relational data structure whose linguistically labeled elements (*meaning points*) and their

ARBEIT	0.000				
ALLGEMEIN	8.332	ANBIET	8.756	AUSGAB	10.392
STADT	10.711	PERSON	11.075	LEHR	11.811
GEBIET	11.831	VERBAND	12.041	UNTERNEHM	12.130
VERKEHR	12.312	HERRSCH	12.362	VERANTWORT	12.543
EINSATZ	13.980	STELLE	14.120	WERB	15.561
ORGANIS	16.146	VERWALT	16.340	MODE	16.842
GESCHFT	16.873	UNTERRICHT	18.275	BITT	19.614
⋮	⋮	⋮	⋮	⋮	⋮

INDUSTRI	0.000				
SUCH	2.051	ELEKTRON	2.106	LEIT	2.369
BERUF	2.507	SCHUL	3.229	SCHREIB	3.329
WIRTSCHAFT	3.659	COMPUTER	3.667	FHIG	3.959
SYSTEM	4.040	ERFAHR	4.294	KENN	5.286
DIPLOM	5.504	TECHNI	5.882	UNTERRICHT	7.041
ORGANIS	8.355	WUNSCH	8.380	BITT	9.429
STELLE	11.708	UNTERNEHM	14.430	STADT	16.330
GEBIET	17.389	VERBAND	17.569	PERSON	18.938
⋮	⋮	⋮	⋮	⋮	⋮

Table 2: Topological environments $E(z_i, r)$ of i =ARBEIT/labour and INDUSTRIE/industry listing labeled points within their respective hypersheres of radius r in the *semantic space* $\langle S, \partial \rangle$ as computed from a random text sample of the 1964 editions (first two pages) of the German daily DIE WELT (175 articles of approx. 7000 lemmatized word tokens and 365 word types).

mutual distances (*meaning differences*) form a system of potential *stereotypes*. Although these representations by labeled points[28] appears to be *symbolic* it has to be remebered that each such point is in fact defined by a *distribution* of wordtype/value-pairs which allow easy switching between these two representational formats when interpreted topologically, as we have done here. Accordingly, based upon the SHS-structure, the *meaning* of a lexical item may be described either as a fuzzy subset of the vocabulary, or as a meaning point vector, or as a meaning point's topological environment. The latter is determined by those points which are found to be most adjacent and hence will delimit the central point's meaning indirectly as its

[28]It should be noted here, that—SHS being *compact*—only a few of the infinitely many points in *semantic space* are in fact identified by labels (i.e. via *lexicalization*) whereas the majority of space localities are not, with the understanding that any lexical item may not only name a point but rather refers to a region of adjacent (but unlabeled) points in space thus allowing for natural language terms' essential *vagueness*.

prototype (*Tab.* 2).

3.6 Following a *semiotic* notion of *understanding* and *meaning constitution*[29], the *SHS*-structure may be considered the core of a two-level conceptual knowledge representation system[30]. Essentially, it separates the format of a basic (stereotype) word meaning representation from its latent (dependency) relational concept organization. Whereas the former is a rather static, topologically structured (associative) memory, the latter can be characterized as a collection of dynamic and flexible structuring procedures to re-organize the memory data by semiotic principles under various aspects[31]

Other than in pre-defined semantic network structures of predicative knowledge, and unlike conceptual representations that link nodes to one another according to what cognitive scientists supposedly know about the way conceptual information is structured in memory[32], the *SHS*-model may be considered—conceptually speaking—mereley as raw data. Taken as an associative base structure, particular procedures may operate on it whose objective would be to select, reorganize, and at the same time convert existing relations into some node-pointer-type structure.

As non-predicative meaning relations of lexical *relevance* and *perspective* depend haevily on con- and cotextual constraints these will more adequately be defined procedurally, i.e. by generative algorithms that induce them on changing data differently rather than trying to make them up by limited (and doubtful) introspection on the analysts' or their testpersons' side. This is achieved by a recursively defined procedure that produces hierarchies of meaning points, structured in n-ary trees, under perspectival aspects according to and in dependence of their meanings' relevancy.

Given one meaning point's position as a start, the algorithm of *least distances* (*LD*) will

1. list all of the starting point's labeled neighbours and stack them by their increasing distances;
2. prime the starting point as head node or root of the tree to be generated, before the algorithm's generic procedure takes over:
3. it will take the top-most entry from the stack, generate a list of its neighbours, determine from it the least distant that has already been primed, and identify it as the ancestor-node to which the new point is linked as descendant.

