

Collative Semantics

Dan Fass

Computing Research Laboratory
New Mexico State University
Las Cruces, NM, USA 88003

0. Abstract

This paper introduces Collative Semantics (CS), a new domain-independent semantics for natural language processing (NLP) which addresses the problems of lexical ambiguity, metonymy, various semantic relations (conventional relations, redundant relations, contradictory relations, metaphorical relations and severely anomalous relations) and the introduction of new information. We explain the two techniques CS uses for matching together knowledge structures (KSs) and why semantic vectors, which record the results of such matches, are informative enough to tell apart semantic relations and be the basis for lexical disambiguation.

1. Introduction

Collative Semantics is a new domain-independent semantics for NLP. This paper focusses on CS, describing the main issues it addresses (lexical ambiguity, metonymy, semantic relations, introduction of new information) and general details of its knowledge representation, knowledge structures, techniques for matching together knowledge structures, and the way it records the results of matching in semantic vectors.

CS has been implemented in a natural language program called *meta5* which has been described in detail elsewhere (Fass, 1986). Briefly, the program produces semantic representations of single sentences requiring lexical disambiguation and metonymic inferencing. While processing such sentences *meta5* computes semantic vectors for the matches between pairs of word-senses which are rich enough to discriminate between instances of conventional relations, semantically redundant relations, contradictory relations, metaphorical relations and severely anomalous relations between word-sense pairs. Semantic vectors also record the assertion of new information. *Meta5* can process all the example sentences given in the next section.

2. Semantic Relations, Novel Information and Metonymy

In this section we provide brief descriptions of those semantic phenomena considered by CS, except lexical ambiguity, starting with semantic relations. Conventional, metaphorical and severely anomalous relations can all be described using the terminology of Richards (1936). The subject term is the "topic," the term it is compared to is the "vehicle," the similarity or resemblance between them is the "ground," and any difference or dissimilarity is the "tension." We also adopt Perrine's (1971) four-fold classification of metaphors into combinations of explicit and implicit topics and vehicles.

In a metaphorical relation there is tension between the topic and vehicle because the topic term is not a type of vehicle term. What is salient (Ortony, 1979) is given by the context, a salient property of the vehicle is found and an analogical match discovered between it and a property from the topic. The remaining properties of the topic and vehicle have similarities and differences and, although these are not central to recognising the metaphorical relation, the higher the proportion of differences to similarities the "better" the metaphor.

For example, in the metaphorical relation between 'car' and 'drink' in

(1) "The car drank gasoline."

the topic is 'car' and the implicit vehicle is 'animal', the agent preference Wilks, 1975) of 'drink'. The tension is caused by a car (the topic) not being a type of animal (the vehicle). What is salient in this context is the action of drinking, given by the main sentence verb. The salient property of the vehicle is one referring to the salient action, i.e. that an animal drinks potable liquids. An analogical match is found between the salient property of the vehicle and a property of the topic : animals drink potable liquids as cars use gasoline. The ground is the expending of liquid. Matching the remaining properties of the topic and vehicle, some pairs of properties are the same (animals and cars are both bounded, three-dimensional and solid) but other pairs express differences (animals are living, cars are nonliving; animals are composed of flesh, cars are made of metal).

In a conventional relation there is no tension because the topic term is a type of vehicle term. The salient property of the vehicle matches identically with a property from the topic. There is a high proportion of similarities to differences amongst matches of other pairs of properties of the topic and vehicle.

For example, in the conventional relation between 'man' and 'drink' in

(2) "The man drank beer."

the topic is 'man' and the implicit vehicle is again 'animal', the preferred agent of 'drink'. A man (the topic) is type of animal (the vehicle). What is salient is drinking so the salient property of animals is again that they drink potable liquids. An identical match is found between animals drinking potable liquids and men drinking potable liquids. Of the remaining properties of animal and man matched together, a large proportion are similar and very few are different.

