

Learning the Scope of Negation via Shallow Semantic Parsing

Junhui Li Guodong Zhou* Hongling Wang Qiaoming Zhu

School of Computer Science and Technology

Soochow University at Suzhou

{lijunhui, gdzhou, redleaf, qmzhu}@suda.edu.cn

Abstract

In this paper we present a simplified shallow semantic parsing approach to learning the scope of negation (SoN). This is done by formulating it as a shallow semantic parsing problem with the negation signal as the predicate and the negation scope as its arguments. Our parsing approach to SoN learning differs from the state-of-the-art chunking ones in two aspects. First, we extend SoN learning from the chunking level to the parse tree level, where structured syntactic information is available. Second, we focus on determining whether a constituent, rather than a word, is negated or not, via a simplified shallow semantic parsing framework. Evaluation on the BioScope corpus shows that structured syntactic information is effective in capturing the domination relationship between a negation signal and its dominated arguments. It also shows that our parsing approach much outperforms the state-of-the-art chunking ones.

1 Introduction

Whereas negation in predicate logic is well-defined and syntactically simple, negation in natural language is much complex. Generally, learning the scope of negation involves two subtasks: negation signal finding and negation scope finding. The former decides whether the words in a sentence are negation signals (i.e., words indicating negation, e.g., *no*, *not*, *fail*, *rather than*), where the semantic information of the words, rather than the syntactic information, plays a critical role. The latter determines the sequences of words in the sentence which are negated by the given negation signal. Compared with negation scope finding, negation signal finding is much simpler and has been well resolved in the literature, e.g. with

the accuracy of 95.8%-98.7% on the three subcorpora of the Bioscope corpus (Morante and Daelemans, 2009). In this paper, we focus on negation scope finding instead. That is, we assume golden negation signal finding.

Finding negative assertions is essential in information extraction (IE), where in general, the aim is to derive factual knowledge from free text. For example, Vincze et al. (2008) pointed out that the extracted information within the scopes of negation signals should either be discarded or presented separately from factual information. This is especially important in the biomedical domain, where various linguistic forms are used extensively to express impressions, hypothesized explanations of experimental results or negative findings. Szarvas et al. (2008) reported that 13.45% of the sentences in the abstracts subcorpus of the BioScope corpus and 12.70% of the sentences in the full papers subcorpus of the Bioscope corpus contain negative assertions. In addition to the IE tasks in the biomedical domain, SoN learning has attracted more and more attention in some natural language processing (NLP) tasks, such as sentiment classification (Turney, 2002). For example, in the sentence “*The chair is not comfortable but cheap*”, although both the polarities of the words “*comfortable*” and “*cheap*” are positive, the polarity of “the chair” regarding the attribute “*cheap*” keeps positive while the polarity of “the chair” regarding the attribute “*comfortable*” is reversed due to the negation signal “*not*”.

Most of the initial research on SoN learning focused on negated terms finding, using either some heuristic rules (e.g., regular expression), or machine learning methods (Chapman et al., 2001; Huang and Lowe, 2007; Goldin and Chapman, 2003). Negation scope finding has been largely ignored until the recent release of

* Corresponding author

the BioScope corpus (Szarvas et al., 2008; Vincze et al., 2008). Morante et al. (2008) and Morante and Daelemans (2009) pioneered the research on negation scope finding by formulating it as a chunking problem, which classifies the words of a sentence as being inside or outside the scope of a negation signal. However, this chunking approach suffers from low performance, in particular on long sentences, due to ignoring structured syntactic information. For example, given golden negation signals on the Bioscope corpus, Morante and Daelemans (2009) only got the performance of 50.26% in PCS (percentage of correct scope) measure on the full papers subcorpus (22.8 words per sentence on average), compared to 87.27% in PCS measure on the clinical reports subcorpus (6.6 words per sentence on average).