   Repeated succesively for each of the meaning points stacked and in turn primed in accordance with this procedure, the *LD*-algorithm will select a particular fragment of the relational structure latently inherent in the semantic space data and depending on the perspectival aspect. i.e. the initially primed meaning point the algorithm is started with.

[29] Rieger (1977)

[30] Rieger (1989b)

[31] This corroborates and extends ideas expressed within the theories of *spreading activation* and their processes of *priming* (Lorch 1982) by allowing the variable and dynamic constitution of paths (along which activation might spread) to be a function of priming, instead of its presupposed condition.

[32] Schank (1982)

noindent Working its way through and consuming all labeled points in SHS—unless stopped under conditions of given target node, number of nodes to be processed, or threshold of maximum distance/minimum criteriality—the LD-algorithm[33] transforms prevailing similarities of meaning as represented by adjacent points to establish a binary, non-symmetric, and transitive relation of lexico-semantic *relevance* between them, conditioned by the *perspective* chosen. Stop conditions may deliberately be formulated either qualitatively (i.e. by naming a target point as final node) or quantitatively (i.e. by the number of nodes to be processed, or the threshold of maximum distance/minimum criteriality). It is this *relevance*-relation induced by the LD-algorithm which constitutes the so-called Δ-*operation* allowing for the hierarchical re-organisation of meaning points as nodes under a primed head in an n-ary tree called *dispositional dependency structure* (DDS)[34].

3.7 To illustrate the feasibility of the Δ-operation's generative procedure, a subset of the relevant, linguistic constraints triggered by the lexical item $x_i,\ i =$ ARBEIT/labour and INDUSTRIE/ industry is given (*Figs.* 7 and 8) in the format of weighted DDS-treegraphs[35].
In addition to the distances given between nodes in the DDSs, a numerical expression has been devised which describes any node's degree of relevance according to the tree structure. As a numerical measure $Cr_i(z_d)$[36], any node's *criteriality* is to be calculated with respect to its position in the tree and its root's (or the chosen aspect's) position in $\langle S, \partial \rangle$. Therefore it has been defined as a function of both, its distance values and its level within its repective tree structure, in the following way:

$$\begin{aligned} Cr_i(z_d)_{\kappa+1} &= Cr_i(z_a)_\kappa * e^{-\frac{\partial(z_d,z_a)}{\partial(z_d,z_i)+1}}\ ;\ 0 \leq Cr_i \leq 1.0 \\ \text{for}\quad Cr_i(z_i)_\kappa &= 1\ \text{ at }\ \kappa = 0 \end{aligned} \tag{14}$$

It may either be understood to measure a head-node's z_i meaning-dependencies on the daughter-nodes z_n or, inversely, to express their meaning-criterialities adding up to an aspect's interpretation as determined by that head's meaning[37]. For a wide range of purposes in processing DDS-trees, differing criterialities of nodes can be used to estimate which paths are more likely being taken against others being followed less likely under priming by certain meaning points, allowing for the numerical assessment of *dependency paths* to trace those intermediate nodes which determine the most relevant associative transitions of any target node under any specifiable aspect or perspective.

---

[33] The *LD*-algorithm is basically a *minimal spanning tree*-algorithm (PRIM 1957) controlled additionally, however, by the respective head-node's position and environment in $\langle S, \partial \rangle$.

[34] Rieger (1985)

[35] As computed from the DIE WELT corpus of newspaper texts.

[36] with the notation of Cr := criteriality-value; z_i := root-node (head); z_a := antecedant-node (mother); z_d := descendant-node (daughter); ∂ := distance-value (between meaning-points); κ := level of tree-structure.

[37] Rieger (1989a)

4 What may it be used for? or the need for CIPS.

From the communicative point-of-view natural language texts, whether stored electronically or written conventionally, will in the foreseeable future provide the major source of scientifically, historically, and socially relevant information. Due to the new technologies, the amount of such textual information continues to grow beyond manageable quantities. Rapid access and availability of *data*, therefore, no longer serves to *solve* an assumed problem of lack of *information* to fill an obvious *knowledge* gap in a given instance, but is instead and will even more so in future *create* a new problem which arises from the abundance of information we are confronted with.
Thus, actual and potential (human) problemsolvers feel the increasing need to employ computers more effectively than hitherto for informational search through masses of natural language material. Although the demand is high for intelligent machinery to assist in or even provide speedy and reliable selection of relevant information under individual aspects of interest within specifyable subject domains, such systems are not yet available.