A severely anomalous relation has the same tension between topic and vehicle as a metaphorical relation : the topic term is not a type of vehicle term. A salient property is found in the vehicle but it does not find an identical or analogical match with any property from the topic. A high proportion of the remaining properties of the topic and vehicle matched together are different and few are the same.

Our description of semantically redundant and contradictory relations is based on Katz's (1964) definitions. Whether the assertion of a particular property onto a subject term is a redundant relation, contradictory relation or new information is a function of the knowledge already in the term's dictionary entry.

A contradictory relation is one in which a modifier asserts a property onto a subject term which is incompatible with a property already in the subject term, e.g.

(3) "John McEnroe is female."

where the assertion [sex1, female1] from the adjective clashes

with the property [sex1, male1] already in the dictionary entry for John McEnroe.

A semantically redundant relation is one in which a modifier asserts a property onto a subject term which is identical to a property already in that term, e.g.

(4) "John McEnroe is male."

is redundant if the dictionary entry for John McEnroe already contains [sex1, male1], the same as the property asserted by the adjective.

If for sentence (4), the dictionary entry of John McEnroe does not previously contain [sex1, male1], then the property asserted by the adjective is recognised as new information.

Metonymy is a figure of speech in which the name of one thing is substituted for that of another related to it (Lakoff & Johnson, 1980), for example artist for artform as in

(5) "John played Bach."

(= the music of Bach) and

(6) "Ted reads Steinbeck."

(= the writings of Steinbeck). Other types of metonymy include container for contained as in

(7) "Mary drank the bottle."

(= the liquid in the bottle), and co-agent for activity in

(8) "Lendl played Wilander."

(= tennis with Wilander), and part for whole (also known as synecdoche).

3. Knowledge Representation and Knowledge Structures

The knowledge representation (KR) of Collative Semantics is a hierarchically structured semantic network with word-senses as the nodes. Every node is also a sense-frame, the KS of Collative Semantics, so-called because sense-frames are frame-like structures representing individual word-senses. This arrangement of a semantic network with frame-like structures as nodes is common to many frame- and semantic network-based systems (e.g. Bobrow & Winograd, 1977; Roberts & Goldstein, 1977; Brachman, 1979; Hirst, 1983).

In developing our set of arc labels we have taken note of research at the epistemological level of semantic networks (Brachman, 1979) so, for example, we distinguish between 'superinstance' which denotes membership of an individual within a class of individuals from 'supertype' which denotes membership of a class of individuals by a class of individuals.

Much like Quillian's (1968) planes, sense-frames are composed of other word-senses which have their own sense-frames. There are no semantic primitives in the sense of Schank's (1973) Conceptual Dependency or Wilks' (1975) Preference Semantics. In CS word-senses perform the function of semantic primitives. They are used to capture broad generalisations among groups of words or word-senses; to represent the meanings of individual word-senses; to represent the underlying meaning of sentences and their parts; to support lexical disambiguation; to support inferencing of various kinds; and to support the discrimination of semantic relations.

In the next section we explain how semantic primitives are used in Preference Semantics to do lexical disambiguation and how word-senses are used in CS to do the same.

4. Techniques for Matching Together Knowledge Structures

Section 4 explains the two techniques used for matching together KSs in CS : computing a graph relation and cell-

matching.

Computing a graph relation is developed from Wilks' (1975) technique for computing satisfied and violated preferences in the Preference Semantics System. The 80-100 semantic primitives are organised into a directed graph (see Wilks 1977, Appendix A for details) in which they appear as nodes. The arcs are unlabelled but appear to denote set inclusion, e.g. MAN, the class of human beings, belongs to *ANI, the class of animate entities. A path searching algorithm operates over the directed graph. Its source and destination nodes are two semantic primitives. Satisfied preferences are paths describing set inclusion; violated preferences are paths describing set exclusion.

Computing a graph relation is a path search algorithm which operates over CS's hierarchically structured semantic network. Its source and destination nodes are two word-senses in that network. Five kinds of path (or graph relation) are sought between those nodes. Two of the paths describe types of set inclusion and hence are the equivalent of a path for a satisfied preference. The remaining three kinds of path are equivalent to a violated preference and describe types of set exclusion.