This paper explores negation scope finding from a parse tree perspective and formulates it as a shallow semantic parsing problem, which has been extensively studied in the past few years (Carreras and Màrquez, 2005). In particular, the negation signal is recast as the predicate and the negation scope is recast as its arguments. The motivation behind is that structured syntactic information plays a critical role in negation scope finding and should be paid much more attention, as indicated by previous studies in shallow semantic parsing (Gildea and Palmer, 2002; Punyakanok et al., 2005). Our parsing approach to negation scope finding differs from the state-of-the-art chunking ones in two aspects. First, we extend negation scope finding from the chunking level into the parse tree level, where structured syntactic information is available. Second, we focus on determining whether a constituent, rather than a word, is negated or not. Evaluation on the BioScope corpus shows that our parsing approach much outperforms the state-of-the-art chunking ones.

The rest of this paper is organized as follows. Section 2 reviews related work. Section 3 introduces the Bioscope corpus on which our approach is evaluated. Section 4 describes our parsing approach by formulating negation scope finding as a simplified shallow semantic parsing problem. Section 5 presents the experimental results. Finally, Section 6 concludes the work.

2 Related Work

While there is a certain amount of literature within the NLP community on negated terms finding (Chapman et al., 2001; Huang and Lowe, 2007; Goldin and Chapman, 2003), there are only a few studies on negation scope finding (Morante et al., 2008; Morante and Daelemans, 2009).

Negated terms finding

Rule-based methods dominated the initial research on negated terms finding. As a representative, Chapman et al. (2001) developed a simple regular expression-based algorithm to detect negation signals and identify medical terms which fall within the negation scope. They found that their simple regular expression-based algorithm can effectively identify a large portion of the pertinent negative statements from discharge summaries on determining whether a finding or disease is absent. Besides, Huang and Lowe (2007) first proposed some heuristic rules from a parse tree perspective to identify negation signals, taking advantage of syntactic parsing, and then located negated terms in the parse tree using a corresponding negation grammar.

As an alternative to the rule-based methods, various machine learning methods have been proposed for finding negated terms. As a representative, Goldin and Chapman (2003) adopted both Naïve Bayes and decision trees to distinguish whether an observation is negated by the negation signal “*not*” in hospital reports.

Negation scope finding

Morante et al. (2008) pioneered the research on negation scope finding, largely due to the availability of a large-scale annotated corpus, the Bioscope corpus. They approached the negation scope finding task as a chunking problem which predicts whether a word in the sentence is inside or outside of the negation scope, with proper post-processing to ensure consecutiveness of the negation scope. Morante and Daelemans (2009) further improved the performance by combing several classifiers.

Similar to SoN learning, there are some efforts in the NLP community on learning the scope of speculation. As a representative, Özgür and Radev (2009) divided speculation

learning into two subtasks: speculation signal finding and speculation scope finding. In particular, they formulated speculation signal finding as a classification problem while employing some heuristic rules from the parse tree perspective on speculation scope finding.

3 Negation in the BioScope Corpus

This paper employs the BioScope corpus (Szarvas et al., 2008; Vincze et al., 2008)¹, a freely downloadable negation resource from the biomedical domain, as the benchmark corpus. In this corpus, every sentence is annotated with negation signals and speculation signals (if it has), as well as their linguistic scopes. Figure 1 shows a self-explainable example. In this paper, we only consider negation signals, rather than speculation ones. Our statistics shows that 96.57%, 3.23% and 0.20% of negation signals are represented by one word, two words and three or more words, respectively. Additional, adverbs (e.g., *not*, *never*) and determiners (e.g., *no*, *neither*) occupy 45.66% and 30.99% of negation signals, respectively.

```
<sentence id="S26.8">These findings <xcope
id="X26.8.2"><cue type="speculation"
ref="X26.8.2">indicate that</cue> <xcope
id="X26.8.1">corticosteroid resistance in bron-
chial asthma <cue type="negation"
ref="X26.8.1">can not</cue> be explained by
abnormalities in corticosteroid receptor charac-
teristics</xcope></xcope>.</sentence>
```

Figure 1: An annotated sentence in the BioScope corpus.

The Bioscope corpus consists of three sub-corpora: the full papers and the abstracts from the GENIA corpus (Collier et al., 1999), and clinical (radiology) reports. Among them, the full papers subcorpus and the abstracts subcorpus come from the same genre, and thus share some common characteristics in statistics, such as the number of words in the negation scope to the right (or left) of the negation signal and the average scope length. In comparison, the clinical reports subcorpus consists of clinical radiology reports with short sentences. For detailed statistics about the three subcorpora, please see Morante and Daelemans (2009).