4.1 Development of earlier proposals[38], only recently resulted in some advance[39] towards an artificial *cognitive information processing system* (*CIPS*) which is capable of learning to understand (identify and interpret) the meanings in natural language texts by generating dynamic conceptual dependencies (for inferencing).
Suppose we have an information processing system with an initial structure of constraints modelled as SHS. Provided the system is exposed to natural language discourse and capable of basic structural processing as postulated, then its (rudimentary) interpretations generated from given texts will not change its subsequent interpretations via altered input-cycles, but the system will come up with differing interpretations due to its modified old and/or established new constraints as structural properties of processing. Thus, it is the structure that determines the system's interpretations, and being subject to changes according to changing environments of the system, constitutes its *autopoetic space*:

> an outopoetic organization constitutes a closed domain of relations specified with respect to the autopoetic organization that these relations constitute, and thus it defines a space in which it can be realized as a concrete system, a space whose dimensions are the relations of production of the components that realize it[40].

Considering a text understanding system as *CIPS* and letting its environment consist of texts being sequences of words, then the system will not only identify these words but—according to its own capacity for α- and δ-abstraction together with its Δ-operation—will at the same time realize the semantic connectedness between their

[38]Rieger (1984)
[39]Rieger (1989c)
[40]Maturana/Varela 1980, p.135

meanings which are the system's state changes or *dispositional dependencies* that these words invoke. They will, however, not only be trigger *DDS* but will at the same time—because of the prototypical or distributed representational format (of the *SHS*) being separated from the dynamic organization of meaning points (in *DDS*)—modify the underlying *SHS*-data according to recurrent *syntagmatic* and *paradigmatic* structures as detected from the textual environment[41].

4.2 In view of a text skimming system under development[42], a basic *cognitive* algorithm has been designed which detects from the textual environment the system is exposed to, those strucural information which the system is able to collect due to the two-level structure of its linguistic information processing and knowledge acquisition mechanisms. These allow for the automatic generation of a pre-predicative and formal representation of conceptual knowledge which the system will both, gather from and modify according to the input texts processed. The system's internal knowledge representation is designed to be made accessible by a front-end with dialog interface. This will allow system-users to make the system *skim* masses of texts for them and display its acquired knowledge graphically in dynamic structures of interdependently formed conceptualisations. These provide variable constraints for the procedural modelling of conceptual connectedness and non-propositional inferencing which both are based on the algorithmic induction of an aspect-dependent relevance relation connecting concepts differently according to differing conceptual perspektives in semantic *Dispositional Dependency Structures (DDS)*. The display of *DDS*s or their resultant graphs may serve the user to acquire an overall idea of what the texts processed are roughly about or deal with along what general lines of conceptual dependencies. They may as well be employed in an knowledge processing environment to provide the user with relevant new keywords for an optimized recall-precision ratio in intelligent retrieval tasks, helping for instance to avoid unnecessary reading of irrelevant texts.

ARBEIT	0.0/1.000	⇐ *Premises* ⇒	0.0/1.000	INDUSTRIE
			2.11/.508	ELEKTRO/electro
ALLGEMEIN/general	8.33/.409		2.03/.285	BERUF/profession
STADT/city	6.79/.229		5.60/.142	UNTERRICHT/instruct
UNTERNEHM/enterpr	5.57/.150		4.51/.088	ORGANISAT/organis
Conclusion ⇒	7.28/.097	WUNSCH	4.88/.052	⇐ *Conclusion*

Table 3: *Semantic inference paths* from the premises ARBEIT/labour and INDUSTRIE/industry to the conclusion WUNSCH/wish/desire

[41] Modelling the principles of such a semiotic system's *autopoietic existence* by means of mathematical *topoi* is one of the objectives of a PhD-thesis (by C. Thiopoulos) just completed at the Deptartment of Computational Linguistics, University of Trier.