Computing satisfied and violated preferences is the basic mechanism for doing lexical disambiguation in Preference Semantics. Collative Semantics has at least the disambiguation power of Preference Semantics because computing a graph relation produces paths describing set inclusion and exclusion over directed graphs with practically the same hierarchical organisation, but whereas the nodes in Wilks' digraph are a restricted set of semantic primitives, the nodes in CS's digraph are an unrestricted set of word-senses.

The result of computing a graph relation is recorded in a five-column **graph relation** array where each column corresponds to one of the five kinds of set inclusion or exclusion. An initialised array contains all zeroes. If computing a graph relation is used and a particular path is found then a 1 is added to the appropriate column of the array. This is a very different way of recording the result of a path search from Preference Semantics (the importance of recording results is considered in the next section).

The second matching technique in CS, cell matching, is a type of multiple comparison algorithm for matching together KSs. A multiple comparison algorithm is a technique used to match together the elements from two sets (often within two KSs) by isolating a pair of elements from the two sets and matching those elements together in some way.

Theories of metaphor in linguistics and psychology which compare two terms in a metaphorical relation presuppose some kind of multiple comparison between their elements. Those elements have been termed "features," "properties" and "attributes" (amongst others) and usually have been contained within KSs, normally some type of frame. For example in linguistic semantics, Levin (1977) attempts to discriminate metaphors from other semantic relations by matching together sets of semantic features and in psychology, Ortony (1979) discusses metaphor as matches between attributes within two schemata. Different kinds of match between elements are considered important. Ortony differentiates "identical" matches of attributes from "similar" matches based on some "structural isomorphism" between the knowledge associated with the two attributes. Similar matches are seen as especially important to metaphor.

Cell matching takes as input two lists of elements called **cells**. A pair of cells are isolated from the two lists using a set of structural constraints and matched together. As cells are composed of word-senses, CS's first matching technique, com-

puting a graph relation, can be used to match cells together. Five different kinds of match are tried between cells, corresponding to the two types of set inclusion and three types of set exclusion. These include, to use Ortony's terms, identical and similar matches.

The results of cell matching are recorded in a seven column **cell-match** array, where each column correspond to a different kind of **cell relation**. The first five columns correspond to the five types of set inclusion and exclusion. Some cells from the two lists fail the structural constraints and cannot be matched at all. The last two columns record such cases. An initialised cell-match array contains all zeroes. When matching two lists of cells together, each occurrence of a cell relation between pairs of cells adds 1 to the appropriate column of the array until all cells have been accounted for. A cell-match array, then, records the complete match of two lists of cells.

5. Recording the Results of Matching Together Knowledge Structures

Section 5 explains how the results of matching together sense-frames using computing a graph relation and cell matching are recorded in semantic vectors. Semantic vectors are informative enough to tell apart semantic relations between word-senses and be the basis for lexical disambiguation.

In other approaches to lexical disambiguation their record of the results of matching together KSs cannot be used to adequately discriminate semantic relations. In Preference Semantics, the record of the result of the path search algorithm is binary : either a satisfied or violated preference. A binary choice is inadequate because a single satisfied preference could be a literal, redundant, contradictory or contrary semantic relation and a single violated preference could be a metaphorical or severely anomalous one (see Fass & Wilks, 1983). The same criticism can be made of marker passing (e.g. Quillian, 1968; Charniak, 1981; Hirst, 1983), in which lexical disambiguation is done by an intersection search algorithm which finds the shortest path between two nodes in a directed graph. The record is a single path, but a single path cannot be used for discriminating among semantic relations. And in numerical weighting schemes (e.g. Waltz & Pollack, 1985) which "takes place on a weighted network of associations, where 'activation energy' is distributed over the network based on some mathematical function of the strength of connections" (Ibid, p.54), the numerical weights on nodes and edges cannot be used to distinguish semantic relations.

Semantic vectors are informative enough to be the basis for lexical disambiguation and tell apart semantic relations between word-senses because they contain graph relation arrays and cell match arrays.