¹ <http://www.inf.u-szeged.hu/rgai/bioscope>

For preprocessing, all the sentences in the Bioscope corpus are tokenized and then parsed using the Berkeley parser² (Petrov and Klein, 2007) trained on the GENIA TreeBank (GTB) 1.0 (Tateisi et al., 2005)³, which is a bracketed corpus in (almost) PTB style. 10-fold cross-validation on GTB1.0 shows that the parser achieves the performance of 86.57 in F1-measure. It is worth noting that the GTB1.0 corpus includes all the sentences in the abstracts subcorpus of the Bioscope corpus.

4 Negation Scope Finding via Shallow Semantic Parsing

In this section, we first formulate the negation scope finding task as a shallow semantic parsing problem. Then, we deal with it using a simplified shallow semantic parsing framework.

4.1 Formulating Negation Scope Finding as a Shallow Semantic Parsing Problem

Given a parse tree and a predicate in it, shallow semantic parsing recognizes and maps all the constituents in the sentence into their corresponding semantic arguments (roles) of the predicate. As far as negation scope finding considered, the negation signal can be regarded as the predicate⁴, while the scope of the negation signal can be mapped into several constituents which are negated and thus can be regarded as the arguments of the negation signal. In particular, given a negation signal and its negation scope which covers $word_m, \dots, word_n$, we adopt the following two heuristic rules to map the negation scope of the negation signal into several constituents which can be deemed as its arguments in the given parse tree.

- 1) The negation signal itself and all of its ancestral constituents are non-arguments.
- 2) If constituent X is an argument of the given negation signal, then X should be the highest constituent dominated by the scope of $word_m, \dots, word_n$. That is to say, X 's parent constituent must cross-bracket or include the scope of $word_m, \dots, word_n$.

² <http://code.google.com/p/berkeleyparser/>

³ <http://www-tsujii.is.s.u-tokyo.ac.jp/GENIA>

⁴ If a negation signal consists of multiply words (e.g., rather than), the last word (e.g., than) is chosen to represent the negation signal.

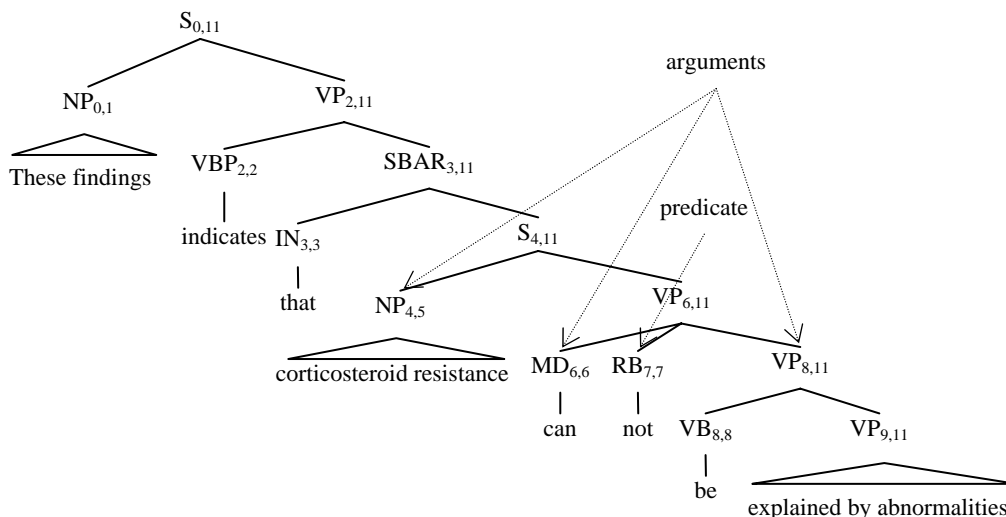


Figure 2: An illustration of a negation signal and its arguments in a parse tree.

