

# Semantic Parsing Based on FrameNet

Cosmin Adrian Bejan, Alessandro Moschitti, Paul Morărescu,  
Gabriel Nicolae and Sanda Harabagiu

University of Texas at Dallas  
Human Language Technology Research Institute  
Richardson, TX 75083-0688, USA

## Abstract

This paper describes our method based on Support Vector Machines for automatically assigning semantic roles to constituents of English sentences. This method employs four different feature sets, one of which being first reported herein. The combination of features as well as the extended training data we considered have produced in the Senseval-3 experiments an F1-score of 92.5% for the unrestricted case and of 76.3% for the restricted case.

## 1 Introduction

The evaluation of the Senseval-3 task for Automatic Labeling of Semantic Roles is based on the annotations made available by the FrameNet Project (Baker et al., 1998). The idea of automatically identifying and labeling frame-specific roles, as defined by the semantic frames, was first introduced by (Gildea and Jurafsky, 2002). Each semantic frame is characterized by a set of *target words* which can be nouns, verbs or adjectives. This helps abstracting the thematic roles and adding semantics to the given frame, highlighting the characteristic semantic features.

Frames are characterized by (1) target words or lexical predicates whose meaning includes aspects of the frame; (2) frame elements (FEs) which represent the semantic roles of the frame and (3) examples of annotations performed on the British National Corpus (BNC) for instances of each target word. Thus FrameNet frames are schematic representations of situations lexicalized by the target words (predicates) in which various participants and conceptual roles are related (the frame elements), exemplified by sentences from the BNC in which the target words and the frame elements are annotated.

In Senseval-3 two different cases of automatic labeling of the semantic roles were considered. The Unrestricted Case requires systems to assign FE labels to the test sentences for which (a) the boundaries of each frame element were given and the target words identified. The Restricted Case requires systems to (i) recognize the boundaries of the FEs

for each evaluated frame as well as to (ii) assign a label to it. Both cases can be cast as two different classifications: (1) a classification of the role when its boundaries are known and (2) a classification of the sentence words as either belonging to a role or not<sup>1</sup>.

A similar approach was used for automatically identifying predicate-argument structures in English sentences. The PropBank annotations ([www.cis.upenn.edu/~ace](http://www.cis.upenn.edu/~ace)) enable training for two distinct learning techniques: (1) decision trees (Surdeanu et al., 2003) and (2) Support Vector Machines (SVMs) (Pradhan et al., 2004). The SVMs produced the best results, therefore we decided to use the same learning framework for the Senseval-3 task for Automatic Labeling of Semantic Roles. Additionally, we have performed the following enhancements:

- we created a multi-class classifier for each frame, thus achieving improved accuracy and efficiency;
  - we combined some new features with features from (Gildea and Jurafsky, 2002; Surdeanu et al., 2003; Pradhan et al., 2004);
  - we resolved the data sparsity problem generated by limited training data for each frame, when using the examples associated with any other frame from FrameNet that had at least one FE shared with each frame that was evaluated;
  - we crafted heuristics that improved mappings from the syntactic constituents to the semantic roles.
- We believe that the combination of these four extensions are responsible for our results in Senseval-3.

The remainder of this paper is organized as follows. Section 2 describes our methods of classifying semantic roles whereas Section 3 describes our method of identifying role boundaries. Section 4 details our heuristics and Section 5 details the experimental results. Section 6 summarizes the conclusions.

---

<sup>1</sup>The second classification represents the detection of role boundaries. The semantic parsing defined as two different classification tasks was introduced in (Gildea and Jurafsky, 2002).

## 2 Semantic role classification

The result of the role classifier on a sentence, as illustrated in Figure 1, is the identification of semantic roles of the FEs when the boundaries of each FE are known. To be able to assign the labels of each FE, we used three sets of features. Feature Set 1, illustrated in Figure 2 was used in the work reported in (Gildea and Jurafsky, 2002).

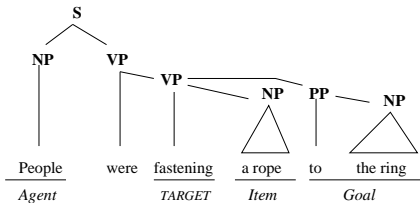


Figure 1: Sentence with annotated frame elements.

