Negation and Speculation Target Identification

Bowei Zou Guodong Zhou Qiaomin Zhu

Natural Language Processing Lab, School of Computer Science and Technology
Soochow University, Suzhou, 215006, China
zoubowei@gmail.com, {gdzhou, qmzhu}@suda.edu.cn

Abstract. Negation and speculation are common in natural language text. Many applications, such as biomedical text mining and clinical information extraction, seek to distinguish positive/factual objects from negative/speculative ones (i.e., to determine what is negated or speculated) in biomedical texts. This paper proposes a novel task, called negation and speculation target identification, to identify the target of a negative or speculative expression. For this purpose, a new layer of the target information is incorporated over the BioScope corpus and a machine learning algorithm is proposed to automatically identify this new information. Evaluation justifies the effectiveness of our proposed approach on negation and speculation target identification in biomedical texts.

Keywords: negation; speculation; target identification

1 Introduction

Negative and speculative expressions are common in natural language text. While negation is a grammatical category which comprises various kinds of devices to reverse the truth value of a proposition, speculation is a grammatical category which expresses the attitude of a speaker towards a statement in terms of degree of certainty, reliability, subjectivity, sources of information, and perspective. It is widely accepted that negation and speculation play a critical role in natural language understanding, especially information extraction from biomedical texts. Szarvas (2008) observes that a significant proportion of the gene names mentioned in a corpus of biomedical articles appear in speculative sentences (638 occurrences out of a total of 1,968). Morante and Sporleder (2012) state that in order to automatically extract reliable information from clinical reports, it is of great importance to determine whether symptoms, signs, treatments, outcomes, or any other clinical relevant factors are present or not.

Recent studies of negation and speculation on biomedical information extraction focus on trigger detection, which aims to detect the signal of a negative or speculative expression, and scope resolution, which aims to determine the linguistic coverage of a negative or speculative trigger in a sentence, in distinguishing unreliable or uncertain information from facts. For example, sentences (1) and (2) include a negative trigger and a speculative trigger, respectively, both denoted in **boldface** with their scopes denoted in square brackets (adopted hereinafter).

adfa, p. 1, 2011.

(1) *Our results show that* [***no*** *transcription of the RAG-1 gene could be detected*].

(2) [*The cardiovascular disease* ***may*** *recur*] *even after cure.*

However, people may wonder what is exactly negated or speculated, e.g. which objects are negated or speculated on a clinical medicine, by which for an authority to make a proper action. This poses strong requirements beyond trigger detection and scope resolution.

From this regard, we propose a novel task, called negation and speculation target identification, to extract the object targeted by a negative or speculative expression. For example, in sentences (1) and (2), the targets are *transcription of the RAG-1 gene* and *the cardiovascular disease*, given the negative and speculative expressions ***no*** and ***may***, respectively (denoted with underline). The main contributions of this paper are the proposal of a new task, the annotation of a new corpus and the proposal of a machine learning approach for such a new task. It is also worthy to mention that negation and speculation target identification can not only complement trigger detection and scope resolution but also help them better infer unreliable or uncertain information from context. For example, in the scenario of sentence (1), if sentence *Some studies claimed that transcription of the RAG-1 gene was detected* is given as its context, we can easily infer that the affirmative statement in the given context is doubtful.

The rest of this paper is organized as follows. Section 2 introduces the related work. In Section 3, we discuss some details on negation and speculation target identification. In Section 4, the annotation guidelines for the target identification corpus are introduced. In Section 5, our machine learning approach is proposed with various kinds of lexical and syntactic features. Section 6 reports the experimental results and gives some discussions. Finally, we draw the conclusion in Section 7.

2 Related Work

There is a certain amount of literature within the natural language processing community on negation and speculation. While earlier studies adopt rule-based approaches (e.g., Light et al., 2004), machine learning-based approaches begin to dominate the research on negation and speculation (e.g., Morante et al., 2008) since the release of the BioScope corpus (Vincze et al., 2008).

Recently, the studies on negation and speculation have been drawing more and more attention, such as the CoNLL'2010 Shared Task on trigger detection and scope resolution of negation and speculation (Farkas et al., 2010), and the ACL 2010 Workshop on negation recognition (Morante and Sporleder, 2010). Even more, a special issue of Computational Linguistics (Morante and Sporleder, 2012) has been published on negation and speculation. However, none of the above shared tasks or workshops aim at identifying the target of a negative or speculative expression.

