Procedural Meaning Representation by Connotative Dependency Structures
An Empirical Approach to Word Semantics for Analogical Inferencing[*]

Burghard B. Rieger

MESY-Group, German Department

Technical University of Aachen

Germany

Natural language understanding systems make use of language and/or world knowledge bases. One of the salient problems of meaning representation and knowledge structure is the modelling of its acquisition and modification from natural language processing. Based upon the statistical analysis of discourse, a formal representation of vague word meanings is derived which constitutes the lexical structure of the vocabulary employed in the texts as a fragment of the connotative knowledge conveyed in discourse. It consists of a distance-like data structure of linguistically labeled space points whose positions give a prototype-representation of conceptual meanings. On the basis of these semantic space data an algorithm is presented which transforms prevailing similarities of conceptual meanings as denoted by adjacent space points to establish a binary, non-symmetric, and transitive relation between them. This allows for the hierarchical reorganization of points as nodes dependent on a head in a binary tree called connotative dependency structure (CDS). It offers an empirically founded operational approach to determine relevant portions of the space structure constituting semantic dispositions which the priming of a meaning point will trigger with decreasing criteriality. Thus, the CDS allows for the execution of associatively guided search strategies, contents-oriented retrieval operations, and source-dependent processes of analogical inferencing.

Introduction

In procedural approaches of linguistic semantics, cognitive psychology and artificial intelligence, natural language understanding systems make use of language and/or world knowledge bases. Defined as lexical structures, memory models or semantic networks, they are formatted according to whatever representational, explanatory or inferential purpose a particular simulation of processes and/or of understanding was

aiming at [1]. The language and world knowledge embodied in these systems, however, is restricted under two aspects: most of it is obtained introspectively and as such not warranted by any operational means or, whenever it seems to, these operations are not the permitting condition for, but a performing result of simple referencing in clear-cut environments.

Based mainly upon the investigators' or the system designers' own or some consulted experts' linguistic competence and/or world knowledge in a subject domain, the data considered semantically relevant to be organized in referential and/or conceptual structures (lists, arrays, networks, topologies, etc.) have a more or less ad hoc character and are confined to representing logically reconstructable propositions. Neglectable as these shortcomings prove to be for strictly extensionally defined environments and fragments of knowledge structure in referential models, data complexity tends to increase to meet exploding difficulties and escalating problems whenever abstract concepts or even vague meanings are to be processed in a not exclusively denotative but also connotative setting of formal semantic representation.

As natural language communication may be characterized by the apparent ease and efficiency however, with which ill-defined concepts and fuzzy meanings are being intended and expressed by speakers, identified and understood by hearers, and successfully used by speakers/hearers in performing inferences of some — not necessarily logical — sort, it is argued here, that any non-trivial simulation of processes of cognition and/or natural language comprehension will have to provide some means of dynamic knowledge representation which permits to account more satisfactorily for one or the other aspect raised above.

> The concept of 'representation of knowledge' seems lucid enough when talking about memories of sentences, numbers, or even faces, for one can imagine how to formulate these in terms of propositions, frames, or semantic networks. But it is much harder to do this for feelings, insights and understandings, with all the attitudes, dispositions, and 'ways of seeing things' that go with them. (The term 'disposition' is used here in its ordinary language sense to mean 'a momentary range of possible behaviours'!) Traditionally, such issues are put aside, with the excuse that we should understand simpler things first. But what if feelings and view points a r e the simpler things — the elements of which the others are composed? Then, I assert, we should deal with dispositions directly, using a 'structural' approach [...]   [2]

In the present case this has been developed in two stages: the semantic space as a distance-like data structure, and an algorithm to transform its distance relations to form source-oriented hierarchies of connotative dependency structures.