– PHRASE TYPE (pt): This feature indicates the syntactic type of the phrase labeled as a frame element, e.g. NP for Agent in Figure 1.
– PARSE TREE PATH (path): This feature contains the path in the parse tree between the predicate phrase and the target word, expressed as a sequence of nonterminal labels linked by direction symbols (up or down), e.g. NP ↑ S ↓ VP ↓ VP for Agent in Figure 1.
– POSITION (pos) – Indicates if the constituent appears before or after the the predicate in the sentence.
– VOICE (voice) – This feature distinguishes between active or passive voice for the predicate phrase.
– HEAD WORD (hw) – This feature contains the head word of the evaluated phrase. Case and morphological information are preserved.
– GOVERNING CATEGORY (gov) – This feature applies to noun phrases only, and it indicates if the NP is dominated by a sentence phrase (typical for subject arguments with active-voice predicates), or by a verb phrase (typical for object arguments).
– TARGET WORD – In our implementation this feature consists of two components: (1) WORD: the word itself with the case and morphological information preserved; and (2) LEMMA which represents the target normalized to lower case and infinitive form for the verbs or singular for nouns.

Figure 2: Feature Set 1 (FS1)

Feature Set 2 was introduced in (Surdeanu et al., 2003) and it is illustrated in Figure 3. The CONTENT WORD (cw) feature illustrated in Figure 3 applies to PPs, SBARs and VPs, as it was reported in (Surdeanu et al., 2003). For example, if the PP is “in the past month”, instead of using “in”, the head of the PP, as a feature, “month”, the head of the NP is selected since it is more informative. Similarly, if the SBAR is “that occurred yesterday”, instead of using the head “that” we select “occurred”, the head of the VP. When the VP “to be declared” is considered, “declared” is selected over “to”.

Feature set 3 is a novel set of features introduced in this paper and illustrated in Figure 4. Some of the new features characterize the frame, e.g. the frame name (FRAME-NAME); the frame FEs, (NUMBER-FEs); or the target word associated with

the frame (TARGET-TYPE). Additional characterization of the FEs are provided by by the GRAMMATICAL FUNCTION feature and by the list of grammatical functions of all FEs recognized in each sentence (LIST\_Grammatical\_Function feature).

– CONTENT WORD (cw) – Lexicalized feature that selects an informative word from the constituent, different from the head word.
PART OF SPEECH OF HEAD WORD (hPos) – The part of speech tag of the head word.
PART OF SPEECH OF CONTENT WORD (cPos) – The part of speech tag of the content word.
NAMED ENTITY CLASS OF CONTENT WORD (cNE) – The class of the named entity that includes the content word
BOOLEAN NAMED ENTITY FLAGS – A feature set comprising: <ul style="list-style-type: none"> <li>– neOrganization: set to 1 if an organization is recognized in the phrase</li> <li>– neLocation: set to 1 a location is recognized in the phrase</li> <li>– nePerson: set to 1 if a person name is recognized in the phrase</li> <li>– neMoney: set to 1 if a currency expression is recognized in the phrase</li> <li>– nePercent: set to 1 if a percentage expression is recognized in the phrase</li> <li>– neTime: set to 1 if a time of day expression is recognized in the phrase</li> <li>– neDate: set to 1 if a date temporal expression is recognized in the phrase</li> </ul>

Figure 3: Feature Set 2 (FS2)

In FrameNet, sentences are annotated with the name of the sub-corpus. There are 12,456 possible names of sub-corpus. For the 40 frames evaluated in Senseval-3, there were 1442 names associated with the example sentences in the training data and 2723 names in the test data. Three of the most frequent sub-corpus names are: “V-trans-other” (frequency=613), “N-all” (frequency=562) and “V-trans-simple”(frequency=560). The name of the sub-corpus indicates the relations between the target word and some of its FEs. For example, the “V-trans-other” name indicated that the target word is a transitive verb, and thus its FEs are likely to have other roles than object or indirect object. A sentence annotated with this sub-corpus name is: “*Night’s coming, you can see the black shadow on [Self-mover.the stones] that [TARGETrush] [Pathpast] and [Pathbetween your feet].*”. For this sentence both FEs with the role of Path are neither objects or indirect objects of the transitive verb.

Feature SUPPORT\_VERBS considers the usage of support expressions in FrameNet. We have found that whenever adjectives are target words, their semantic interpretation depends on their co-occurrence with verbs like “take”, “become” or “is”. Support verbs are defined as those verbs that combine with a state-noun, event-noun or state-adjective to create a verbal predicate, allowing arguments of the verb to serve as FEs of the frame evoked by the noun or the adjective.

The CORENESS feature takes advantage of a more recent implementation concept of core FEs

(vs. non-core FEs) in FrameNet. More specifically, the FrameNet developers classify frame elements in terms of how central they are to a particular frame, distinguishing three levels: core, peripheral and extra-thematic.

The features were used to produce two types of examples: positive and negative examples. For each FE of a frame, aside from the positive examples rendered by the annotations, we considered as negative examples all the annotations of the other FEs for the same frame. The positive and the negative examples were used for training the multi-class classifiers.

