

On Generating Semantic Dispositions in a Given Subject Domain

Stereotype representation and dynamic structuring of fuzzy
word meanings for contents-driven semantic processing*

Burghard B. Rieger

Abstract

Modelling system structures of word meanings and/or world knowledge is to face the problem of their mutual and complex relatedness. In linguistic semantics, cognitive psychology, and knowledge representation most of the necessary data concerning lexical, semantic and/or external world information is still provided introspectively. In a rather sharp departure from that form of data acquisition the present approach has been based on the empirical analysis of discourse that real speakers/writers produce in actual situations of performed or intended communication in prescriptive contexts or subject domains. The approach makes essential use of statistical means to analyse usage regularities of words to map their fuzzy meanings and connotative interrelations in a format of stereotypes. Their dependencies are generated algorithmically as multi-perspective dispositions that render only those relations accessible to automatic processing which can — under differing aspects differently — be considered relevant. Generating such semantic dispositional dependencies dynamically by an procedure would seem to be an operational prerequisite to and a promising candidate for the simulation of contents-driven (*analogically-associative*), instead of formal (*logically-deductive*) inferences in semantic processing.

1 Introduction

Current semantic theories of word meanings and/or world knowledge representation regard memory in human or artificial systems of cognition and/or understanding as

*This paper (an intermediate version of which was read on ICCH/83) reports on the empirical foundations of a project in computational semantics on the automatic analysis and representation of natural language meanings in texts. This project was supported by the North Rhine Westphalia Ministry of Science and Research under grant IV A2 FA 8600. Published in: Agrawal, J.C./Zunde, P. (Eds.): Empirical Foundations of Information and Software Science. New York/London (Plenum Press) 1985, pp. 273-291.

a highly complex structure of interrelated concepts. The cognitive principles underlying these structures are poorly understood yet. As the problem of their mutual and complex relatedness has more and more been recognized, different methods and formats have been proposed with differing success to model these interdependencies. However, the work of psychologists, AI-researchers, and linguists active in that field still appears to be determined by their respective discipline's general line of approach rather than by consequences drawn from these approaches' intersecting results in their common field of interest.

In linguistic semantics, cognitive psychology, and knowledge representation most of the necessary data concerning lexical, semantic and/or external world information is still provided introspectively. Researchers are exploring (or make test-persons explore) their own linguistic/cognitive capacities and memory structures to depict their findings (or let hypotheses about them be tested) in various representational formats (lists, arrays, trees, nets, active networks, etc.). It is widely accepted that model structures resulting from these analyses do have a more or less ad hoc character and tend to be confined to their limited theoretical or operational performances within a specified subject domain and/or implemented system. Thus, these approaches — by definition — can only map what of the world's fragment under investigation is already known to the analysts, not, however, what of it might be conveyed in texts unknown to them. Being basically *interpretative* and in want of operational control, such knowledge representations will not only be restricted quite naturally to undisputed informational structures which consequently can be mapped in accepted and well established (concept-hierarchical, logically deductive) formats, but they will also lack the flexibility and dynamics of more *constructive* model structures which are needed for automatic meaning analysis and representation from input texts to allow for a component to build up and/or modify a system's own knowledge, however shallow and vague that may appear compared to human understanding.

Other than these more orthodox lines of introspective data acquisition in meaning and knowledge representation research, the present approach has been based on the algorithmic analysis of discourse that real speakers/writers produce in actual situations of performed or intended communication on a certain subject domain. The approach makes essential use of procedural means to map fuzzy word meanings and their connotative interrelations in the format of conceptual stereotypes. Their varying dependencies constitute dynamic dispositions¹ that render only those concepts accessible which may — within differing contexts differently — be considered relevant under a specified perspective or aspect. Thus — under the notion of *lexical relevance* and *semantic disposition* — a new meaning relation may operationally be defined between elements in a conceptual representation system which in itself

¹Instead of formally introducing any of the algorithms developed and tested so far for the purposes at hand, an impression of their performance and application shall in the sequel be given by way of some — hopefully illustrative — figures and examples. For more detailed introductions the reader is referred to the bibliography at the end of this paper where additional informations on the MESY-project in general and its procedural approach in particular may be found in a number of the author's recent publications.

may empirically be reconstructed from natural language discourse. Such dispositional dependency structures would seem to be an operational prerequisite to and a promising candidate for the simulation of contents-driven (*analogically-associative*), instead of formal (*logically-deductive*) inferences in semantic processing.

After these (1.) introductory lines and more for illustrative purposes rather than for a detailed and qualifying discussion, some of the standard concept and/or word-meaning *representational formats* in memory models and knowledge systems (2.) will be compared in order to motivate our rather strict departure from them in developing and using (3.) some *statistical means* for the analysis of texts and the representation of the data obtained which will briefly be introduced as the semantic space model. Starting from the notion of *priming* and *spreading activation* in memory as a cognitive model for comprehension processes, we will (4.) deal with our procedural method of representing semantic *dispositions* by way of inducing a relation of lexical *relevance* among labeled concept representations in semantic space². Concluding (5.), two or three problem areas connected with word meaning and concept processing will be touched which might be tackled anew and perhaps be brought to a more adequate though still tentative solution under an empirically founded approach in procedural semantics.

2 Representational formats in knowledge systems

Lexical structures in linguistic semantics, memory models in cognitive psychology, and semantic networks in AI-research have in common that they use as basic format of their models some structure of directed graphs. Probably one of the most familiar forms of concept representation which experimental psychologists like e.g. [1] and [2] have set up and tested in the course of their developments of memory models is shown in Fig. 2.1

Here we have a hierarchy of labeled concept nodes with predicates and properties linked to them which are herited by directly dependent nodes. The hypotheses formulated and tested in experiments predict that test persons will take more time to identify and decide given propositions with an increasing number of node- and level-transitions to be processed in the course of interpretation. Evaluating a sentence like "A canary can sing" will take less time than to decide whether the sentence "A robin can breathe" is true or not. Thus, reaction-time serves as an indicator for the proposed model structure either to be correct or in need of modification.

In early artificial intelligence research a different type of knowledge representation was developed for question-answering-systems. A fragment of the most common schema of the semantic network type [3] is shown in Fig. 2.2. Here again we have labeled concept nodes linked to one another by pointers representing labeled rela-

²The system of both, the text analysing algorithm leading to the semantic space structure and the generative procedure operating on that structure to yield the DDS-trees, is implemented in FORTRAN, CDC-ASSEMBLER, and SIMULA on the CDC-Cyber 175 of the Technical University of Aachen Computing Center.