The first rule ensures that no argument covers the negation signal while the second rule ensures no overlap between any two arguments. For example, in the sentence “*These findings indicate that corticosteroid resistance can not be explained by abnormalities*”, the negation signal “*can not*” has the negation scope “*corticosteroid resistance can not be explained by abnormalities*”. As shown in Figure 2, the node “*RB_{7,7}*” (i.e., *not*) represents the negation signal “*can not*” while its arguments include three constituents {*NP_{4,5}*, *MD_{6,6}*, and *VP_{8,11}*}. It is worth noting that according to the above rules, negation scope finding via shallow semantic parsing, i.e. determining the arguments of a given negation signal, is robust to some variations in parse trees. This is also empirically justified by our later experiments. For example, if the *VP_{6,11}* in Figure 2 is incorrectly expanded by the rule $VP_{6,11} \rightarrow MD_{6,6} + RB_{7,7} + VB_{8,8} + VP_{9,11}$, the negation scope of the negation signal “*can not*” can still be correctly detected as long as {*NP_{4,5}*, *MD_{6,6}*, *VB_{8,8}*, and *VP_{9,11}*} are predicted as the arguments of the negation signal “*can not*”.

Compared with common shallow semantic parsing which needs to assign an argument with a semantic label, negation scope finding does not involve semantic label classification and thus could be divided into three consequent phases: argument pruning, argument identification and post-processing.

4.2 Argument Pruning

Similar to the predicate-argument structures in common shallow semantic parsing, the negation signal-scope structures in negation scope finding can be also classified into several certain types and argument pruning can be done by employing several heuristic rules to filter out constituents, which are most likely non-arguments of a negation signal. Similar to the heuristic algorithm as proposed in Xue and Palmer (2004) for argument pruning in common shallow semantic parsing, the argument pruning algorithm adopted here starts from designating the negation signal as the current node and collects its siblings. It then iteratively moves one level up to the parent of the current node and collects its siblings. The algorithm ends when it reaches the root of the parse tree. To sum up, except the negation signal and its ancestral constituents, any constituent in the parse tree whose parent covers the given negation signal will be collected as argument candidates. Taking the negation signal node “*RB_{7,7}*” in Figure 2 as an example, constituents {*MD_{6,6}*, *VP_{8,11}*, *NP_{4,5}*, *IN_{3,3}*, *VBP_{2,2}*, and *NP_{0,1}*} are collected as its argument candidates consequently.

4.3 Argument Identification

Here, a binary classifier is applied to determine the argument candidates as either valid arguments or non-arguments. Similar to argument

identification in common shallow semantic parsing, the structured syntactic information plays a critical role in negation scope finding.

Basic Features

Table 1 lists the basic features for argument identification. These features are also widely used in common shallow semantic parsing for both verbal and nominal predicates (Xue, 2008; Li et al., 2009).

Feature	Remarks
b1	Negation: the stem of the negation signal, e.g., not, rather_than. (<i>can_not</i>)
b2	Phrase Type: the syntactic category of the argument candidate. (<i>NP</i>)
b3	Path: the syntactic path from the argument candidate to the negation signal. (<i>NP<S>VP>RB</i>)
b4	Position: the positional relationship of the argument candidate with the negation signal. “left” or “right”. (<i>left</i>)

Table 1: Basic features and their instantiations for argument identification in negation scope finding, with $NP_{4,5}$ as the focus constituent (i.e., the argument candidate) and “*can not*” as the given negation signal, regarding Figure 2.

Additional Features

To capture more useful information in the negation signal-scope structures, we also explore various kinds of additional features. Table 2 shows the features in better capturing the details regarding the argument candidate and the negation signal. In particular, we categorize the additional features into three groups according to their relationship with the argument candidate (AC, in short) and the given negation signal (NS, in short).

Some features proposed above may not be effective in argument identification. Therefore, we adopt the greedy feature selection algorithm as described in Jiang and Ng (2006) to pick up positive features incrementally according to their contributions on the development data. The algorithm repeatedly selects one feature each time which contributes most, and stops when adding any of the remaining features fails to improve the performance. As far as the negation scope finding task concerned, the whole feature selection process could be done by first running the selection algorithm with the basic features (b1-b4) and then incrementally picking up effective features from (ac1-ac6, AC1-AC2,

ns1-ns4, NS1-NS2, nsac1-nsac2, and NSAC1-NSAC7).

