

An Algorithm for Aspects of Semantic Interpretation Using an Enhanced WordNet

In Proceedings of NAACL-2001, Pittsburgh, June 2001

Fernando Gomez

Dept. of Computer Science
University of Central Florida, Orlando, FL 32816
gomez@cs.ucf.edu

Abstract

An algorithm for semantic interpretation is explained¹. The algorithm is based on predicates defined for WordNet verb classes. The algorithm is driven by the definition of these predicates whose thematic roles are linked to the WordNet ontology for nouns and to the syntactic relations that realize them. The algorithm has been tested in the identification of the meaning of the verb, thematic roles, and temporal and spatial adjuncts.

1 Introduction

The semantic interpretation algorithm explained in this paper offers a solution to the following interpretation problems: determination of the meaning of the verb, identification of thematic roles and adjuncts, and attachments of prepositional phrases (PPs). An interesting aspect of the algorithm is that the solution of all these problems is interdependent. The interpretation algorithm uses WordNet (Miller et al., 1993) as its lexical knowledge-base. Predicates, or verbal concepts, have been defined for WordNet verb classes, which have been reorganized considerably following the criteria imposed by the interpretation algorithm. WordNet ontology for nouns has also undergone some reorganization and redefinition to conform with the entries in the thematic roles of the predicates. One of the views that guides this research is that the syntax of many verbs is determined by their meaning. Some verbs that are highly ambiguous, say more than 10 senses, and light verbs, which do not lexicalize anything, will need special definitions.

Briefly, the algorithm is as follows. For every verb in a sentence, WordNet provides a list of verb synsets for which we have defined predicates. These predicates can be viewed as contenders for the meaning of the verb. As syntactic relations are parsed, the interpreter checks each predicate in order to see if the predicate has a thematic role which is realized by the syntactic relation. If so, the interpreter records this fact and gets the next syntactic relation. The predicate that realizes the most syntactic relations in the sentence is selected as the meaning of the verb.

This paper is organized as follows. The first part of the paper - sections 2 to 5 - explains the methodology for building predicates for WordNet verb classes, and the second part - sections 6 to 9 - describes the semantic interpretation algorithm, testing and conclusions.

2 Defining Predicates for WordNet Verb Classes

One of the problems of linking WordNet verb classes (Fellbaum, 1993) to the semantic interpretation process is that WordNet does not present a classification of verbs based on semantic decomposition, but rather one based on relational semantic analysis, in which lexical items are taken as irreducible meaning units. Although a full semantic decomposition of verbs (Jackendoff, 1990; Talmy, 1985; Dowty, 1979), much less one that reduces verbs to a small set of primitives (Schank, 1975), is not very relevant to semantic interpretation, some abstract semantic predicates are needed to anchor classes of verbs and their subclasses. These abstract semantic predicates and their more specific sub-predicates contain the selectional restrictions and syntactic relations for the thematic roles defining them. The entries in the predicates

¹This paper is a modified version of the technical report *Semantic Interpretation of Unrestricted Domains using WordNet*, CS-TR-98-02

will be used by the semantic interpreter to attach modifiers and to link syntactic relations to semantic ones. Although WordNet does not explicitly recognize conceptual components, these are implicitly present in the relations linking subordinate verbs to superordinate verbs. The main relation linking a subordinate verb to its superordinate, is that of troponymy (Fellbaum, 1993), meaning that *to V1 is to V2 in some particular manner*, where *manner* covers a diverse class of semantic relations, including the intention of the agent of the action, the way in which the action is carried out, the instrument, etc. Thus, we have taken a top-down approach that defines generic abstract predicates subsuming semantically and syntactically a large class of verbs. WordNet verb classes have been mapped into these generic predicates. Some of this mapping has required us to define new classes and to reclassify and/or redefine some WordNet classes and subclasses. Two major consequences derive from anchoring verb classes in abstract semantic predicates: coalescing many WordNet senses into a predicate, and mapping the same WordNet synset into distinct predicates. For instance, all the 5 synsets listed by WordNet for “travel”: “travel1, go, move, locomote;” “travel2, journey;” “travel3, take a trip, make a trip;” “travel4, journey;” and “travel5 (undergo transportation, as in a vehicle)” can be coalesced into the abstract semantic predicate *change-of-location-by-animate*. This predicate defines a class of verbs containing the most generic properties shared by all members of the class. The *differentia* between this predicate and its subpredicates are given by one or more of the following: a) specific selectional restrictions for the thematic roles, b) different syntactic realizations of the thematic roles, and c) specific sets of inferences associated with the subpredicates. For instance, the *instrument* of the predicate *drive-a-vehicle*, not to be confused with the verb “drive,” is always a vehicle, while the *instrument* of *change-of-location-by-animate* can be an animate, an animate body part, etc. The *instrument* of *drive-a-vehicle* is never realized by a subject, while the *instrument* of the generic predicate can be realized by a subject, e.g., “This bus goes to Cambridge every Wednesday.” Note that in the sentence “This car/bus drives well” the meaning of “drive” is