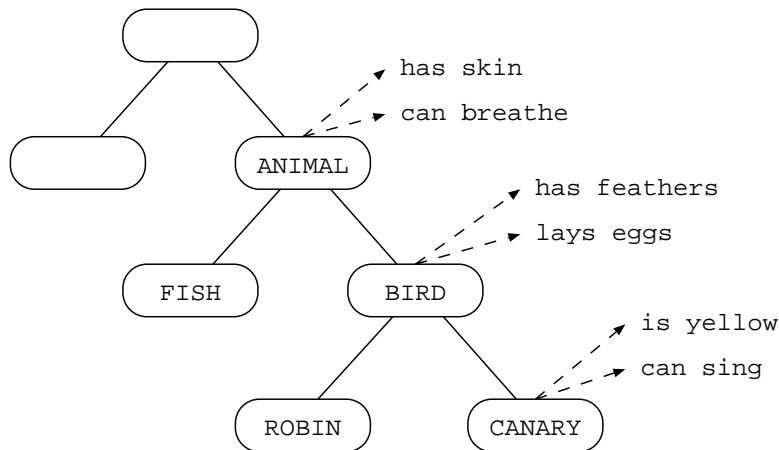


Figure 2.1: A graph model of concept representation.

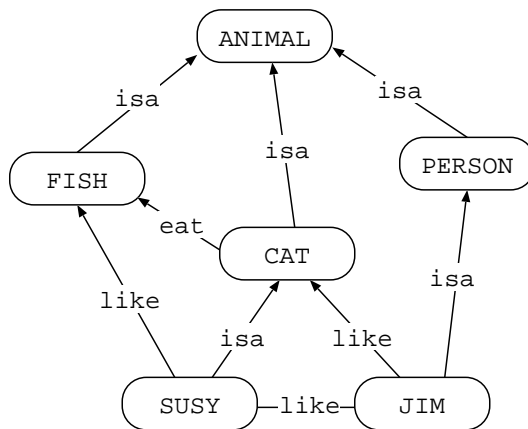


Figure 2.2: A semantic network for knowledge representation.

tions which form a network instead of a tree structure. This enables the system to answer questions like: "Is Susy a cat?" correctly by identifying the SUSY-node, its ISA-relation pointer and the CAT-node. Moreover, the pointer structure allows for the processing of paths laid through the network, initiated by questions like: "Susy, cat?" which will prompt the answer "Susy is a cat. Cat eats fish. Cat is an animal. Fish is an animal."

A schematic representation of concept relatedness as envisaged by cognitive theorists who work along more procedural lines of memory models [4] is shown in Fig. 2.3. Their distance-relational conception lends itself readily to the notion of stereotype representation for concepts that do not have intersubjectively identifiable sharp boundaries [5].

Instead of binarily decidable category membership, stereotypical concepts or prototypes are determined by way of their adjacency to other prototypes. Taken as

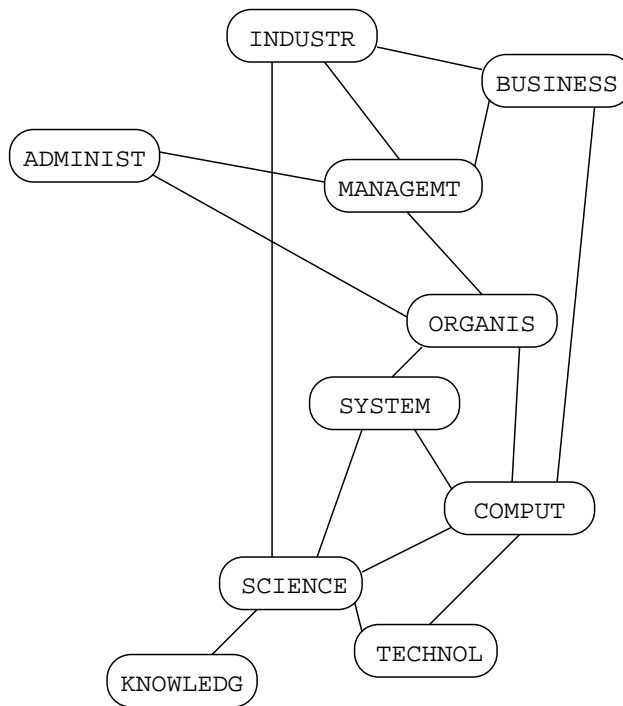


Figure 2.3: A schematic representation of concept relatedness.

a memory model, stimulation of a concept will initiate *spreading activation* to prime the more adjacent concepts more intensely than those farther away in the network structure, thus determining a realm of concepts related by their primed semantic affinity. In the given example, the stimulation of the concept-node MANAGEMENT will activate that of BUSINESS first, then INDUSTRY and ORGANISATION with about the same intensities, then ADMINISTRATION and so on, with the intensities decreasing as a function of the activated nodes' distances.

These three schemata of model structures — although obviously concerned with the simulation of symbol understanding processes — are designed to deal primarily with static aspects of meaning and knowledge. Thus, in interpreting input symbols/strings, pre-defined/stored meaning relations and constructions can be identified and their representations be retrieved. Without respective grounding made explicit and represented in that structure, however, possibly distorted or modified instantiations of such relations or relevant supplementary semantic information can hardly be recognized or be provided within such representational systems. As the necessary data is not taken from natural language discourse in communicative environments but elicited in experimental settings by either exploring one's own or the test persons' linguistically relevant cognitive and/or semantic capacities, usage similarities of different and/or contextual variations of identical items are difficult to be ascertained. This is rather unsatisfactory from a linguist's point-of-view who thinks that his discipline is an empirical one and, hence, that *descriptive semantics* ought

to be based upon linguistic data produced by real speaker/hearers in factual acts of communicative performance in order to let new meaning representations (or fragments of them) replace (or improve) older ones to change/update a static memory structure.

3 Statistical tools for discourse analysis

It has been shown elsewhere³ sufficiently large sample of pragmatically homogeneous texts, called *corpus*, 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 lexical items 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 *semantic regularities* which may be detected empirically from the texts.

The empirical analysis of discourse and the formal representation of vague word meanings in natural language texts as a system of interrelated concepts is based on the WITTGENSTEINIAN [8] notion of *language games* and their *functions*⁴. 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.