Feature	Remarks
argument candidate (AC) related	
ac1	the headword (ac1H) and its POS (ac1P). (<i>resistance, NN</i>)
ac2	the left word (ac2W) and its POS (ac2P). (<i>that, IN</i>)
ac3	the right word (ac3W) and its POS (ac3P). (<i>can, MD</i>)
ac4	the phrase type of its left sibling (ac4L) and its right sibling (ac4R). (<i>NULL, VP</i>)
ac5	the phrase type of its parent node. (<i>S</i>)
ac6	the subcategory. (<i>S:NP+VP</i>)
combined features (AC1-AC2)	
b2&fc1H, b2&fc1P	
negation signal (NS) related	
ns1	its POS. (<i>RB</i>)
ns2	its left word (ns2L) and right word (ns2R). (<i>can, be</i>)
ns3	the subcategory. (<i>VP:MD+RB+VP</i>)
ns4	the phrase type of its parent node. (<i>VP</i>)
combined features (NS1-NS2)	
b1&ns2L, b1&ns2R	
NS-AC-related	
nsac1	the compressed path of b3: compressing sequences of identical labels into one. (<i>NP<S>VP>RB</i>)
nsac2	whether AC and NS are adjacent in position. “yes” or “no”. (<i>no</i>)
combined features (NSAC1-NSAC7)	
b1&b2, b1&b3, b1&nsac1, b3&NS1, b3&NS2, b4&NS1, b4&NS2	

Table 2: Additional features and their instantiations for argument identification in negation scope finding, with $NP_{4,5}$ as the focus constituent (i.e., the argument candidate) and “*can not*” as the given negation signal, regarding Figure 2.

4.4 Post-Processing

Although a negation signal in the BioScope corpus always has only one continuous block as its negation scope (including the negation signal itself), the negation scope finder may result in discontinuous negation scope due to independent prediction in the argument identification phase. Given the golden negation signals, we observed that 6.2% of the negation scopes predicted by our negation scope finder are discontinuous.

Figure 3 demonstrates the projection of all the argument candidates into the word level. According to our argument pruning algorithm in Section 4.2, except the words presented by

the negation signal, the projection covers the whole sentence and each constituent (LAC_i or RAC_j in Figure 3) receives a probability distribution of being an argument of the given negation signal in the argument identification phase.

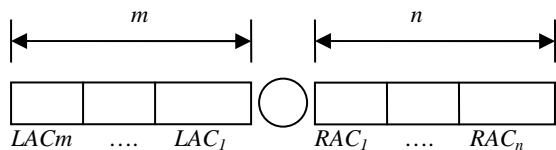


Figure 3: Projecting the left and the right argument candidates into the word level.

Since a negation signal is deemed inside of its negation scope in the BioScope corpus, our post-processing algorithm first includes the negation signal in its scope and then starts to identify the left and the right scope boundaries, respectively.

As shown in Figure 3, the left boundary has $m+1$ possibilities, namely the negation signal itself, the leftmost word of constituent LAC_i ($1 \leq i \leq m$). Supposing LAC_i receives probability of P_i being an argument, we use the following formula to determine LAC_{k^*} whose leftmost word represents the boundary of the left scope. If $k^*=0$, then the negation signal itself represents its left boundary.

$$k^* = \arg \max_k \prod_{i=1}^k P_i * \prod_{i=k+1}^m (1 - P_i)$$

Similarly, the right boundary of the given negation signal can be decided.

5 Experimentation

We have evaluated our shallow semantic parsing approach to negation scope finding on the BioScope corpus.

5.1 Experimental Settings

Following the experimental setting in Morante and Daelemans (2009), the abstracts subcorpus is randomly divided into 10 folds so as to perform 10-fold cross validation, while the performance on both the papers and clinical reports subcorpora is evaluated using the system trained on the whole abstracts subcorpus. In addition, SVMlight⁵ is selected as our classifier. In particular, we adopt the linear kernel and the training parameter C is fine-tuned to 0.2.

⁵ <http://svmlight.joachims.org/>

The evaluation is made using the accuracy. We report the accuracy using three measures: *PCLB* and *PCRB*, which indicate the percentages of correct left boundary and right boundary respectively, *PCS*, which indicates the percentage of correct scope as a whole.