[42] Rieger (1988a)

Dispositional dependencies appear to be a prerequisit not only to source-oriented, contents-driven *search* and *retrieval* procedures which may thus be performed effectively on any *SHS*-structure. Due to its procedural definition, it also allows to detect varying dependencies of identically labeled nodes under different aspects which might change dynamically and could therefore be employed in conceptual, pre-predicative, and *semantic* inferencing as opposed to propositional, predicative, and *logic* deduction.
For this purpose a procedure was designed to operate simultaniously on two (or more) *DDS*-trees by way of (simulated) parallel processing. The algorithm is started by two (or more) meaning points which may be considered to represent conceptual *premises*. Their *DDS* can be generated while the actual inferencing procedure begins to work its way (breadth-first, depth-first, or according to highest criteriality) through both (or more) trees, tagging each encountered node. When the first node is met that has previously been tagged by activation from another premise, the search procedure stops to activate the dependency paths from this *concluding* common node back to the *premises*, listing the intermediate nodes to mediate (as illustrated in *Tab.* 3) the *semantic inference paths* as part of the dispositional dependencies structures *DDS* concerned.

4.3 It is hoped that our system will prove to provide a flexible, source-oriented, contents-driven method for the *multi-perspective* induction of dynamic conceptual dependencies among stereotypically represented concepts which—being linguistically conveyed by natural language discourse on specified subject domains—may empirically be detected, formally be presented, and continuously be modified in order to promote the learning and understanding of meaning by *cognitive information processing systems* (*CIPS*) for machine intelligence. As the analytical apparatus allows—as shown—to switch easily between either the *symbolic* or the *distributed* interpretation of representational formats used here, research is under way to emulate what sofar has been analysed as numerical constraints of (correlational) item distributions within a structural model of semantic usage regularities. It is presently being investigated to be remodelled in some connectionist architecture with the advantage of semiotically well established and linguistically well founded empirical data providing testable numerical parameters by weights and grades of activation.

References

Barwise, J./ Perry, J.(1983): Situations and Attitudes. Cambridge, MA (MIT)

Braspenning, P.J. (1989): "Out of Sight, Out of Mind → Blind Idiot. A review of Connectionism in the courtroom." *AiCommunications AICOM*, Vol.2,3/4, pp.168–176

Collins, A.M./ Loftus, E.F. (1975): A spreading activation theory of semantic processing. *Psychological Review* 6(1975) 407–428

Feldman, J.A. (1989): "Connectionist Representation of Concepts" in: Pfeiffer/ Schreter/ Fogelman-Soulié/ Steels, pp.25–45

Forsyth, R./ Rada, R. (1986): Machine Learning. Chichester (Ellis Horwood)
Goldblatt, R. (1984): Topoi. The Categorial Analysis of Logic. (Studies in Logic and the Foundations of Mathematics 98), Amsterdam (North Holland)
Heidegger, M. (1927): Sein und Zeit. Tübingen (M.Niemeyer)
Hinton, G.E./ McClelland, J.L./ Rumelhart, D.E. (1986): "Distributed Representation" in: Rumelhart/ McClelland, pp.77–109
Husserl, E. (1976): Ideen II (*Husserliana* III/1), DenHaag (M. Nijhoff)
Lorch, R.F. (1982): "Priming and Search Processes in Semantic Memory: A test of three models of Spreading activation", *Journal of Verbal Learning and Verbal Behavior* 21, pp.468–492
Maturana, H./ Varela, F. (1980): Autopoiesis and Cognition. The Realization of the Living. Dordrecht (Reidel)
Minsky, M./ Papert, S. (1969): Perceptrons. Cambridge, MA (MIT-Press)
Norvig, P. (1987): Unified Theory of Inference for Text Understanding. (EECS-Report UCB/CSD 87/339) University of California, Berkeley
Peirce, C.S. (1906): "Pragmatics in Retrospect: a last formulation" (CP 5.11 – 5.13), in: The Philosophical Writings of Peirce. Ed. by J. Buchler, NewYork (Dover), pp.269–289
Pfeiffer, R./ Schreter, Z./ Fogelman-Soulié/ Steels, L. (1989)(Eds.): Connectionism in Perspective. Amsterdam/ NewYork/ Oxford/ Tokyo (North-Holland)
Prim, R.C. (1957): "Shortest connection networks and some generalizations", *Bell Systems Technical Journal* 36, pp.1389–1401
Rieger, B. (1977): "Bedeutungskonstitution. Einige Bemerkungen zur semiotischen Problematik eines linguistischen Problems" *Zeitschrift für Literaturwissenschaft und Linguistik* 27/28, pp.55–68
Rieger, B. (1981): Feasible Fuzzy Semantics. In: Eikmeyer, H.J./ Rieser, H. (Eds): Words, Worlds, and Contexts. New Approaches in Word Semantics. Berlin/ NewYork (de Gruyter), pp.193–209
Rieger, B. (1984): "The Baseline Understanding Model. A Fuzzy Word Meaning Analysis and Representation System for Machine Comprehension of Natural Language." in: O'Shea, T.(Ed): Proceedings of the 6th European Conference on Artificial Intelligence (ECAI 84), NewYork/ Amsterdam (Elsevier Science), pp.748–749
Rieger, B. (1985): Lexical Relevance and Semantic Disposition. On stereotype word meaning representation in procedural semantics. In: Hoppenbrouwes, G./ Seuren. P./ Weijters, T. (Eds.): Meaning and the Lexicon. Dordrecht (Foris), pp.387–400
Rieger, B. (1985b): "On Generating Semantic Dispositions in a Given Subject Domain" in: Agrawal, J.C./ Zunde, P. (Eds.): Empirical Foundation of Information and Software Science. NewYork/ London (Plenum Press), pp.273–291
Rieger, B. (1988a): TESKI – A natural language TExt-SKImmer for shallow understanding and conceptual structuring of textually conveyed knowledge. LDV/CL-Report 10/88, Dept. of Computational Linguistics, University of Trier
Rieger, B. (1988b): "Definition of Terms, Word Meaning, and Knowledge Structure. On some problems of semantics from a computational view of linguistics". in: Czap, H./ Galinski, C. (Eds.): Terminology and Knowledge Engineering (Supplement). Frankfurt (Indeks Verlag), pp.25–41
Rieger, B. (1989a): "Situations and Dispositions. Some formal and empirical tools for semantic analysis" in: Bahner, W. (Ed.): Proceedings of the XIV. Intern.Congress of Linguists (CIPL), Berlin (Akademie) [in print]