Semantic Space Structure

Theoretical approaches in formal semantics tend to deny a dynamic linguistic meaning structure, but assume the existence of an external system structure of a world, or possible worlds, whose pre-formatted entities may referentially be related to language terms constituting their denotation. Structural approaches in linguistic semantics tend to deny the possibility of denotational, but presuppose the knowledge (and comprehension) of language systems whose semantic relations among their items are being described intra-lingually by means of syntagmatic and paradigmatic oppositions along

certain dimensions in semantic fields. Other than these two, our present way of approach strives to presuppose as little and to reconstruct empirically as much as possible of the relational (not necessarily logically reconstructable) structure that in the course of discourse is constituted by the regular use of language terms as a system of linguistically labeled empirical objects, called meanings.

We consider the natural language users' ability to intend and comprehend meanings in verbal interaction a phenomenologically undoubtable, empirically well established, and theoretically defensible basis for any semantic study of natural language performance. It is assumed that the usage regularities followed and/or established by employing different lexical items differently for communicative purposes in discourse may be analysed not only to describe the lexical structure of vocabulary items used, but also to model a fragment of the concomitantly conveyed common knowledge or semantic memory structure constituted.

This is achieved by an algorithm that takes lemmatized strings of natural language discourse of a certain domain as input and produces as output a distance-like data structure of linguistically labeled points whose positions represent their meanings. As the statistical means for the empirical analysis of prevailing interdependencies between lexical items in text strings have elsewhere [3] been developed and discussed to some extent [4], and as the formal representation of vague word meanings derived from these analyses has previously [5] been outlined and illustrated, too [6], an informal description will suffice here.

The algorithm applied so far consists of a consecutive mapping of lexical items onto fuzzy subsets of the vocabulary according to the numerically specified statistical regularities and the differences these items have been used with in the discourse analysed. The resulting system of sets of fuzzy subsets may be interpreted topologically as a n-dimensional hyperspace with a natural metric. Its n linguistically labeled elements (representing meaning points) and their mutual distances (representing meaning differences) form discernable clouds and clusters [7]. These determine the overall structuredness of a domain by measurable semantic (paradigmatic and/or syntagmatic) properties on the lexical items concerned.

Connotative Dependency Structure

Stimulated by the theory of spreading activation in memory models [8] in conjunction with the psychological account of language understanding in procedural semantics [9] a dynamic meaning representation can be developed of the basis of the prototypical, but static representations provided by the semantic hyperspace structure. This is achieved by a recursively defined algorithm which has formally been introduced elsewhere [10] so that it may verbally be described here as a procedure to generate a potential of latent relations among meaning points in the semantic space.

In a way, this procedure reconstructs for this model what recent theories of cognition and language comprehension have introduced in network models of semantic memory: paths of excitation that may be activated from any primed node and which spread along node relating links over the whole network with decreasing intensities. Compared to the execution of spreading activation processes in network models, however, the present procedure — speaking in model genetical terms — must be considered
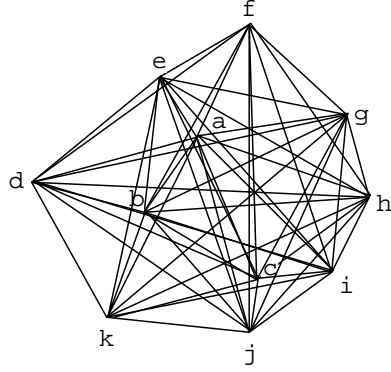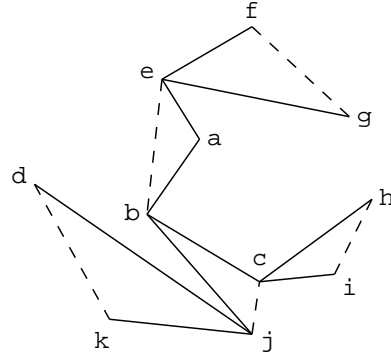
Fig. 1

Fig. 2

of prior status. The semantic hyperspace is not a transitively related network of nodes, but a symmetrically related data structure of linguistically labeled n-tuples of numerical values. Therefore, priming of any item would immediately activate every other item rendering the process of spreading activation undiscriminating for semantic representation. So, the new procedure, first, has to establish links between items and evaluate them by processing the data base provided in order to let these links eventually serve as directed paths along which possible activation might spread. Operating on the distance-like data of the semantic space, the algorithm's generic procedure will start with any meaning point being primed to determine those two other points, the sum of distances between which form a triangle of minimum edges' lengths. Repeated successively for each of these meaning points listed and in turn primed in accordance with this procedure, particular fragments of the relational structure inherent in the semantic space will be selected depending on the aspect, i.e. the primed point the algorithm is initially started with. Working its way through and consuming all labeled points in the space system, the procedure transforms prevailing similarities of meanings as represented by adjacent points to establish a binary, non-symmetric, and transitive relation between them. It allows for the hierarchical rearrangement of meaning points as nodes under a primed head in the format of a binary tree, called connotative dependency structure (CDS).