The statistics which have been used so far for the systematic analysis not of propositional strings but of their elements, namely words in natural language texts, is basically descriptive. Developed from and centred around a correlational measure to specify intensities of co-occurring lexical items used in natural language discourse, 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 correlation coefficient appropriately modified for the purpose has been used as a mapping function. It allows to compute the relational interdependence of any two lexical items from their textual frequencies. Those items which co-occur frequently in 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*. Different degrees of *word-*

³See also [7] where the principle of semantization is introduced as a process which can be emulated by procedural means to constitute meanings by consecutive restrictions of elementary choices among entities on the levels of pragmatics, via semantics and syntactics down to morphophonetics; whereas these elements and/or entities on each of the semiotic levels are generated by an inversely operating procedure, which allows recurrent combinations of elements to be identified against those combinatorial possibilities not realized on that level these combinations constitute the new elements which on the next level may be combined, etc.

⁴"A meaning of a word is a kind of employment of it. For it is what we learn when the word is incorporated into our language. That is why there exists a correspondence between the concept *rule* and *meaning*. [...] Compare the meaning of a word with the *function* of an official. And *different meanings* with *different functions*. When language games change, then there is a change in concepts, and with the concepts the meanings of words change." [8] No. 61–65, p. 10e

repugnancy and *word-affinity* — indicated by numerical values ranging from -1 to $+1$ — may thus be ascertained without recurring to an investigator's or his test-persons' word and/or world knowledge (semantic competence), but can instead solely be based upon the usage regularities of lexical items observed in a corpus of pragmatically homogeneous texts, spoken or written by real speakers/hearers in actual or intended acts of communication (communicative performance).

Let K be such a corpus that consists of t texts belonging to a specific *language-game*, i.e. satisfying the condition of pragmatic homogeneity, and let V be the vocabulary of i lexical entries x being used

$$K := \{t \mid t = 1, \dots, T\} \quad \text{and} \quad V := \{x_i \mid i = 1, \dots, n\}$$

with U being the overall length of all texts t in K

$$U_k = \sum_{t=1}^T u_t, \quad 1_t \leq u_t \leq U$$

and H_i the total frequency of the lexical entry x_i in K

$$H_i = \sum_{t=1}^T h_{it}, \quad 1_{it} \leq u_{it} \leq H_i$$

Then the modified correlation coefficient will read

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

where $h_{it}^* = \frac{H_i}{U} u_t$ and $h_{jt}^* = \frac{H_j}{U} u_t$

For the sake of illustrating the analysing algorithm's performance, we will consider a simplified case where the vocabulary V employed in the texts shall be limited to only three word-types, namely x_i , x_j and x_k which have a certain overall token-frequency. Then the modified correlation coefficient will measure the regularities of usage by the *affinities* and *repugnancies* that may hold between anyone lexical item and all the others employed in the discourse analysed. That will yield for any item an n -tuple of correlation-values, in this case for the lexical item x_i with $n = 3$ the triple of values α_{ii} , α_{ij} , α_{ik} . These correlation-values are now interpreted as being coordinates that will define for each lexical item x_i , x_j , and x_k one point $y(\alpha_i)$, $y(\alpha_j)$, and $y(\alpha_k)$ respectively in a three-dimensional space structure spanned by the three axis i , j , and k as illustrated in Fig. 3.1. As the positions of these points now obviously depend on the regularities the lexical items concerned have been used with in the texts of the corpus, the y -points are called *corpus-points* of i , j and k in the α - or *corpus-space*.

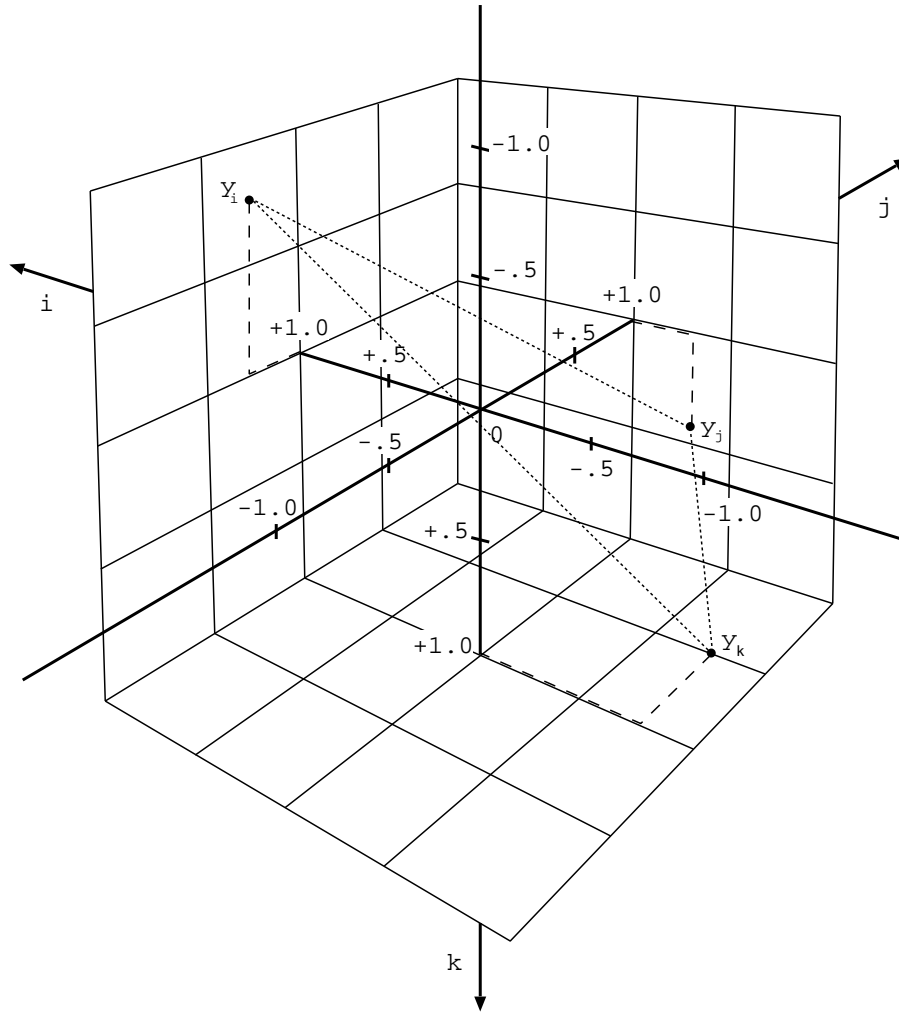


Figure 3.1: Representation of corpus-points in a corpus-space.