not *change-of-location-by-animate*.

On the other hand, *migrate* differs from *change-of-location-by-animate* only by the specific inferences associated with this predicate. The inferences derived from “Three hundred Irishmen migrated to Boston last year” would be very different had the main verb been “traveled,” yet there are no differences between the selectional restrictions and syntactic realizations for the predicates *migrate* and *travel*. Subpredicates inherit all the thematic roles not listed in their definitions from their parent predicates. The number and nature of the thematic roles depend on the generic predicates and subpredicates, and not on some general criteria regardless of each predicate (Fillmore, 1968). See (Gomez, 1998) for a discussion.

3 The syntax of roles

The syntax of roles in the predicates that will be defined below is given by the following grammar:

```
S -> (ROLE L); ROLE -> agent|theme|..
L -> (SM) (SR) L| (SM) (SR)
SM -> ONTOLOGICAL-CAT SM|ONTOLOGICAL-CAT
SM -> -ONTOLOGICAL-CAT SM|-ONTOLOGICAL-CAT
SR -> SYNTACTIC-REL SR|SYNTACTIC-REL
SYNTACTIC-REL -> (prep PREP)|subj|
|obj|obj2|subj-if-obj|subj-if-no-obj|
|obj-if-obj2|predicate-complement|
|complement-phrase|adjective-phrase
PREP -> ANY-PREP PREP| ANY-PREP
ONTOLOGICAL-CAT ->thing|physical-thing|...
ANY-PREP -> on|in|....
```

Each thematic role is followed by any number of list pairs. The first list contains the selectional restrictions, a subset of the ontological categories in WordNet, in order of preference (Wilks, 1975) for the thematic role, and the second list contains the syntactic relations (henceforth, SRs) that may realize the thematic role. For any given sentence, only one of the SRs in the list realizes the role. An ontological category preceded by the sign “-” in a selectional restriction means that the thematic role is not realized by that ontological category. The entry *obj* refers to the first postverbal NP, *obj2* to the second postverbal NP. *Subj-if-obj* refers to the subject of a sentence that also has an object (the verb is used transitively), and *subj-if-no-obj* refers to the subject of a sentence containing

no object (the verb is used intransitively). Obj-if-obj2 refers to the obj of a sentence having also an obj2 (the verb is used ditransitively). Thus, subj refers to the subject of a sentence without expressing any context about the transitivity or intransitivity of the verb.

4 Verbs of Communication

In this section, we explain the predicate *communicate* so that the algorithm may be followed by the reader. Other definitions of predicates may be found in (Gomez, 1998).

```
[communicate
(is-a (interact))
(wn-map (communicate2) (transfer2))
(agent (human-agent animal) (subj))
(theme ( abstraction) (cp obj obj2)
(thing) ((prep about of))
(abstraction) ((prep on)))
(recipient (human-agent animal) (obj
obj-if-obj2 (prep to)))
(medium-of-expression
(written-communication)(subj (prep in)))]
```

The *wn-map* entry means that all the WordNet verb forms under “communicate2” and “transfer2” are in principle coalesced into the predicate *communicate*. These two synsets include over 2000 verb forms. The synset *transfer2* in WordNet is a subclass of *move2*, *displace4*, which includes physical things as the things being transferred or moved. However, our predicate *communicate* and its subpredicates encompass the transfer of information, thoughts and abstractions in general by an animate being to another, excluding the transfer of physical things. Subpredicates of the predicate *communicate* will be defined as explained below. But, if they were not defined, then any form that is mapped by WordNet into any of these two synsets would be mapped into the predicate *communicate*. Of course, if the form is also mapped by WordNet into another synset besides any of the senses of “communicate2,” or “transfer2” then it will be mapped into whatever predicate we identify for that synset.