5.2 Experimental Results on Golden Parse Trees

In order to select beneficial features from the additional features proposed in Section 4.3, we randomly split the abstracts subcorpus into training and development datasets with proportion of 4:1. After performing the greedy feature selection algorithm on the development data, features {NSAC5, ns2R, NS1, ac1P, ns3, NSAC7, ac4R} are selected consecutively for argument identification. Table 3 presents the effect of selected features in an incremental way on the development data. It shows that the additional features significantly improve the performance by 11.66% in PCS measure from 74.93% to 86.59% ($\chi^2; p < 0.01$).

Feature	PCLB	PCRB	PCS
Baseline	84.26	88.92	74.93
+NSAC5	90.96	88.92	81.34
+ns2R	91.55	88.92	81.92
+NS1	92.42	89.50	83.09
+ac1P	93.59	89.50	84.26
+ns3	93.88	90.09	84.84
+NSAC7	94.75	89.80	85.42
+ac4R	95.04	90.67	86.59

Table 3: Performance improvement (%) of including the additional features in an incremental way on the development data (of the abstracts subcorpus).

However, Table 3 shows that the additional features behave quite differently in terms of PCLB and PCRB measures. For example, PCLB measure benefits more from features NSAC5, ns2R, NS1, ac1P, and NSAC7 while PCRB measure benefits more from features NS1 and ac4R. It also shows that the features (e.g., NSAC5, ns2R, NS1, NSAC7) related to neighboring words of the negation signal play a critical role in recognizing both left and right boundaries. This may be due to the fact that neighboring words usually imply sentential information. For example, “*can not be*” indicates a passive clause while “*did not*” indicates an active clause. Table 3 also shows that the recognition of left boundaries is much easier than that of right boundaries. This may be due

to the fact that 83.6% of negation signals have themselves as the left boundaries in the abstracts subcorpus.

Table 4 presents the performance on the abstracts subcorpus by performing 10-fold cross-validation. It shows that the additional features significantly improve the performance over the three measures ($\chi^2; p < 0.01$).

Feature	PCLB	PCRB	PCS
Baseline	84.29	87.82	74.05
+selected features	93.06	88.96	83.10

Table 4: Performance (%) of negation scope finding on the abstracts subcorpus using 10-fold cross-validation.

5.3 Experimental Results on Automatic Parse Trees

The GTB1.0 corpus contains 18,541 sentences in which 11,850 of them (63.91%) overlap with the sentences in the abstracts subcorpus⁶. In order to get automatic parse trees for the sentences in the abstracts subcorpus, we train the Berkeley parser with the remaining 6,691 sentences in GTB1.0. The Berkeley parser trained on 6,691 sentences achieves the performance of 85.22 in F1-measure on the other sentences in GTB1.0. For both the full papers and clinical reports subcorpora, we get their automatic parse trees by using two Berkeley parsers: one trained on 6,691 sentences in GBT1.0, and the other trained on all the sentences in GTB1.0.

To test the performance on automatic parse trees, we employ two different configurations. First, we train the argument identification classifier on the abstracts subcorpus using *automatic parse trees* produced by Berkeley parser trained on 6,691 sentences. The experimental results are presented in the rows of *autoparse(t&t)* in Table 5 and Table 6. Then, we train the argument identification classifier on the abstracts subcorpus using *golden parse trees*. The experimental results are presented in the rows of *autoparse(test)* in Table 5 and Table 6.

We also report an oracle performance to explore the best possible performance of our system by assuming that our negation scope finder can always correctly determine whether a candidate is an argument or not. That is, if an ar-

gument candidate is outside or cross-brackets with the golden negation scope, then it is a non-argument. The oracle performance is presented in the rows of *oracle* in Table 5 and Table 6.