Two y -points in this space will consequently be the more adjacent to each other, the less their usages differ. These differences may be calculated by a distance measure δ between any two y -points

$$\delta(y_i, y_j) = \left(\sum_{k=1}^n (\alpha(x_i, x_k) - \alpha(x_j, x_k))^2 \right)^{1/2}, \quad 0 \leq \delta_{ij} \leq 2\sqrt{n}$$

as illustrated in Fig. 3.1 by dotted lines. The distance-values are real, non-negative numbers which represent a new characteristic. For any item $y_i, y_j,$ and y_k an n -tuple of δ -values, i.e. for y_i the triple $\delta_{ii}, \delta_{ij}, \delta_{ik}$ is obtained which may be interpreted as new coordinates. These will again for each item $x_i, x_j,$ and x_k define new points $z(\delta_i), z(\delta_j),$ and $z(\delta_k)$ in a new n -dimensional space, called semantic space, as illustrated in Fig. 3.2. The positions of such points in the semantic space will clearly depend on all the differences (δ - or distance-values) in all the regularities of usage (α - or

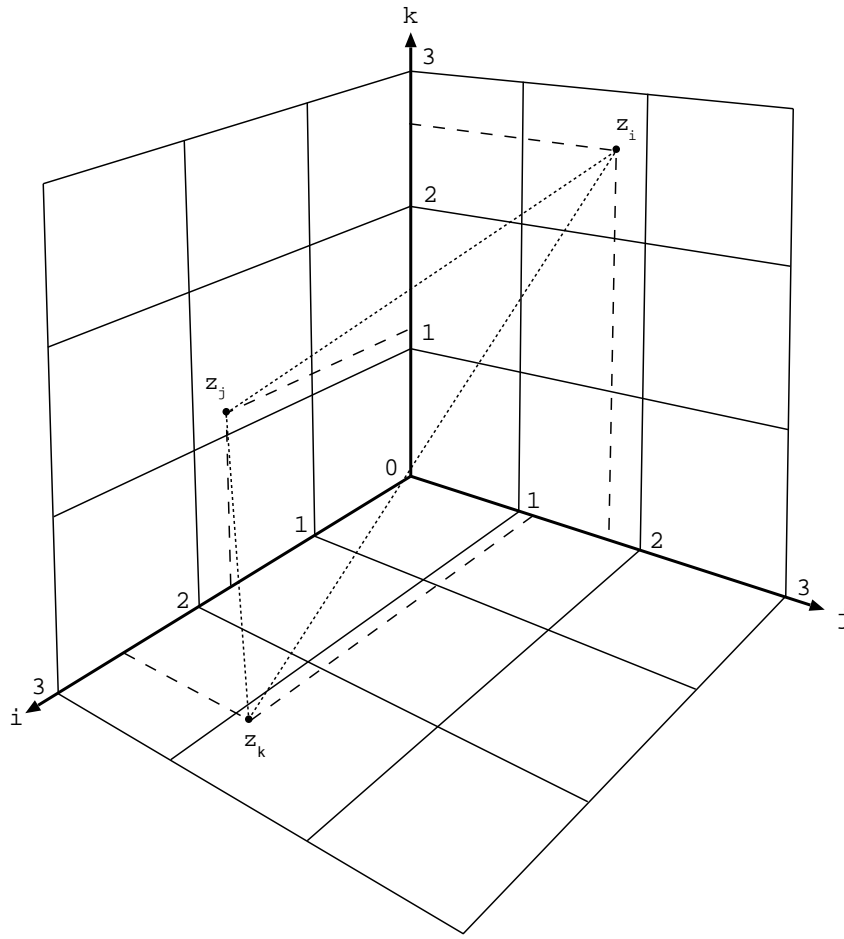


Figure 3.2: Placement of points of Figure 3.1 in semantic space.

correlation-values) any lexical item shows in the texts analysed. Thus, each lexical item is mapped onto a fuzzy subset of the vocabulary according to the numerically specified regularities these items have been used with in the discourse analysed. Measuring the differences of any one's lexical item's usage regularities against those of all others allows for the above interpretation and consecutive mappings of items onto theoretical constructs.

These new entities are abstract representations of what *meanings* may be composed of, i.e. a number of operationally defined elements whose varying contributions are to be derived directly from the differing usage regularities that the corresponding lexical items produce in the texts analysed. As theoretical constructs, these entities constitute *meaning* from a more holistic approach to lexical system description. Translating the WITTGENSTEINIAN notion of *meaning* into a mathematically operational form of empirical feasibility, these new *meaning*-components can procedurally be characterized as a function of all the differences of all regularities any one of the vocabulary's items is used with compared to any other item in the same corpus of

discourse.

The resulting system of sets of fuzzy subsets of the vocabulary represent a *structured lexicon*. It is a relational data structure which may be interpreted topologically as a hyperspace with a natural metric, called *semantic space*. Its linguistically labelled elements represent *meaning points*, and their mutual distances represent *meaning differences*. The position of a meaning point may be described by its semantic environment. This is determined by those other points in the semantic hyperspace which — within a given diameter — are most adjacent to the central one chosen to be illustrated according to the following Euclidean metric

$$\delta_2(z_i, z_j) = \left(\sum_{k=1}^n (\delta(y_i, y_k) - \delta(y_j, y_k))^2 \right)^{1/2}, \quad 0 \leq \delta_{2ij} \leq 2\sqrt{n}$$

Figure 3.3 shows the topological environment $E\langle\text{GESCHAEFT}\rangle$, i.e. those points being situated within the hypersphere of a certain diameter of the meaning point GESCHAEFT/business as computed from a corpus of German newspaper texts comprising some 8000 tokens of 360 types in 175 texts from the 1964 editions of the daily DIE WELT [9].

Having seen that topological environments of that sort do in fact assemble meaning points of a certain semantic affinity solely by the performance of the text analysing algorithms and without any competent language user's interference, a number of questions arose whose answers should at least be mentioned:

- Are there regions of *point density* in the semantic space, forming clouds and clusters which might indicate a semantic (syntagmatic and/or paradigmatic) structuredness?
- Can such regions be detected and described automatically by statistical methods of multi-varied and cluster analysis, and how would they look like?
- Could the internal relation according to which certain meaning points cluster be specified in terms of the logical-declarative vs. analogical-associative opposition of semantic relatedness?