The entry *agent* indicates that the agents of *communicate* are entities belonging to the classes *animal* or *human-agent* (a human, or a social-group). It also indicates that the agent is realized syntactically by the subject of the

sentence. The *theme* role contains three pairs of lists. The first pair expresses this role when it is syntactically realized by a cp (complement phrase), obj (first postverbal NP) or obj2 (second postverbal NP). The selectional restrictions for *theme* when it is syntactically realized by cp, obj or obj2 is *abstraction*.

The second pair of lists indicate that the *theme* may also be syntactically realized by the prepositions “of” and “about,” and the selectional restriction is “thing,” which includes every entity in the ontology. Lastly, the final pair indicates that the *theme* can be realized by the preposition “on,” and the selectional restriction is *abstraction*. The sign “-” preceding a selectional restriction means that any noun sense that is subsumed by the semantic category preceded by “-” does not realize that thematic role. Thus, if one had written “-human-agent” in the selectional restriction for the *theme* when it is syntactically realized by obj, the interpreter would not have taken “President,” or “Congress” as the *theme* in “He briefed the President/Congress.” Of course, in this case this is unneeded because *human-agent* is not an *abstraction* in the ontology. The *recipient* role is syntactically realized by obj, (e.g., “He informed the President.”), by obj-if-obj2, (e.g., “She told the judge the truth.”), and the preposition “to,” (e.g., “She told the truth to the judge.”), and the selectional restrictions are *human-agent* and *animal*. The role *medium-of-expression* indicates the medium used for the transfer of information and is syntactically realized by a subject (e.g., “The first book tells of her youth and early studies,” or the preposition “in” (e.g., “In her first book, she tells of young women who react violently against a dehumanized society.”)

Subpredicates of the predicate *communicate* are defined. For instance, consider the WordNet synset “teach1, learn, instruct” which has *communicate2* as its superclass. This synset forms a class in WN with over 40 entries. Below, it is the hierarchy for the predicate *teach* and some of its subpredicates, which are listed followed by the corresponding synsets in WordNet.

```
teach(wn-map(teach1))(is-a(communicate))
* train(wn-map(train1 train2))
* indoctrinate(wn-map(indoctrinate1))
* lecture(wn-map(lecture1))
```

* educate(wn-map(educate1))

Space limitation impedes us to list the definitions for these predicates, but some of the differences between *teach* and its superpredicate, *communicate* are as follows. The definition for the predicate *teach* is identical to that of *communicate*, except that the *theme* is not realized by the preposition “of.” Its subpredicates *train*, *lecture*, and *educate* do not realize syntactically the *theme* with an obj, and so on. Educate1 is not a subclass of *teach* in WordNet, but of *polish2*, which in turn it is a subclass of *better2* and this of *change1*. We have classified educate1 as a subpredicate of *teach* because it shares more thematic roles and inferences with this predicate than with any other. The complete hierarchy for *communicate* that we have defined includes over 124 predicates.

5 Semi-Automatic Construction of Predicates

As of this writing, we have defined over 1300 predicates and mapped 85% of WordNet verb classes into predicates. The final product that is coming out from this effort differs very considerably from WordNet verb classes because of the different criteria for predicate classes. The definition of predicates has proceeded on a top down manner, namely defining predicates and subpredicates for WordNet classes and subclasses, respectively. This has required a considerable amount of work, because of the need to redefine and reorganize the WordNet verb classes, and, to a lesser extent, the ontology for nouns. For instance, the concept *written-communication*, which has many subconcepts, is categorized in Wordnet only as an *abstraction*. Thus, the interpreter failed to interpret such simple sentences as “She burned the letter/She put the letter on the table,” because “letter” does not have *physical-thing* as one of its hypernyms (superconcepts). The changes to the ontology of nouns took much of our initial effort because many of interpreter failures were due to the noun ontology. We are finishing a report that describes these changes, which have resulted into a noun ontology whose content and structure have been determined by the semantic interpretation algorithm, rather than by our intuitions about how to categorize ontologically a concept.