Similar to target identification in biomedical texts, opinion target extraction (OTE) in sentiment analysis aims to identify the topics on which an opinion is expressed (Pang and Lee, 2008). Nevertheless, above opinion target is related to a sentiment word instead of a negative or speculative expression. Among others, in semantic role labeling (Carreras and Màrquez, 2005), a target may act as some semantic role. However, such correspondence does not always exist since a target is dominated by a negative or speculative expression, while a semantic role is dominated by a predicate.

Even though the studies on negation and speculation have received much interest in the past few years, open access annotated resources are rare, usually with limitation in information and small scale in size. For example, the Hedge Classification corpus (Medlock and Briscoe, 2007) only contains the annotation for hedge triggers in 1537 sentences and does not contain the scope information. The BioScope corpus (Vincze et al., 2008) annotates the linguistic scopes of negative and speculative triggers in biomedical texts. Obviously, none of above resources is suitable for negation and speculation target identification.

3 Target of Negation and Speculation

A negative and speculative expression always attaches to an object or its attribute which is negated or speculated. In this paper, we define such an object as the target of negation or speculation (except particular illustration, we use "target" for simplicity).

According to above definition of target, it seems that almost all targets should be entities. However, the fact is that some predicates can be also negated or speculated by a verbal negation or speculation expression. For example, in sentence (2), the cardiovascular disease is the speculation target, meaning whether this disease could recur. Here, the direct speculative object is an event (recur). In such a situation, we consider the agent of a negative or speculative expression as the target. Statistics on 100 samples randomly chosen from our target identification corpus (For details please refer to Section 4) shows that only 42 targets are entities, while the remaining 58 targets are the agents of verbal negative or speculative expressions. To better illustrate the concept of target, we clarify its difference with scope and subject.

Target vs. Scope: Both scope and target are extremely important to capture the negative and speculative meanings. While scope refers to the grammatical part in a sentence that is negated or speculated, target is concerned with the negative or speculative object rather than the grammatical coverage of a negative or speculative cue. For example, in sentence (1), while the scope is *no transcription of the RAG-1 gene could be detected*, representing a negative proposition, the target is *transcription of the RAG-1 gene*, representing a negated object. In addition, it should be noted that a target is not always in a scope. For example, in sentence (3), while the scope *without voting* negates the evaluating way for prize, *The Prize of Best Employee* is target.

(3) <u>*The Prize of Best Employee*</u> *is awarded* [***without*** *voting*], *unexpectedly.*

(4) *Company management has* ***not*** *yet decided on* <u>*the Prize of Best Employee*</u>.

Target vs. Subject: The target and the subject in a sentence may not be the same. A target represents the object described by a negative or speculative expression, while a subject is a constituent that conflates nominative case with the topic. The former is from semantic perspective on negation and speculation, while the latter is from the syntactic perspective. In sentences (3) and (4), both the negated targets are *the Prize of Best Employee*, no matter whether or not they are the subjects in a sentence.

4 Corpus

Due to the lack of corpus annotation for target identification on negation and speculation, a new layer of the target information is added to the BioScope corpus (Vincze et al., 2008)[1], a freely available resource which has already been annotated with the linguistic scopes of negative and speculative triggers in biomedical texts.

4.1 Annotation Guidelines

In BioScope corpus, only the sentences including the speculative or negative information are chosen for target annotation. In annotation, the most general two basic guidelines are: 1) if a noun phrase can be inferred as the object described by a negative or speculative expression, it is the target. 2) Otherwise, target is the agent of the sentence concluding negative or speculative trigger. During the process of annotation, more than 70% of sentences can be annotated by the two basic guidelines.

In the following, we introduce the specific guidelines developed throughout the annotation process with examples to deal with the specific characteristics in target identification on negation and speculation.

Guideline 1: In sentence (5), the target should be something (maybe drug or therapy), but does not appear in the sentence. In this situation, we annotated *it* as the target, since this paper is only concerned with the target in a sentence, and as for what it is actually, there is no need for annotation.

(5) *It is **not** effective for all tuberculosis patients.*

Guideline 2: When there is a raising verb (e.g., seem, appear, be expected, be likely, etc.) in a sentence, as in sentence (6), we prefer to mark the logical agent as the target rather than the formal one.