Having checked a great number of environments, it was ascertained that they do in fact assemble meaning points of a certain semantic affinity. Further investigation revealed [9] that there are regions of higher point density in the semantic space, forming clouds and clusters. These were detected by multivariate and cluster-analyzing methods [10] which showed, however, that both, the paradigmatically and syntagmatically related items formed what may be named *connotative clouds* rather than what is known to be called *semantic fields* [11]. Although its internal relations appeared to be unspecifiable in terms of any logically deductive or concept hierarchical system, their elements' positions revealed a high degree of stable structures which suggested a regular form of contents-dependant associative connectedness [12] which gave rise to the idea of having the variable *relevance* of related meanings and/or concepts be defined procedurally [13], [14], [15].

<u>GESCHAEFT</u> /business	0.000		
WERB/advertism	2.837	KENNTNIS/knowledge	3.028
BITTE/request	3.284	TECHNIK/technic	3.527
PERSON/person	3.930	BUCH/book	4.232
FAEHIG/capable	4.471	ORGANISAT/org	4.526
INFORMAT/info	4.568	ERFAHR/experience	4.708
ALLGEMEIN/general	4.816	BRITAIN/britain	4.838
KONTAKT/contak	4.902	UNTERRICHT/instruc	4.919
ANGEBOT/offer	5.047	AUSGAB/expens	5.064
RAUM/space	5.098	DIPLOM/diploma	5.155
VERBAND/assocn	5.183	COMPUTER/compute	5.212
STADT/city	5.216	ELEKTRON/electr	5.311
LEHR/teach	5.321	LEIT/lead, control	5.404
WEG/way	5.464	STELLE/position	5.498
WIRTSCHAFT/econom	5.503	MODE/fashion	5.537
JOURNAL/journal	5.621	BILDUNG/education	5.657
GEBIET/area	5.697	SUCH/search	5.733
SYSTEM/system	5.752	EINSATZ/activity	5.813
ARBEIT/labour	5.834	AUFTRAG/order	5.872
WUNSCH/wisch	5.880	PROGRAMM/program	5.880
AUSLAND/abroad	5.881	INDUSTRIE/industry	5.909
...		...	

Figure 3.3: The topological environment of $E\langle\text{GESCHAEFT}\rangle$.

4 Representation of semantic disposition

Following a more semiotic understanding of meaning constitution, the present semantic space model may be considered the core structure of a word meaning/world knowledge representation system which separates the format of a basic (stereotype) meaning representation from its latent (dependency) relational organization. Whereas the former is a rather static, topologically structured (associative) memory representing the data that text analysing algorithms provide, the latter can be characterized as a collection of dynamic and flexible structuring processes to re-organize these data under various principles. Other than declarative knowledge that can be represented in pre-defined semantic network structures, meaning relations of lexical relevance and semantic dispositions which are heavily dependent on context and domain of knowledge concerned will more adequately be defined procedurally, i.e. by generative algorithms that induce them on changing data only and whenever necessary. This is achieved by a recursively defined procedure that produces hierarchies of meaning points, structured under given aspects according to and in dependence of their meanings' relevancy.

Taking up the heuristics provided by *Spreading Activation Theory* in semantic

memory, cognitive structures, and concept representation as advanced by [16], [17], and [18], the notion of spreading activation can be employed not only to denote activation of related concepts in the process of *priming* studied in subsequent publications like [19] and [20] but — generically somewhat prior to that — may also signify the very procedure which induces these relations between concepts. Originally developed as a procedural model to cope with observed latencies of activated concepts in comprehension processes, *priming* and *spreading activation* is based on network-type models or world-knowledge structures as illustrated briefly above. Essentially defined by nodes, representing concepts, meanings or objects, and pointers which relate them conceptually, semantically, or logically to one another, these formats have a considerable advantage over the semantic space structure outlined above: one of the problems of distance-like data structures in semantic processing is that — *distance* being a symmetric relation — well-known search strategies for retrieval, matching, and inferencing purposes cannot be applied because these are based upon some non-symmetric relations, as realized by pointer structures in well-known word meaning and/or world knowledge representations.

In order to make such procedures operate on the semantic space data, its structure has to be transformed into some hierarchical organisation of its elements. For this purpose, the semantic space model has to be re-interpreted as a sort of conceptual raw data and associative base structure. What appeared to be a disadvantage first, now turns out to be an advantage over more traditional formats of representation. Other than these approaches which have to presuppose the structural format of the semantic memory models that are to be tested in word recall and/or concept recognition experiments, the semantic space provides some of the necessary data for the procedural definition of dynamic, instead of static model structures that allow variable stereotype instead of fixed categorial concept representations. Thus, the concept nodes as abstract mappings of meanings of lexical items are not just linked to one another according to what cognitive scientists supposedly know about the way conceptual information is structured in memory, but it is this very structure that is already considered to by a dynamic format of stereotype concept organization. Defined as procedures that operate on the semantic space data, this is tantamount to a dynamic re-structuring of meaning points and — depending on the controlling parameters — the generation of paths between them along which — in case of priming — activation might spread whenever a meaning point is stimulated.

Unlike the ready-set and fixed relations among nodes, an algorithm has been devised which operates on the semantic space data structure as its base to induce dependencies between its elements, i.e. among subsets of the meaning points. The recursively defined procedure detects fragments of the semantic space according to the meaning point it is started with and according to the semantic similarities, i.e. the distance relations it encounters during operation, constituting what we termed semantic *relevance*. Stop-conditions may deliberately be formulated either qualitatively (naming a target point) or quantitatively (number of points to be processed).