But, now that most predicate classes have been defined and that the reorganization and structure of the ontology for nouns is close to completion, there is considerable room for the semi-automatic construction of predicates for those individual verbs that, because of their high polysemy, require predicate definitions tailored to them. For our developing purposes, we have used an electronic version of the *The World Book Encyclopedia* (World Book, Inc., Chicago, 1987). We have partitioned the articles into subject classes. We have grouped all articles that deal with biographies, all those about animals, about countries, diseases, etc. Each class forms a corpus in which the verb senses and the noun ontology exhibit certain common features. Diverse functions skim the articles extracting all sentences in which a verb (or any word) is used. The sentence is passed to the parser and, then, to the semantic interpreter for interpretation.

When a failure in interpretation occurs because of one of the following reasons: a) the sense of the verb was not correctly identified, b) no meaning was assigned to a syntactic relation, or c) a thematic role or an adjunct was incorrectly identified, several functions are activated. These functions skim the classes of articles searching for sentences which share some syntactic and semantic properties with the ones that the interpreter fails to interpret. Thus, if the interpreter fails to identify some thematic roles in a sentence, the syntactic similarity algorithm selects all those sentences containing the syntactic relations that realize the unidentified thematic roles. Then, the algorithm chooses from those sentences those in which the ontological category of the head nouns of the syntactic relations are semantically similar as defined by Resnik (Resnik, 1999). Then, the user selects from these ontological categories the selectional restrictions for the thematic roles of the predicate being defined. In many cases, the ontological categories proposed by the acquisition algorithm are the exact ones, requiring little intervention on the part of the user.

6 Interpretation Algorithm

The semantic interpretation algorithm is activated by the parser after parsing a clause. Thus, interpretation does not start until the end of the clause is reached. The interpreter requires

from the parser an identification of the NP complements and PPs of the clause, or sentence. Our parser recognizes clausal and NP complements, relative clauses, resolves gaps resulting from questions and relativization, and handles coordination and subordination. It does not resolve structural ambiguity, which is the sole task of the interpreter. Our mapping of WordNet verb synsets to predicates provides a list containing the predicates for the verb of the clause. The goals of the algorithm are to select one predicate from that list, attach PPs and identify thematic roles. All these tasks are simultaneously achieved. For each syntactic relation in the clause (starting with the NP complements) and for every predicate in the list of predicates, the algorithm checks if the predicate explains the SR (syntactic relation). A predicate *explains* an SR if there is a thematic role in the predicate realized by the SR and the selectional restrictions of the thematic role subsume the ontological category of the head noun of the syntactic relation. This process is repeated for each SR in the clause and each predicate in the list of predicates. Then, the predicate that explains the most SRs is selected as the meaning of the verb. The thematic roles of the predicate have been identified as result of this process. In case of ties, the predicate that has the greater number of thematic roles realized is selected. The algorithm in greater detail is as follows.

Step I. Get the definitions of the predicates which the verb of the clause may stand for. Put those predicates in the list VC-L. (Our mapping of WordNet synsets to predicates produces that list.)

Step II. Check which thematic roles in every VC in VC-L are realized by the NP complements of the clause, starting with the subject:

1. Match Subj (pre-verbal NP) Let HN-OF-SUBJ be the head noun of the subject of the clause.

1.1 *For* every VC in VC-L, initialize the list Matches to nil. *Match* the *subj* selectional restrictions in VC against the senses of HN-OF-SUBJ. Store the result in the list Matches. Save Matches under VC. End *For*. (The *match* algorithm is explained below.)

1.2 If the clause has an object, *For* every VC in VC-L initialize the list Matches to nil. *Match* the *subj-if-obj* selectional restrictions in the VC

against the senses of HN-OF-SUBJ. Store the result in Matches. Save Matches under VC. End *For*.

1.3 Else If the clause has no object, *For* every VC in VC-L initialize the list Matches to nil. *Match* the *subj-if-no-obj* selectional restrictions in the VC against the senses of HN-OF-SUBJ. Store the result in Matches. Save Matches under VC. End *For*.

2. Match obj and obj2 (post-verbal NPs). Let HN-OF-OBJ be the head noun of the *obj* in the clause. Let HN-OF-OBJ2 be the head noun of *obj2* in the clause.