Table 5 and Table 6 show that:

- 1) Automatic syntactic parsing lowers the performance of negation scope finding on the abstracts subcorpus in all three measures (e.g. from 83.10 to 81.84 in PCS). As expected, the parser trained on the whole GTB1.0 corpus works better than that trained on 6,691 sentences (e.g. 64.02 Vs. 62.70, and 89.79 Vs. 85.21 in PCS measure on the full papers and the clinical reports subcorpora, respectively). However, the performance decrease shows that negation scope finding is not as sensitive to automatic syntactic parsing as common shallow semantic parsing, whose performance might decrease by about ~10 in F1-measure (Toutanova et al., 2005). This indicates that negation scope finding via shallow semantic parsing is robust to some variations in the parse trees.
- 2) *autoparse(test)* consistently outperforms *autoparse(t&t)* on both the abstracts and the full papers subcorpora. However, it is surprising to find that *autoparse(t&t)* achieves better performance on the clinical reports subcorpus than *autoparse(test)*. This may be due to the special characteristics of the clinical reports subcorpus, which mainly consists of much shorter sentences with 6.6 words per sentence on average, and better adaptation of the argument identification classifier to the variations in the automatic parse trees.
- 3) The performance on all three subcorpora indicates that the recognition of right boundary is much harder than that of left boundary. This may be due to the longer right boundary on an average. Our statistics shows that the average left/right boundaries are 1.1/6.9, 0.1/3.7, and 1.2/6.5 words on the abstracts, the full papers and the clinical reports subcorpora, respectively.
- 4) The oracle performance is less sensitive to automatic syntactic parsing. In addition, given the performance gap between the performance of our negation scope finder and the oracle performance, there is still much room for further performance improvement.

⁶ There are a few cases where two sentences in the abstracts subcorpus map into one sentence in GTB.

	Abstracts			Papers			Clinical		
	PCLB	PCRB	PCS	PCLB	PCRB	PCS	PCLB	PCRB	PCS
autoparse(t&t)	91.97	87.82	80.88	85.45	67.20	59.26	97.48	88.30	85.89
autoparse(test)	92.71	88.33	81.84	87.57	68.78	62.70	97.48	87.73	85.21
oracle	99.72	94.59	94.37	98.94	84.13	83.33	99.89	98.39	98.39

Table 5: Performance (%) of negation scope finding on the three subcorpora by using automatic parser trained with 6,691 sentences in GTB1.0.

	Papers			Clinical		
	PCLB	PCRB	PCS	PCLB	PCRB	PCS
autoparse(t&t)	85.98	67.99	60.32	97.48	92.66	90.48
autoparse(test)	87.83	70.11	64.02	97.36	92.20	89.79
oracle	98.94	83.86	83.07	99.77	97.94	97.82

Table 6: Performance (%) of negation scope finding on the two subcorpora by using automatic parser trained with all the sentences in GTB1.0.

Method	Abstracts	Papers	Clinical
M et al. (2008)	57.33	n/a	n/a
M & D (2009)	73.36	50.26	87.27
Our baseline	73.42	53.70	88.42
Our final system	81.84	64.02	89.79

Table 7: Performance comparison over the PCS measure (%) of our system with other state-of-the-art ones.

Table 7 compares our performance in PCS measure with related work. It shows that even our baseline system with four basic features as presented in Table 1 performs better than Morante et al. (2008) and Morante and Daelemans(2009). This indicates the appropriateness of our simplified shallow semantic parsing approach and the effectiveness of structured syntactic information on negation scope finding. It also shows that our final system significantly outperforms the state-of-the-art ones using a chunking approach, especially on the abstracts and full papers subcorpora. However, the improvement on the clinical reports subcorpus is less apparent, partly due to the fact that the sentences in this subcorpus are much simpler (with average length of 6.6 words per sentence) and thus a chunking approach can achieve high performance. Following are two typical sentences from the clinical reports subcorpus, where the negation scope covers the whole sentence (except the period punctuation). Such sentences account for 57% of negation sentences in the clinical reports subcorpus.

- (1) No evidence of focal pneumonia .
- (2) No findings to account for symptoms .

6 Conclusion

In this paper we have presented a simplified shallow semantic parsing approach to negation scope finding by formulating it as a shallow semantic parsing problem, which has been extensively studied in the past few years. In particular, we regard the negation signal as the predicate while mapping the negation scope into several constituents which are deemed as arguments of the negation signal. Evaluation on the Bioscope corpus shows the appropriateness of our shallow semantic parsing approach and that structured syntactic information plays a critical role in capturing the domination relationship between a negation signal and its negation scope. It also shows that our parsing approach much outperforms the state-of-the-art chunking ones. To our best knowledge, this is the first research on exploring negation scope finding via shallow semantic parsing.