HUMAN: This feature indicates whether the syntactic phrase is either (1) a personal pronoun or (2) a hyponym of sense 1 of PERSON in WordNet
SUPPORT_VERBS that are recognized for adjective or noun target words have the role of predicate for the FEs. For example, if the target word is "clever" in the sentence "Smith is very clever, but he's no Einstein", the the FE "Smith" is an argument of the support verb "is" rather than of the target word. The values of this feature are either (1) The POS of the head of the VP containing the target word or (2) NULL if the target word does not belong to a VP
TARGET-TYPE: the lexical class of the target word, e.g. VERB, NOUN or ADJECTIVE
LIST_CONSTITUENT (FEs): This feature represents a list of the syntactic constituents covering each FE of the frame recognized in a sentence. For the example illustrated in Figure 1, the list is: [NP, NP, PP]
Grammatical Function: This feature indicates whether the FE is: - an External Argument (Ext) - an Object (Obj) - a Complement (Comp) - a Modifier (Mod) - Head noun modified by attributive adjective (Head) - Genitive determiner (Gen) - Appositive (Appos)
LIST_Grammatical_Function: This feature represents a list of the grammatical functions of the FEs recognized in the sentence.
NUMBER_FES: This feature indicates how many FEs were recognized in each sentence.
FRAME_NAME: This feature indicates the name of the semantic frame for which FEs are labeled
COVERAGE: This feature indicates whether there is a syntactic structure in the parse tree that perfectly covers the FE
CORENESS: This feature indicates whether the FE instantiates a conceptually necessary participant of a frame. For example, in the REVENGE frame, Punishment is a core element. The values are: (1) core; (2) peripheral and (3) extrathematic. FEs that mark notions such as Time, Place, Manner and Degree are peripheral. Extrathematic FEs situate an event against a backdrop of another event, by evoking a larger frame for which the target event fills a role.
SUB_CORPUS: In FrameNet, sentences are annotated with the name of the subcorpus they belong to. For example, for a verb target word, V-swh represents a subcorpus in which the target word is a predicate to a FE included in a relative clause headed by a wh-word.

Figure 4: Feature Set 3 (FS3)

Our multi-class classification allows each FE to be initially labeled with more than one role when several classifiers decide so. For example, for the ATTACHING frame, an FE may be labeled both as *Goal* and as *Item* if the classifiers for the *Goal* and *Item* select it as a possible role. To choose the final label, we select the classification which was assigned the

largest score by the SVMs.

PARSE TREE PATH WITH UNIQUE DELIMITER – This feature removes the direction in the path, e.g. VBN–VP–ADVP
PARTIAL PATH – This feature uses only the path from the constituent to the lowest common ancestor of the predicate and the constituent
FIRST WORD – First word covered by constituent
FIRST POS – POS of first word covered by constituent
LAST WORD – Last word covered by the constituent
LAST POS – POS of last word covered by the constituent
LEFT CONSTITUENT – Left sibling constituent label
LEFT HEAD – Left sibling head word
LEFT POS HEAD – Left sibling POS of head word
RIGHT CONSTITUENT – Right sibling constituent label
RIGHT HEAD – Right sibling head word
RIGHT POS HEAD – Right sibling POS of head word
PP PREP – If constituent is labeled PP get first word in PP
DISTANCE – Distance in the tree from constituent to the target word

Figure 5: Feature Set 4 (FS4)

### 3 Boundary Detection

The boundary detection of each FE was required in the Restricted Case of the Senseval-3 evaluation. To classify a word as belonging to an FE or not, we used all the entire Feature Set 1 and 2. From the Feature Set 3 we have used only four features: the Support-Verbs feature; the Target-Type feature, the Frame-Name feature and the Sub-Corpus feature. For this task we have also used Feature Set 4, which were first introduced in (Pradhan et al., 2004). The Feature Set 4 is illustrated in Figure 5. After the boundary detection was performed, the semantic roles of each FE were assigned using the role classifier trained for the Restricted Case

### 4 Heuristics

Frequently, syntactic constituents do not cover exactly FEs. For the Unrestricted Case we implemented a very simple heuristic: when there is no parse-tree node that exactly covers the target role  $r$  but a subset of adjacent nodes perfectly match  $r$ , we merge them in a new  $NP_{merge}$  node. For the Restricted Case, a heuristic for adjectival and nominal target words  $w$  adjoins consecutive nouns that are in the same noun phrase as  $w$ .