2.1 If the clause has an *obj2*, then

2.1.1 *For* each verbal concept VC in VC-L, initialize Matches to nil. *Match* the *obj-if-obj2* selectional restrictions in VC against the senses of HN-OF-OBJ. Store the results in Matches. Save Matches under VC. End *For*.

2.1.2 *For* each verbal concept VC in VC-L, initialize Matches to nil. *Match* the *obj2* selectional restrictions in VC-L against the senses of HN-OF-OBJ2. Save Matches under VC. End *For*.

2.2 Else if the clause has no *obj2*, *For* each verbal concept VC in VC-L, initialize Matches to nil. *Match* the *obj* selectional restrictions in VC against the senses of HN-OF-OBJ. Save Matches under VC. End *For*.

Step III. Match post-verbal PPs. *For* every VC in VC-L do *For* every PP starting with the one placed immediately after the verb (left to right) do: If the preposition of the PP is listed in the representation of VC and the selectional restrictions in the VC for that preposition *match* the senses of the head noun of the PP, then save this result under VC. End *For*. End *For*. A *match* of a PP results in identifying the thematic role for that VC.

The Match Algorithm

The algorithm *match* that matches the semantic categories in the selectional restrictions against the senses of a noun is given below. The algorithm consists of an outer loop and two un-nested inner loops. The inner loops iterate in *each* noun sense. One of the inner loops handles the exclusionary semantic categories (those preceded by “-”) and another the confirmatory ones (not preceded by “-”). The outer loop iterates in the selectional restrictions in the thematic roles until a confirmatory category subsumes a noun

sense, or the list of selectional restrictions is exhausted. Let SLR (selectional restrictions) for a thematic role be = $slr_1, slr_2, \dots, slr_n$ and HP, (the senses of a noun after being mapped into the core ontology) be = hp_1, hp_2, \dots, hp_m . Each hp_i has the form (ontological-category noun1 noun2 ... noun*i*). The first element in hp_i is the ontological category.

```

Set match-slr = nil.
Loop1: For each r in SLR do
If r is an exclusionary semantic category
  then
  For each hp in HP do
    If r subsumes the first element in hp
      then delete hp from HP
  End for
If r is a confirmatory semantic category
  then For each hp in HP do
    If r subsumes the first element in hp
      then begin
        Save the match;
        Set match-slr = t;
      end
    End for
If match-slr is t then exit Loop1
End Loop1

```

Step IV. After running the algorithm, the VCs are ranked placing first the one that explains the most syntactic relations in the sentence and last the one that explains the least number of syntactic relations. If several VCs are tied, then they ranked by placing first the VC that has the greater number of thematic roles realized by the clause and last the one that has the least number. If, after this filter, there are still VCs that are tied, the algorithm selects the VC that corresponds to the most frequent verb sense as defined by WordNet. Although two VCs may explain the same number of syntactic relations, one of them may have more thematic roles realized. This is due to the fact that the same syntactic relation may realize more than a thematic role. For instance, the subject of verbs that express a change of location by an animate being may realize both the *theme* and the *agent*. Thus, in the sentence “Peter ran today” several VCs are going to explain all the syntactic relations (e.g., run-a-firm, operate-a-machine, ..), but *moving-by-running* will have more thematic roles realized because the subject realizes the

agent and *theme*. Hence, *moving-by-running* is preferred over the other VCs which have not realized their *themes*.

Step V. Determine adjuncts. Every syntactic relation that has not been mapped to a thematic role must be an adjunct or an NP modifier. The entries for adjuncts are stored in the root node *action* and are inherited by all VCs. Adjuncts are identified *after the meaning of the verb has been determined* because adjuncts are not part of the argument structure of the predicate. Thus, in “Peter eats every day at school,” the meaning of “every day” and “at school” are identified after the predicate has been determined. Prepositions that are listed in the *action* node as *weakly* claimed by the verb may modify the verb or a NP. Heuristic rules decide in those cases. Note that even if the parser recognizes “every day” as an object, not as a time NP, the algorithm will not take “every day” as a *theme* of the predicate because *time* will not *match* the selectional restrictions of *ingest*. However, the interpreter will still recognize “every day” as a temporal adjunct, because the entry for temporal adjunct in the node *action* says that it can be realized by a time-np or an obj. The entry obj is there just in case that the parser makes a mistake.