The process of detection and identification which the algorithm performs may be illustrated in view of a two-dimensional space configuration of 11 points $\langle \delta\{a, b, c, d, e, f, g, h, i, j, k\}\rangle$ (Fig. 1).

Submitted to the search procedure of least triangle under initial priming of the point a the algorithm will identify the number of triangles in Fig. 2 and produce the binary tree as shown in Fig. 3. For the effective use in procedural meaning representation and semantic processing, the CDS-trees may additionally be evaluated by connotative criterialities [10]. The criteriality is a numerical expression of the degree or intensity by which any CDS-node is dependent on the head, calculated as a function both of the involved meaning points' topology and their relative distances in the semantic space. The head's criteriality being 1.0, this value is splitted among every two dependent nodes, and consequently decreases from level to level in the tree structure approximating 0.

4

a

e   b

f   g   c   j

i   h   k   d

Fig. 3

.086 INDUSTRIE

.169 UNTERNEHM

.083 SUCH

.290 STADT

.062 SCHREIB

.121 WUNSCH

.059 SCHUL

.558 ANGEBOT

.075 LEHR

.141 VERKEHR

.066 GESCHAEFT

.268 GEBIET

.078 VERWALT

.127 EINSATZ

.049 WIRTSCHAFT

1.000 ARBEIT

.076 BERUF

.151 STELLE

.075 ELEKTRON

.276 AUSGABE

.065 UNTERRICHT

.125 ORGANISAT

.060 BITTE

.442 PERSON

.053 HERRSCHAFT

.088 ALLGEMEIN

.035 MODE

.166 VERBAND

.040 AUSLAND

.078 VERANTWORT

.038 FOLGE

Fig. 4

5

```
                                                              .075 GEBIET
                                              .150 SCHUL
                                                              .075 VERBAND
                              .290 ELEKTRON
                                                              .073 UNTERNEHM
                                              .130 SCHREIB
                                                              .067 STADT
          .536_SUCH
         ┌ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┐                               .105 BITTE
         ┊                    ┊            .137 UNTERRICHT
         ┊                    ┊                               .032 TECHNIK
         ┊            .246_BERUF ─ ─ ─ ─ ┐                    .058 STELLE