Given one meaning point's position as a start, the algorithm will — other than in [10] and [11] — *first* list all its neighbouring points by increasing distances, *second*

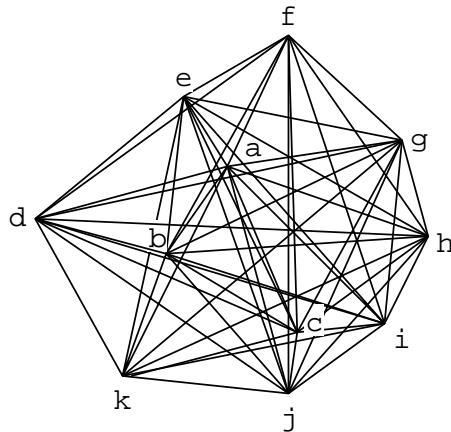


Figure 4.1: A distance-like configuration of eleven sample points a to k .

provide similar lists for each of these neighbours, and *third* prime the starting point as dominant node to mark the tree's root. Then, the algorithm's generic procedure will take the first entry from the first list, determine from the appropriate second list its most adjacent neighbour among those points already primed, in order to identify it as the ancestor (mother-node) to which the new descendant (daughter-node) is linked whose label then gets deleted from the first list. Repeated successively for each of the meaning points listed and in turn primed in accordance with this procedure, the algorithm will select a particular fragment of the relational structure latently inherent in the semantic space data under a certain perspective, i.e. the aspect or initially primed meaning point the algorithm is started with. Working its way through and consuming all labeled points in the space structure — unless stopped under conditions of given target points, number of points to be processed, or threshold of maximal distance — the algorithm transforms prevailing similarities of meanings as represented by adjacent points to establish — in the process of priming — a binary, non-symmetric, and transitive relation between them. This relation allows for the hierarchical re-organization of meaning points as descendant nodes under a primed head or root in an n -ary DDS-tree [12]. Weighted numerically as a function of a node's distance values and level of its tree-position, this measure either expresses a concept's dependencies as given by the root's descendants in that tree, or, inversely, it evaluates their criterialities for that concept as specified and determined by that tree's root.

Without introducing the algorithms formally, some of their operative characteristics can well be illustrated in the sequel by a few simplified examples. Beginning with the schema of a distance-like data structure as shown in the two-dimensional configuration of 11 points, labeled a to k (Fig. 4.1) the stimulation of three different starting points a , b and c results in the dependency structures which the algorithm of least distance selects (Fig. 4.2) as distance detection (first row), as a step-list representation of the selecting process of points activated (second row), then as

their n -ary tree representations (third row) and finally as their transformations to binary-tree structures (fourth row) of points related respectively to be primed.

It is apparent that stimulation of other points within the same configuration of basic data points will result in similar but nevertheless differing trees, depending on the aspect under which the structure is accessed, i.e. the point initially stimulated to start the algorithm with.

Applied to the semantic space data of 360 defined meaning points calculated from the textcorpus of the 1964 editions of the German newspaper DIE WELT, the *Dispositional Dependency Structures (DDS)* of AUFTRAG/order and GESCHAEFT/business are given in Figs. 4.3 and 4.4 as generated by the procedure described. Different stop conditions given for the generation of the DDS resulted in different trees: DDS⟨AUFTRAG⟩ qualitative stop by target node GESCHAEFT, grade 7, depth 13, 64 nodes; and DDS⟨GESCHAEFT⟩ quantitative stop by number of nodes to be processed, grade 4, depth 10, 60 nodes. In the DDS⟨AUFTRAG⟩ (Fig. 4.3) we find only one descendant (LEIT/lead) on level 1, three as connotative alternates on level 2, one of which (ELEKTRON/electronic) has even 7 descendants on level 3, etc. In the DDS⟨GESCHAEFT⟩ (Fig. 4.4) there are two descendant connotative alternates (WERB/advertisment; KENNTNIS/knowledge) on level 2, each of which has four descendants on level 3, etc. Attention is drawn to the dependencies of the direct descendants (BITTE/request) → (PERSON/person) → (HAUS/house). As in DDS⟨AUFTRAG⟩ this dependency is found in exactly the same order in the DDS⟨GESCHAEFT⟩ but here it is situated farther from the root, starting on the tree's sixth level only, instead of its third.

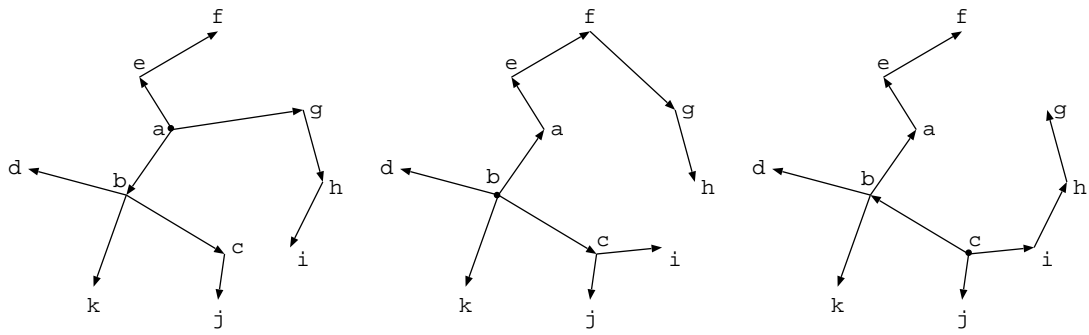
To calculate such differences, a numerical measure of *criteriality* Cr_i to its mother-node z_d under a given aspect i can be defined as a function of its distance value $\delta_2(z_d, z_a)$, the tree's root z_r , and its level g concerned.

$$Cr_i(z_d)_{g+1} = Cr_i(z_a)_g \exp\left(\frac{\delta_2(z_d, z_a)}{\delta_2(z_d, z_r) + 1}\right)$$

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 of certain meaning points activated.

5 Conclusion

It goes without saying that generating DDS-trees is a prerequisite to source-oriented, contents-driven *search* and *retrieval* procedures which may thus be performed effectively on the semantic space structure. Given the meaning point AUFTRAG/order being stimulated, and GESCHAEFT/business as the target point to be searched for, then, the DDS⟨AUFTRAG⟩ will be generated as illustrated above, providing with decreasing criterialities the range of *semantic dispositions* inherent in the semantic space data under the aspect of, and triggered by the priming of AUFTRAG/order. The tree generating process being stopped after hitting and incorporating as its last