7 Testing

In the last two years, we have conducted diverse tests and the system has been demoed in a major conference. We plan to demo it in this conference, which is perhaps the best test we can offer. Some of the tests have evaluated the performance of the algorithm in identifying subpredicates of major predicates such as *change-of-location-by-animate* with over 1200 verb forms in Wordnet, *cause-change-of-location* with over 1600 verb forms, *transfer-of-possession* with over 1500 forms, *transfer-of-information* with over 1000 forms, and *judge* with 600 forms and others.

Recently, we performed a bottom-up evaluation. We selected 30 verbs randomly from the set of 165 verbs having more than 10 senses in Wordnet. These were the verbs selected: *accept break bring carry catch charge come deliver dress fall feel fly grow hold lead leave meet pass pick present ride run serve strike suffer take throw turn walk work*.

Table 1: Test Results

	Test1	Test2	Test3
Predicate	96%	95%	91%
PP Attachment	96%	96%	NA
Thematic-Roles	95%	95%	NA
Adjuncts	97%	97%	NA

Some of these verbs actually have more than 40 senses. We tested the performance of the algorithm in identifying the predicate (meaning of the verb), the attachment of PPs, thematic roles and adjuncts. For our testing corpus, we used an electronic encyclopedia. For each verb tested a program selected ten sentences. This produced 300 sentences - 10 for each verb.

A problem with this test, test1, is that it does not test all predicates for each of the verbs selected. The reason is that the sentences randomly selected by the testing program may have the same predicates. Selecting one or two sentences per predicate is not a good evaluation procedure either because the interpreter may identify predicate p in sentence $s1$, but it may fail to identify p in sentence $s2$. Moreover, it may identify p in $s2$, but it may fail to identify its thematic roles, adjuncts, etc. Then, we conducted test2 in which other 300 sentences were randomly selected but the selection program discarded those sentences with a predicate already tested in test1. The results were very similar to test1. Yet, some of the less frequent predicates for some of the verbs were not selected by this procedure either. The results of both tests are in Table 2.

Metaphoric uses of the verb were a cause of failure. The failure for thematic roles and PP attachment were mainly due to missing prepositions in the thematic role entry of the predicate and overgeneralizing the definition of a subpredicate. That is, stating that the subpredicate inherits all or most of its thematic roles from its super-predicate when that is not the case.

Test3 is a very recent test that was conducted on the 5000 words that were sense tagged for SIGLEX99. We tested the system on determining the meaning of the verb. Thematic roles, PP attachment and adjuncts are not provided in the data. We excluded from the test the verbs: "be," "do," "have." We selected the verbs that

occurred 3 or more times in the text. The worse performance of the system on this text is explained by the fact that some of the senses in some of the verbs ("make," "go," and "come" especially) have not yet been covered.

8 Related Research

Besides the obvious influence of WordNet, this work is very much related to Palmer's VerbNet project (Dang et al., 1998), and has benefited from (Levin, 1993) and (Pritchett, 1992). The lexical view expressed in this paper that the syntax of many verbs is determined by their meaning have been also presented by (Pinker, 1989; Grimshaw, 1990; Levin, 1993) and others. A stronger point for which we offer some evidence in this paper, is that, if the above view is correct, then the generic predicates drive the semantic interpretation process which will need only a minimal syntactic analysis as that expressed by D-theory (Marcus et al., 1983) and minimal commitment model parsers (Weinberg, 1993; Gorrell, 1991). The other guiding principle of this research is that the meaning of the verb depends not only on its selectional restrictions, but also on the syntactic relations that realize them. A similar view has been presented in (Pustejovsky, 1995). The lexical aspect of this research is also related to the FrameNet project (Baker et al., 1998), which is producing frame-semantic descriptions for English lexical items.

Hirst (Hirst, 1992) also used an eliminative approach to resolve verb meaning and thematic roles simultaneously. The algorithm has also some similarities to the one reported in (Gomez et al., 1997), but there are major differences. For instance, the algorithm in (Gomez et al., 1997) relies on what the authors call VM-rules to determine the meaning of the verb. These rules, which are stored in the verb forms and are order sensitive, are tried as the parser parses syntactic relations. If one of the VM-rules fires, determining the meaning of the verb, then the thematic relations and attachment of prepositions are resolved by accessing the predicate definition. However in the present algorithm, the meaning of the verb and thematic relations are synergistically determined by the scoring algorithm. This makes unnecessary the VM-rules and grounds the algorithm solely on the declarative representations of the predicates, whose

selectional restrictions are based on a general ontology of English, namely WordNet.