         ┊            └ ─ ─ ─ ─ ─ ─ ─ ─ ┐┊
         ┊                              ┊└ ─.109 ORGANISAT    .051 WUNSCH
         └ ─ ─ ─ ─ ─ ─ ─ ─ ┐
         1.000 INDUSTRIE    │                                 .080 AUSGABE
                           │             .151 FAEHIG
                           │                                  .071 ALLGEMEIN
                           │     .283 COMPUTER
                           │                                  .068 PERSON
                           │             .132 SYSTEM
                           │                                  .064 ANGEBOT
         .464 LEIT ────────┘                                  .059 VERKEHR
                          │              .115 DIPLOM
                          │                                   .056 EINSATZ
         .181 ERFAHR ─────┘                                   .035 ARBEIT
                            .066 KENNTNIS
                                                              .031 VERANTWORT
```
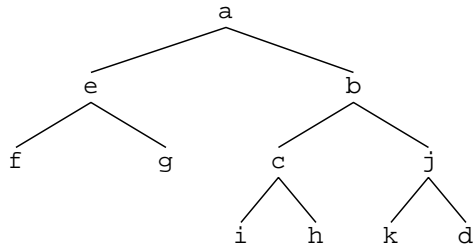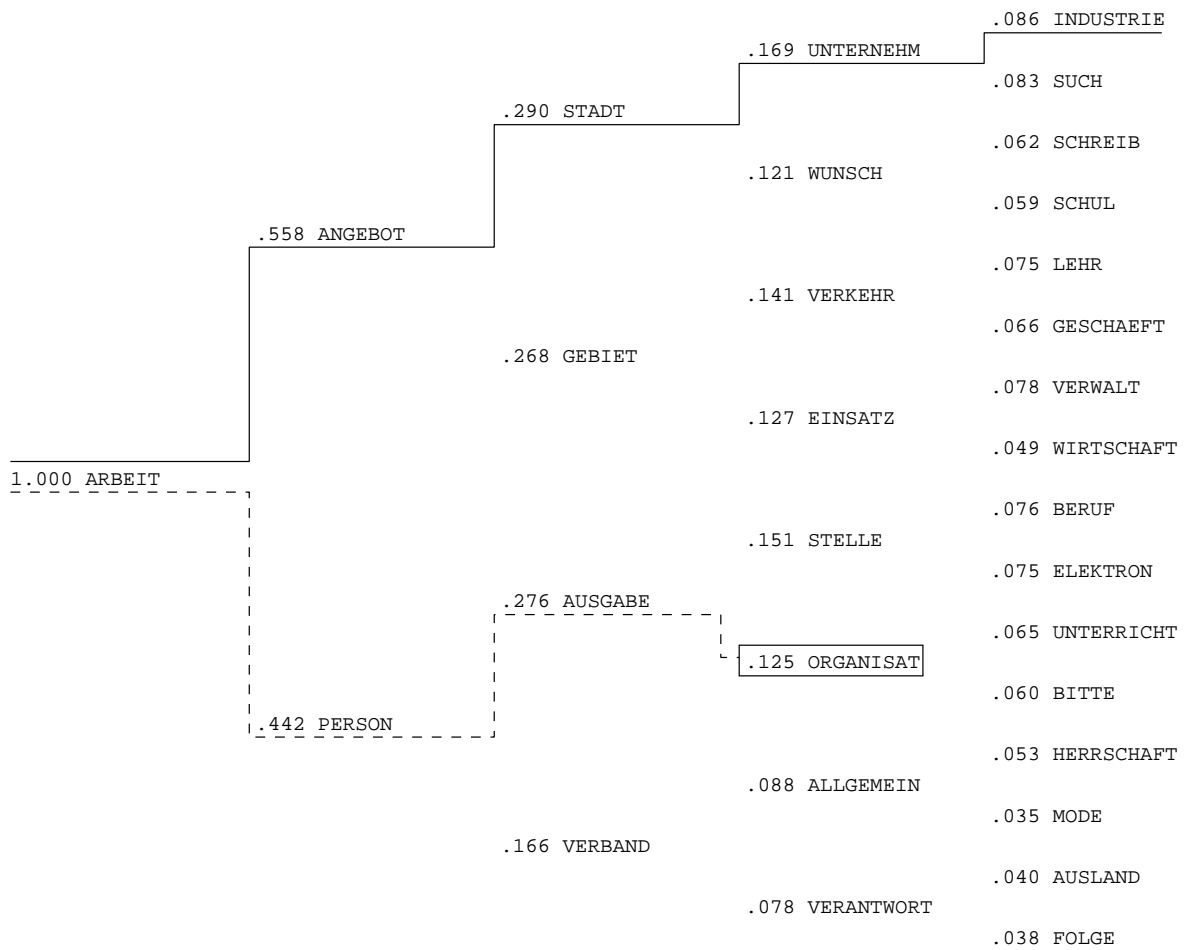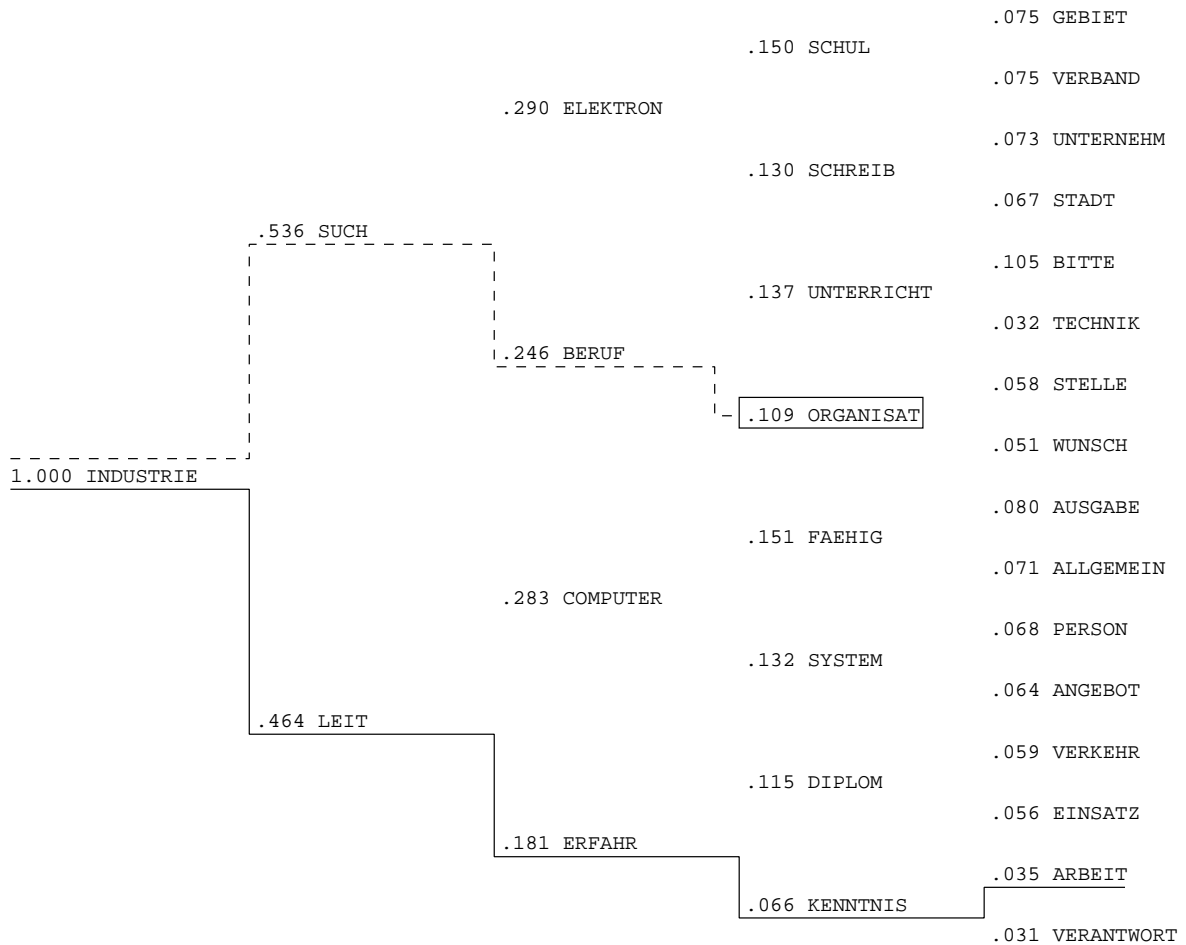
Fig. 5