### 5 Experimental Results

In the Senseval-3 task for Automatic Labeling of Semantic Roles 24,558 sentences from FrameNet were assigned for training while 8,002 for testing. We used 30% of the training set (7367 sentences) as a *validation-set* for selecting SVM parameters that optimize accuracy. The number of FEs for

which labels had to be assigned were: 51,010 for the training set; 15,924 for the validation set and 16,279 for the test set. We used an additional set of 66,687 sentences (hereafter *extended data*) as extended data produced when using the examples associated with any other frame from FrameNet that had at least one FE shared with any of the 40 frames evaluated in Senseval-3. These sentences were parsed with the Collins’ parser (Collins, 1997). The classifier experiments were carried out using the SVM-light software (Joachims, 1999) available at <http://svmlight.joachims.org/> with a polynomial kernel<sup>2</sup> (degree=3).

### 5.1 Unrestricted Task Experiments

For this task we devised four different experiments that used four different combination of features: (1) FS1 indicates using only Feature Set 1; (2) +H indicates that we added the heuristics; (3) +FS2+FS3 indicates that we add the feature Set 2 and 3; and (4) +E indicates that the *extended data* has also been used. For each of the four experiments we trained 40 multi-class classifiers, (one for each frame) for a total of 385 binary role classifiers. The following Table illustrates the overall performance over the *validation-set*. To evaluate the results we measure the *F1*-score by combining the precision *P* with the recall *R* in the formula  $F1 = \frac{2 \times P \times R}{P + R}$ .

FS1	+H	+H+FS2+FS3	+H+FS2+FS3+E
84.4	84.9	91.7	93.1

### 5.2 Restricted Task Experiments

In order to find the best feature combination for this task we carried out some preliminary experiments over five frames. In Table 1, the row labeled B lists the *F1*-score of boundary detection over 4 different feature sets: FS1, +H, +FS4 and +E, the extended data. The row labeled R lists the same results for the whole Restricted Case.

Table 1: Restrictive experiments on *validation-set*.

	+FS1	+H	+H+FS2+FS3	+H+FS4+E
B	80.29	80.48	84.76	84.88
R	74.9	75.4	78	78.9

Table 1 illustrates the overall performance (boundary detection and role classification) of automatic semantic role labeling. The results listed in Tables 1 and 2 were obtained by comparing the FE boundaries identified by our parser with those annotated in FrameNet. We believe that these results are more

<sup>2</sup>In all experiments and for any classifier, we used the default SVM-light regularization parameter (e.g.,  $C = 1$  for normalized kernels) and a cost-factor  $j = 100$  to adjust the rate between Precision and Recall.

indicative of the performance of our systems than those obtained when using the scorer provided by Senseval-3. When using this scorer, our results have a precision of 89.9%, recall of 77.2% and an *F1*-score of 83.07% for the Restricted Case.

Table 2: Results on the *test-set*.

	Precision	Recall	<i>F1</i>
Unrestricted Case	94.5	90.6	92.5
Boundary Detection	87.3	75.1	80.7
Restricted Case	82.4	71.1	76.3

To generate the final Senseval-3 submissions we selected the most accurate models (for unrestricted and restricted tasks) of the validation experiments. Then we re-trained such models with all training data (i.e. our training plus validation data) and the setting (parameters, heuristics and extended data) derived over the *validation-set*. Finally, we run all classifiers on the *test-set* of the task. Table 2 illustrates the final results for both sub-tasks.

## 6 Conclusions

In this paper we describe a method for automatically labeling semantic roles based on support vector machines (SVMs). The training benefits from an extended data set on which multi-class classifiers were derived. The polynomial kernel of the SVMs enable the combination of four feature sets that produced very good results both for the Restricted Case and the Unrestricted Case. The paper also describes some heuristics for mapping syntactic constituents onto FEs.

## References

- Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of the COLING-ACL, Montreal, Canada*.
- Michael Collins. 1997. Three generative, lexicalized models for statistical parsing. In *Proceedings of the ACL-97*, pages 16–23.,
- Daniel Gildea and Daniel Jurafsky. 2002. Automatic labeling of semantic roles. *Computational Linguistic*, 28(3):496–530.
- T. Joachims. 1999. Making Large-Scale SVM Learning Practical. In B. Scholkopf, C. Burges, and MIT-Press. A. Smola (ed.), editors, *Advances in Kernel Methods - Support Vector Learning*.
- Sameer Pradhan, Kadri Hacioglu, Valeri Krugler, Wayne Ward, James H. Martin, and Daniel Jurafsky. 2004. Support vector learning for semantic argument classification. *Journal of Machine Learning Research*.
- Mihai Surdeanu, Sanda M. Harabagiu, John Williams, and John Aarseth. 2003. Using predicate-argument structures for information extraction. In *Proceedings of (ACL-03)*.