9 Conclusions

We have presented an algorithm for the semantic interpretation of unrestricted domains. The algorithm is based on the WordNet general ontology for nouns and on a definition of predicates which have been linked to the noun ontology and organized in a hierarchy in which subpredicates inherit thematic roles from its superpredicates. The algorithm, which has been implemented in Common Lisp and is running on Sparc 5 workstations, has been tested in its ability to attach PPs, identify thematic roles and adjuncts.

Acknowledgements

I thank Christiane Fellbaum for encouraging this work, and Carlos Segami for implementing most aspects of the algorithm.

References

- C. F. Baker, C. J. Fillmore, and J.B. Lowe. 1998. The berkeley framenet project. In *Proc. of COLING-ACL-98*, pages 79–85, Montreal, Quebec.
- Hoan Trang Dang, Karin Kipper, Martha Palmer, and Joseph Rosenzweig. 1998. Investigating regular sense extensions based on intersective levin classes. In *In COLING-ACL*, pages 293–299, Montreal, Quebec.
- D. R. Dowty. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel, Dordrecht.
- C. Fellbaum. 1993. English verbs as a semantic net. Technical report, Princeton. CSL Report 43, revised March 1993.
- C. J. Fillmore. 1968. The case for case. In E. Bach and R.T. Harms, editors, *Universals in Linguistic Theory*, pages 1–88. Holt, Rinehart and Winston, New York, NY.
- F. Gomez, C. Segami, and R. Hull. 1997. Determining prepositional attachment, prepositional meaning, verb meaning and thematic roles. *Computational Intelligence*, 13(1):1–31.
- F. Gomez. 1998. Linking wordnet verb classes to semantic interpretation. In *COLING-ACL Workshop on Usage of WordNet in NLP Systems*, pages 58–64, U. de Montreal.
- P. Gorrell. 1991. Subcategorization and sentence processing. In R.C. Berwick and S.P. Abney and C. Tenny, editors, *Principle-Based Parsing: Computation and Psycholinguistics*. Kluwer Academic, Dordrecht, The Netherlands.
- J. Grimshaw. 1990. *Argument Structure*. MIT Press, Cambridge, Mass.
- G. Hirst. 1992. *Semantic interpretation and the resolution of ambiguity*. Cambridge University Press, New York.
- R. Jackendoff. 1990. *Semantic Structures*. MIT Press, Cambridge, Mass.
- B. Levin. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press, Chicago.
- M.D. Marcus, D. Hindle, and M. Fleck. 1983. D-theory: Talking about talking about trees. In *Proc. of the Annual Meeting of the ACL*.
- G.A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller. 1993. Introduction to WordNet: An on-line lexical database. Technical report, Princeton. CSL Report 43, revised March 1993.
- S. Pinker. 1989. *Learnability and Cognition*. MIT Press, Cambridge, Mass.
- B. L. Pritchett. 1992. *Grammatical Competence and Parsing Performance*. The University of Chicago Press, Chicago, Illinois.
- J. Pustejovsky. 1995. *The Generative Lexicon*. MIT Press, Cambridge, Mass.
- P. Resnik. 1999. Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. *Journal of Artificial Intelligence Research*, 11:95–130.
- R. Schank. 1975. *Conceptual Information Processing*. North Holland, Amsterdam.
- L. Talmy. 1985. Lexicalization patterns: semantic structure in lexical forms. In T. Shopen, editor, *Language typology and syntactic description. Vol. 3: Grammatical categories and the lexicon*. Cambridge University Press.
- A. Weinberg. 1993. Parameters in the theory of sentence processing: Minimal commitment theory goes east. *Journal of Psycholinguistic Research*, 22(3):339–364.
- Y.A. Wilks. 1975. Preference semantics. In E.L. Keenan, editor, *Formal Semantics*

of Natural Language. Cambridge University
Press, Cambridge, UK.