Future research will focus on joint learning of negation signal and its negation scope findings. Although Morante and Daelemans (2009) reported the performance of 95.8%-98.7% on negation signal finding, it lowers the performance of negation scope finding by about 7.29%-16.52% in PCS measure.

Acknowledgments

This research was supported by Projects 60683150, 60970056, and 90920004 under the National Natural Science Foundation of China, Project 20093201110006 under the Specialized Research Fund for the Doctoral Program of Higher Education of China.

References

- Xavier Carreras and Lluís Màrquez. 2005. Introduction to the CoNLL-2005 Shared Task: Semantic Role Labeling. In *Proceedings of CoNLL 2005*.
- Wendy W. Chapman, Will Bridewell, Paul Hanbury, Gregory F. Cooper, and Bruce G. Buchanan. 2001. A Simple Algorithm for Identifying Negated Findings and Diseases in Discharge Summaries. *Journal of Biomedical Informatics*, 34: 301-310.
- Nigel Collier, Hyun Seok Park, Norihiro Ogata, et al. 1999. The GENIA project: corpus-based knowledge acquisition and information extraction from genome research papers. In *Proceedings of EACL 1999*.
- Daniel Gildea and Martha Palmer. 2002. The Necessity of Parsing for Predicate Argument Recognition. In *Proceedings of ACL 2002*.
- Ilya M. Goldin and Wendy W. Chapman. 2003. Learning to Detect Negation with 'Not' in Medical Texts. In *Proceedings of SIGIR 2003*.
- Yang Huang and Henry Lowe. 2007. A Novel Hybrid Approach to Automated Negation Detection in Clinical Radiology Reports. *Journal of the American Medical Informatics Association*, 14(3): 304-311.
- Zheng Ping Jiang and Hwee Tou Ng. 2006. Semantic Role Labeling of NomBank: A Maximum Entropy Approach. In *Proceedings of EMNLP 2006*.
- Junhui Li, Guodong Zhou, Hai Zhao, Qiaoming Zhu, and Peide Qian. Improving Nominal SRL in Chinese Language with Verbal SRL Information and Automatic Predicate Recognition. In *Proceedings of EMNLP 2009*.
- Roser Morante, Anthony Liekens, and Walter Daelemans. 2008. Learning the Scope of Negation in Biomedical Texts. In *Proceedings of EMNLP 2008*.
- Roser Morante and Walter Daelemans. 2009. A Metalearning Approach to Processing the Scope of Negation. In *Proceedings of CoNLL 2009*.
- Arzucan Özgür; Dragomir R. Radev. 2009. Detecting Speculations and their Scopes in Scientific Text. In *Proceedings of EMNLP 2009*.
- Slav Petrov and Dan Klein. 2007. Improved Inference for Unlexicalized Parsing. In *Proceedings of NAACL 2007*.
- Vasin Punyakanok, Dan Roth, and Wen-tau Yih. 2005. The Necessity of Syntactic Parsing for Semantic Role Labeling. In *Proceedings of IJCAI 2005*.
- György Szarvas, Veronika Vincze, Richárd Farkas, and János Csirik. 2008. The BioScope corpus: annotation for negation, uncertainty and their scope in biomedical texts. In *Proceedings of BioNLP 2008*.
- Yuka Tateisi, Akane Yakushiji, Tomoko Ohta, and Jun'ichi Tsujii. 2005. Syntax Annotation for the GENIA Corpus. In *Proceedings of IJCNLP 2005, Companion volume*.
- Kristina Toutanova, Aria Haghighi, and Christopher D. Manning. 2005. Joint Learning Improves Semantic Role Labeling. In *Proceedings of ACL 2005*.
- Peter D. Turney. 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In *Proceedings of ACL 2002*.
- Veronika Vincze, György Szarvas, Richárd Farkas, György Móra, and János Csirik. 2008. The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes. *BMC Bioinformatics*, 9(Suppl 11):S9.
- Nianwen Xue and Martha Palmer. 2004. Calibrating Features for Semantic Role Labeling. In *Proceedings of EMNLP 2004*.
- Nianwen Xue. 2008. Labeling Chinese Predicates with Semantic Roles. *Computational Linguistics*, 34(2):225-255.