Step	z_d	z_a
0	a	a
1	e	a
2	b	a
3	c	b
4	f	e
5	g	a
6	d	b
7	h	g
8	i	h
9	k	b
10	j	c

Step	z_d	z_a
0	b	b
1	a	b
2	k	b
3	d	b
4	e	a
5	c	b
6	j	c
7	i	c
8	f	e
9	g	f
10	h	j

Step	z_d	z_a
0	c	c
1	j	c
2	i	c
3	b	c
4	h	i
5	k	b
6	a	b
7	g	h
8	d	b
9	e	a
10	f	e

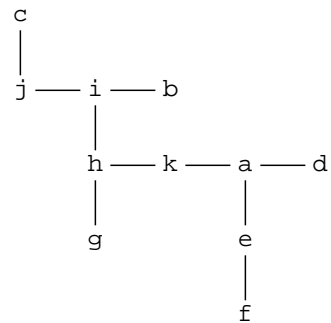
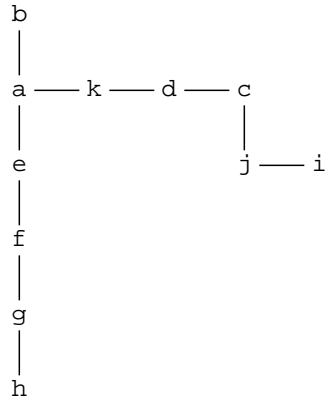
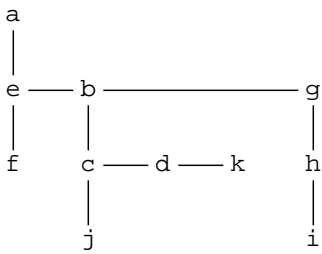
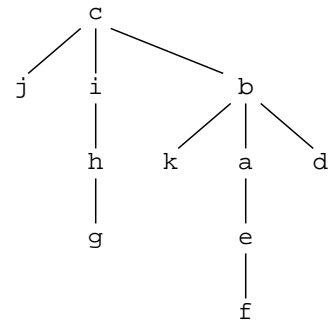
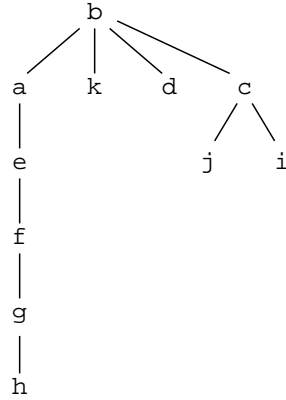
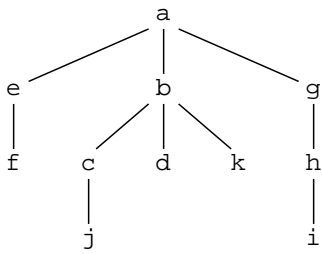


Figure 4.2: Dependency structures resulting from stimulation of three different starting points a , b , and c in Figure 4.1.

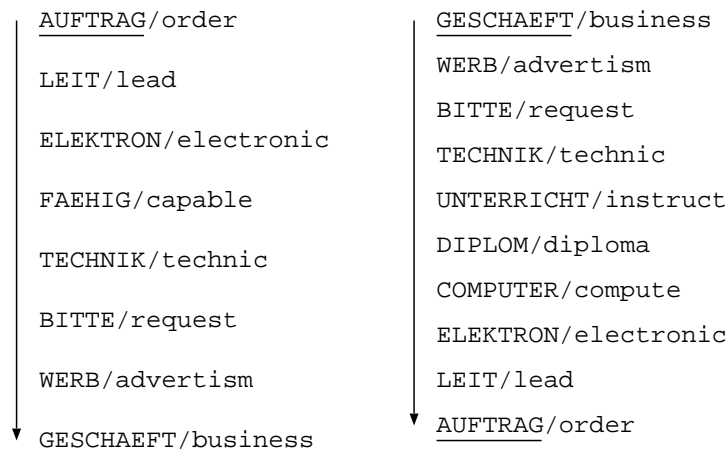


Figure 5.1: Dependency path from AUFTRAG (= order) to GESCHAEFT (= business) and vice versa.

node the target item, its *dependency path* will be activated. This is to trace those intermediate nodes which determine the associative transitions of any target node under any specifiable aspect. Looking up GESCHAEFT/business as a target node under the aspect of AUFTRAG/order its *dependency path* (in Fig. 4.3 above, and given separately in Fig. 5.1 below) consists of WERBUNG/advertise, BITTE/request and TECHNIK/technic, FAEHIG/capable, ELEKTRON/electronic, LEIT/lead which — not surprisingly though — proves to be approximately the dependency path of AUFTRAG/order under the aspect of GESCHAEFT/business but in inverted order and FAEHIG/capable replaced by COMPUTER/compute, DIPLOM/diploma, and UNTERRICHT/instruct.

Using source-oriented search and retrieval processes as described, an *analogical*, contents-driven form of inference — as opposed to *logical deduction* — may operationally be devised by way of parallel processing of two (or more) dependency-trees. For this purpose the algorithms are started by the two (or more) meaning points considered to represent the *premises*, of say, AUFTRAG/order and GESCHAEFT/business. Their DDS-trees will be generated before the inferencing procedure begins to work its way (breadth-first or depth-first) through both (or more) trees, tagging each encountered node. When in either tree the first node is met that has previously been tagged by activation from another priming source, the search procedure stops to activate the dependency paths from this *concluding* common node — in our case FAEHIG/capable for breadth-first and DIPLOM/diploma for depth-first searches — in the DDS-trees concerned and separately presented in Figs. 5.2 and 5.3.

To conclude with, some extrapolating ideas of possible applications and/or new views of older problems might be in order. It appears that the DDS-procedure provides a flexible, source-oriented, contents-driven method for the *multi-perspective*

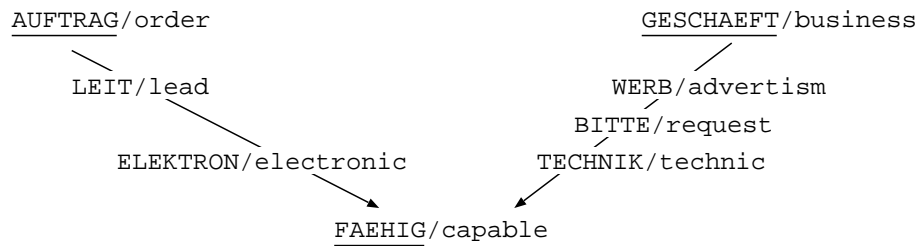


Figure 5.2: Inference paths from AUFTRAG and GESCHAEFT activating FAEHIG (= capable) in breadth-first searching mode.

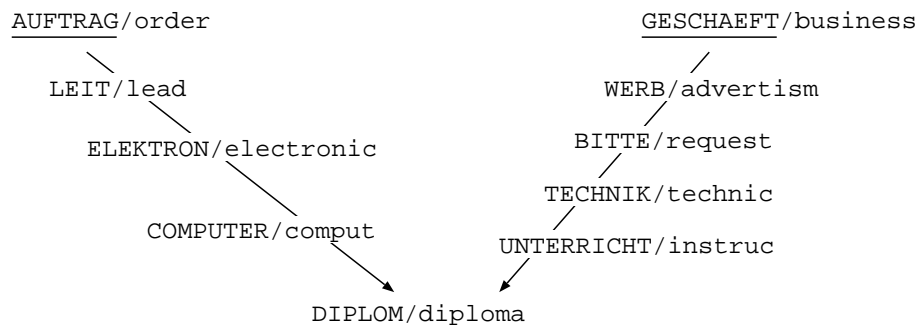


Figure 5.3: Inference paths from AUFTRAG and GESCHAEFT activating DIPLOM (= diploma) in depth-first searching mode.

induction of a relevance relation among stereotypically represented concepts which are linguistically conveyed by natural language discourse on specified subject domains.