Examples of connotative dependency trees are given below where the upper fragments of the CDS's of ARBEIT/labour (Fig. 4) and INDUSTRIE/industry (Fig. 5) are shown as computed from the semantic space structure derived of a sample of German newspaper texts from the 1964 daily editions of 'Die Welt'.

It goes without saying that the generating of CDS-trees is a prerequisit to source-oriented search and retrieval procedures which may thus be performed effectively on the semantic space structure. Given, say, the meaning point ARBEIT/labour to be primed, and, say, INDUSTRIE/industry as the target point to be searched, the CDS (ARBEIT) will be generated first. It provides semantic dispositions of decreasing criteriality under the aspect of ARBEIT in the semantic space data. Then, the tree will be searched (breadth-first) for the target node INDUSTRIE. When this is hit, its dependency path will be activated to back-track those intermediate nodes which determine the connotative transitions of INDUSTRIE under the aspect of ARBEIT, namely UNTERNEHMEN/business, STADT/town, ANGEBOT/offer as underlined in Fig. 4.

The priming of INDUSTRIE and the targetting of ARBEIT leads to the activation of quite a different dependency path mediating ARBEIT under the aspect of IN-DUSTRIE, namely by KENNTNIS/knowledge, ERFAHR/experience, LEIT/control,

```
    1.0 ARBEIT                                      INDUSTRIE 1.0

         .442 Person                              Such .536

              .276 Ausgabe                    Beruf .246

                   .125  | ORGANISATION |   .109
```
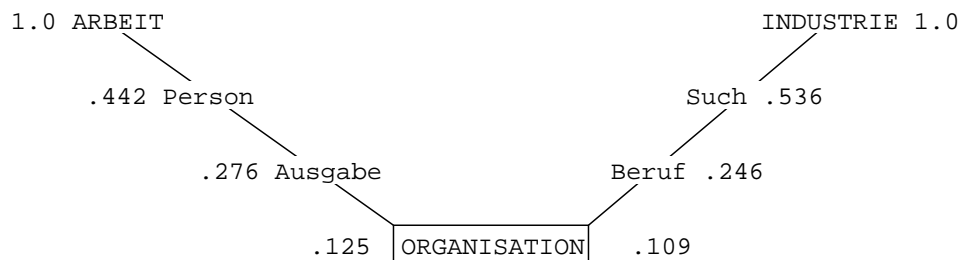
Fig. 6

as underlined in Fig. 5. Using these source-oriented search and retrieval processes, an analogical, contents-dependent form of inferencing, as opposed to logical deduction, may operationally be devised by way of parallel processing of two (or more) CDS-trees. For this purpose an algorithm is started by the two (or more) meaning premises of, say, ARBEIT and INDUSTRIE. Their CDS-trees will be generated before the inferencing procedure begins to work its way (breadth-first) through the trees' levels, taking highest criterialities first in tagging each encountered node. When the first node is met which has previously been tagged already, the search procedure stops to activate the dependency paths from this concluding common node — here, ORGANISAT/organization in the CDS-trees concerned, as illustrated in Fig. 4 and Fig. 5 by dotted lines, separately presented in Fig. 6.

References

[1] Rieger, B.B.: Preface, in: Rieger, B.B. (ed.), Empirical Semantics I (Brockmeyer, Bochum 1981), II–XIII

[2] Minsky, M.: K-Lines — a theory of memory. MIT-AI-Memo 516 (1979)

[3] Rieger, B.: Probleme der automatischen Textanalyse und unscharfen Wortsemantik, in: Krallmann, D. (ed.), Dialogsysteme und Textverarbeitung (LDV-Fittings, Essen 1980) 55–76

[4] Rieger, B.B.: Fuzzy Word Meaning Analysis and Representation in Linguistic Semantics, Proceedings of COLING 80 (Tokyo 1980), 76–84

[5] Rieger, B.B.: Feasible Fuzzy Semantics. On some problems of how to handle word meanings empirically, in: Eikmeyer, H.J./Rieser, H. (eds.), Words, Worlds, and Contexts. New Approaches in Word Semantics (de Gruyter, Berlin/New York 1981) 193–209

[6] Rieger, B.: Unscharfe Wortbedeutungen, in: Hellmann, M.W. (ed.), Ost-West-Wortschatzvergleich (Schwann, Düsseldorf) forthcoming

[7] Rieger, B.B.: Clusters in Semantic Space, Delatte (ed.), Actes du Congrès International Informatique et Sciences Humaines 1981 (LASLA, Liège) forthcoming

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

[9] Miller, G.A./Johnson-Laird, P.N.: Language and Perception (Univ./@ Press, Cambridge 1976)

[10] Rieger, B.B.: Connotative Dependency Structures in Semantic Space, in: Rieger, B.B. (ed.), Empirical Semantics II, (Brockmeyer, Bochum 1981) 622–710