Graph relation arrays are used heavily for lexical disambiguation because a graph relation array records two types of set inclusion, equivalent to a satisfied preference in Preference Semantics, and three types of set exclusion, equivalent to a violated preference. It can be used for lexical disambiguation just as Preference Semantics uses satisfied and violated preferences.

Cell match arrays record the multiple comparison of elements in two KSs in terms of seven different kinds of match. A semantic vector usually contains two or more of these arrays. Cell match arrays can be used for differentiating semantic relations because each relation produces a distinctive pattern of numbers in the columns of the arrays. See (Fass, 1986) for examples.

6. Summary

The knowledge representation and knowledge structure Collative Semantics is similar to that of many existing frame- or semantic-network based systems. What makes CS different and able to address the issues it does (lexical ambiguity, metonymy, semantic relations, introduction of new information) is the techniques it uses for matching together knowledge structures and the recording of those matches into semantic vectors.

7. Bibliography

- Bobrow, D., & Winograd, T. (1977) "An Overview of KRL, A Knowledge Representation Language." *Cognitive Science*, 1, pp. 3-46.
- Brachman, R.J. (1979) "On The Epistemological Status of Semantic Networks." In N.V. Findler (Ed.) *Associative Networks : Representation and Use of Knowledge By Computers*, New York : Academic Press, pp. 3-50.
- Charniak, E. (1981) "Passing Markers : A Theory of Contextual Influence in Language Comprehension." Technical Report CS-80, Dept. of Computer Science, Brown Univ., Providence, Rhode Island.
- Fass, D.C. (1986) "Collative Semantics : A Description of the Meta5 Program." Memorandum M CCS-86-23, Computing Research Lab., New Mexico State Univ., New Mexico.
- Fass, D.C. & Wilks, Y.A. (1983) "Preference Semantics, Ill-Formedness and Metaphor." *American Journal of Computational Linguistics*, 9, pp. 178-187.
- Hirst, Graeme (1983) "Semantic Interpretation against Ambiguity." Technical Report CS-83-25, Dept. of Computer Science, Brown University.
- Katz, J.J. (1964) "Analyticity and Contradiction in Natural Language." In J.A. Fodor & J.J. Katz (Eds.) *The Structure of Language : Readings in the Philosophy of Language*, Englewood Cliffs, N.J. : Prentice Hall.
- Lakoff, G., & Johnson, M. (1980) *Metaphors We Live By*, London : Chicago University Press.
- Levin, S.R. (1977) *The Semantics of Metaphor*, Baltimore : John Hopkins University Press.
- Ortony, A. (1979) "Beyond Literal Similarity." *Psychological Review*, 86, pp. 161-180.
- Perrine, L. (1971) "Psychological Forms of Metaphor." *College English*, 33, pp. 125-138.
- Quillian, M.R. (1968) "Semantic Memory." In M. Minsky (Ed.) *Semantic Information Processing*, Cambridge, Mass : MIT Press, pp. 216-270.
- Richards, I.A. (1936) *The Philosophy of Rhetoric*, London : Oxford University Press.
- Roberts, B.R., & Goldstein, I.P. (1977) "The FRL Manual," MIT AI Memo 409.
- Schank, R.C. (1973) "Identification of Conceptualizations Underlying Natural Language." In R.C. Schank & K.M. Colby (Eds.) *Computer Models of Thought and Language*, San Francisco : W.H. Freeman, pp. 187-247.
- Waltz, D.L., & Pollack, J.B. (1985) "Massively Parallel Parsing : A Strongly Interactive Model of Natural language Interpretation." *Cognitive Science*, 9, pp. 51-74.
- Wilks, Y.A. (1975) "A Preferential Pattern-Seeking Semantics for Natural Language Inference." *Artificial Intelligence*, 6, pp. 53-74.
- Wilks, Y.A. (1977) "Good and Bad Arguments about Semantic Primitives." *Communication and Cognition*, 10, pp. 181-221.