- Applied to any distance-like data structures of knowledge and/or meaning representation systems, the DDS-procedure allows for the generation of possible paths of spreading activation which branch across *semantic space*, submitting relevant portions of it to associatively guided search strategies and retrieval operations.
- Replacing the storage of fixed and ready-set networks by a contents-driven induction of relevance related nodes, the problem of identifying stored meaning constructions with distorted instantiations of them, can be circumvented. Triggered by any identifiable label, the DDS will be generated according to the database provided and the resultant tree-structure will therefore vary according to the possibly varying status of the data in the semantic space structure.
- In view of tacit knowledge and implied information the DDS-procedure offers an empirically based approach and a dynamic representation of *semantic*

dispositions which — in language understanding systems — might serve as connotative default values in identifying and/or interpreting input labels and solving ambiguity and/or vagueness problems of input strings.

- Among other extensions, it is hoped to develop a numerical expression for measuring the amount of meaning conveyed by any string interpreted by DDS-processing. Other than in classical information theory where the amount of information is a function of probabilities of signs calculated from empirical distributions of sets of symbols that have to be finite, the amount of meaning will have to be based upon a numerical measure similar to the *criterialities* calculated as a function of structural properties of open sets and dynamically organized systems of symbols instead.

References

- [1] COLLINS, A.M./QUILLIAN, M.R. (1969): Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* 8 (1969) 240–247
- [2] KLIX, F. (1976): Strukturelle und funktionelle Komponenten des menschlichen Gedächtnisses. In: Klix, F. (Ed.): *Psychologische Beiträge zur Analyse kognitiver Prozesse*. Berlin (Akademie Verlag) 1976, 57–98
- [3] WINOGRAD, T. (1975): Frame Representation and the Declarative/Procedural Controversy. In: Bobrow, D.G./Collins, A. (Eds): *Representation and Understanding. Studies in Cognitive Science*. New York/San Francisco/London (Academic Pr.) 1975, 185–210
- [4] COLLINS, A.M./LOFTUS, E.F. (1975): A spreading activation theory of semantic processing. *Psychological Review* 6 (1975) 407–428
- [5] ROSCH, E. (1975): Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* 104 (1975) 192–233
- [6] RIEGER, B.B. (1981): Feasible Fuzzy Semantics. In: Eikmeyer, H.J./Rieser, H. (Eds): *Words, Worlds, and Contexts. New Approaches in Word Semantics*. Berlin/New York (de Gruyter) 1981, 193–209
- [7] RIEGER, B.B. (1977): Bedeutungskonstitution. Bemerkungen zur semiotischen Problematik eines linguistischen Problems, *Zeitschrift für Linguistik und Literaturwissenschaft* 27/28 (1977) 55–68
- [8] WITTGENSTEIN, L. (1969): *Über Gewißheit — On Certainty*. New York/San Francisco/London (Harper & Row)
- [9] RIEGER, B.B. (1983): Clusters in Semantic Space. In: Delatte, L. (Ed.): *Actes du Congrès International Informatique et Sciences Humaines, Liéges (LASLA)* 1983, 805–81

- [10] RIEGER, B.B. (1981): Connotative Dependency Structures in Semantic Space. In: Rieger, B. (Ed.): *Empirical Semantics. A collection of new approaches in the field*, Vol. II. Bochum (Brockmeyer) 1981, 622–711
- [11] RIEGER, B.B. (1982): Procedural Meaning Representation. In: Horecky, J. (Ed.): *COLING 82. Proceedings of the 9th Intern. Conference on Computational Linguistics*, Amsterdam/New York (North Holland) 1982, 319–324
- [12] RIEGER, B.B. (1984): Inducing a relevance relation in a distance-like data structure of fuzzy word meaning representation. In: Allen, R.F. (Ed.): *Proceedings of the 4th Intern. Conf. on Databases in the Humanities and Social Sciences (ICDBHSS/83)*. New Brunswick (Rutgers U.P.) 1984 (in print)
- [13] RIEGER, B.B. (1984): Lexical Relevance and Semantic Disposition. On stereotype word meaning representation in procedural semantics. In: Hoppenbrouwes/Seuren/Weijters (Eds): *Meaning and the Lexicon. Proceedings of the 2nd Intern. Colloquium on the Interdisciplinary Study of the Semantics of Natural Language*, Nijmegen (N.I.S. Press) 1984 (in print)
- [14] RIEGER, B.B. (1984): Semantic Relevance and Aspect Dependency in a Given Subject Domain. In: Walker, D.E. (Ed.): *COLING 84 — Proceedings of the 10th Intern. Conference on Computational Linguistics*, Stanford (Stanford U.P.) 1984, 298–301
- [15] RIEGER, B.B. (1984): Semantische Dispositionen. Prozedurale Wissensstrukturen mit stereotypisch repräsentierten Wortbedeutungen. In: Rieger, B. (Ed.): *Dynamik in der Bedeutungskonstitution*. Hamburg (Buske) 1984 (in print)
- [16] QUILLIAN, M.R. (1968): Semantic Memory. (unpubl. Carnegie Inst. of Technology doctoral dissert. 1966) in part in: Minsky, M. (Ed.): *Semantic Information Processing*. Cambridge, Mass. (MIT Press) 1968, 216–270
- [17] OLSON, D.R. (1970): Language and thought: aspects of a cognitive theory of semantics. *Psychological Review* 77, 4 (1970) 257–273
- [18] SWINNEY, D.A. (1979): Lexical Processing during Sentence Comprehension, *Journal of Verbal Learning and Verbal Behavior* 18 (1979) 733–743
- [19] 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 (1982) 468–492
- [20] FLORES D’ARCAIS, G.B./JARVELLA, C. (Eds)(1983): *The Progress of Language Understanding*. New York/Sydney/Toronto (Wiley Sons)
- [21] ZADEH, L.A. (1965): Fuzzy sets. *Information and Control* 8 (1965) 338–353