Rieger, B. (1989b): Unscharfe Semantik. Die empirische Analyse, quantitative Beschreibung, formale Repräsentation und prozedurale Modellierung vager Wortbedeutungen in Texten. Frankfurt/ Bern/ NewYork (P. Lang)

Rieger, B. (1989c): "Reconstructing Meaning from Texts. A Computational View of Natural Language Understanding" (2nd German-Chinese-Electronic-Week (DCEW 89)), LDV/CL-Report 5/89, Dept. of Computational Linguistics, University of Trier

Rieger, B. (1990): "Unscharfe Semantik: numerische Modellierung von Wortbedeutungen als 'Fuzzy'-Mengen" in: Friemel, H.-J. / Müller-Schönberger, G. / Schütt, A. (Eds.): Forum '90. Wissenschaft und Technik. (Informatik Fachberichte 259), Berlin/ Heidelberg/ NewYork/ Tokyo (Springer) 1990, pp.80–104

Rieger, B./ Thiopoulos, C. (1989): Situations, Topoi, and Dispositions. On the phenomenological modelling of meaning. in: Retti, J./ Leidlmair, K. (Eds.): 5th Austrian Artificial Intelligence Conference. (ÖGAI 89) Innsbruck; (KI–Informatik-Fachberichte Bd.208) Berlin/ Heidelberg/ NewYork (Springer), pp.365–375

Rosenblatt, F. (1962): Principles of Neurodynamics. London (Spartan)

Rumelhart, D.E./ McClelland, J.L (1986): Parallel Distributed Processing. Explorations in the Microstructure of Cognition. 2 Vols. Cambridge, MA (MIT)

Schank, R.C. (1982): Dynamic Memory. A Theory of Reminding and Learning in Computers and People. Cambridge/ London/ NewYork (Cambridge UP)

Sklansky, J./ Wassel, G. (1981): Pattern Classifiers and Trainable Machines. Berlin/ Heidelberg/ NewYork (Springer)

Varela, F. (1979): Principles of Biological Autonomy. NewYork (North Holland)

Wiener, N. (1956): The Human Use of Human Beings. Cybernetics and Society. NewYork (Doubleday Anchor)

Winograd, T. (1983): Language as a Cognitive Process. Vol.1 Syntax. Reading, MA (Addison-Wesley)

Winograd,T./ Flores, F. (1986): Understanding Computers and Cognition: A New Foundation for Design. Norwood, NJ (Ablex)

Wittgenstein, L. (1958): The Blue and Brown Books. Ed. by R. Rhees, Oxford (Blackwell)

Wittgenstein, L. (1969): Über Gewißheit - On Certainty. NewYork/ San Francisco/ London (Harper & Row), [No.61–65], p.10e

Zadeh,L.A. (1965): Fuzzy sets. *Information and Control* 8(1965), pp.338–353