(6) *It **seems** that the treatment is successful.*

Guideline 3: A target can be partly determined on the basis of syntax. Our manual statistics on syntactic category shows that noun phrases exist in 97.59% and 98.14% of the targets on negation and speculation respectively. Besides, in the annotation process, we extend their scopes to the biggest syntactic unit as much as possible due to following two facts:

First, taking into account the information integrality of a target, it seems better to include all the elements attached to the target, such as prepositional phrases, determiners, adjectives, and so on. In sentence (7), with *blood lymphocytes* as the head word of the target, the two prepositional phrases (introduced by *from* and *with*) that represent the target's attributes are also included within the target:

(7) *In contrast, blood lymphocytes from patients with granulomatous diseases have **little** effect on children.*

Second, the status of a modifier is sometimes uncertain. For example, the negative trigger ***no*** in sentence (8) could modify two different semantic elements: On one hand, it may modify *primary*, with the meaning *the glucocorticoid metabolism is*

---

[1] http://www.inf.u-szeged.hu/rgai/bioscope

impaired. On the other hand, it may modify *impairment*, with the meaning *there is no impairment of the glucocorticoid metabolism at all*. We cannot resolve such ambiguity on the basis of contextual information. Fortunately, we can avoid such ambiguity with the maximal length annotation strategy. Furthermore, if the category of target could be directed further fine by a modifier, the target should contain the modifier.

(8) *There is* ***no*** *<u>primary impairment of glucocorticoid metabolism</u> in the asthmatics.*

Guideline 4: When the trigger is a conjunction, we extend the target all members of the coordination.

(9) *In common sense, symptoms include <u>fever, cough</u>* ***or*** *<u>itches</u>.*

Guideline 5: If a target contains an omitted part, for simplicity, we avoid completing it.

(10) *Finally, recombinant GHF-1 interacted directly with c-Jun proteins but* ***not*** *<u>c-Fos</u>.*

Guideline 6: If there are punctuation marks or conjunctions at the head or end of a target, we ignore them. Nevertheless, for coherence, the punctuation marks or conjunctions in the middle of the target are included (see sentence (9)).

4.2 Corpus Annotation

We have annotated 1,668 negative sentences and 2,678 speculative sentences by two independent annotators following the guidelines in Section 4.1 over the BioScope corpus. Table 1 summarizes the chief characteristics of the corpus.

Table 1. Statistics of target identification corpus.

		Negation	Speculation
#Sentence		1668	2678
%In scope		54.98%	63.71%
%Out of scope		45.02%	36.29%
Average length of sentences		29.73	31.16
Average length of targets		4.36	5.27
Relation to target and keyword	%Before	54.80%	43.05%
	%After	45.02%	49.48%
%Noun phrase target		98.02%	97.46%

During the annotation process, annotators can only refer to the negative or speculative triggers but not their corresponding scopes. This is necessary to ensure that the annotation is not biased by scope information provided of BioScope corpus. The 3rd and 4th rows in Table 1 show the ratio of the cases whether the targets are involved in the scope. It indicates that the targets are not always in the scope.

In target identification corpus, if there is more than one trigger in a sentence, we treat them as different instances. The 7th and 8th rows show the ratio of the cases whether the target position is in front of the trigger or behind it. Such close ratios show that the positional relation between the trigger and its target is not apparent. Additionally, according to our statistics, there are 97.59% and 98.14% of the targets including a noun phrase on negation and speculation respectively (the 9th row in Table 1). This is the reason that we regard the noun phrases as target candidates in our experiments.

The annotators are not allowed to communicate with each other until the annotation process is finished, but they could appeal to linguists when needed. Differences between the two annotated results are also resolved by linguists.

A checking step can ensure that the annotation is grammatical. In this step, every instance has been processed with a syntactic parser (refer to Section 6.1). If the maximal syntactic parsing sub-tree of all target terminal nodes has other terminal nodes, the annotation system would require annotator to confirm. For example in Figure 1, the least common ancestor of "*cotransfection, TCF-1*" is the syntactic category node "*S*", but "*S*" includes other terminal nodes (e.g., "*Upon*"). This kind of instance would be re-annotated. About 17% of negative instances and 12% of speculative instances are re-annotated.

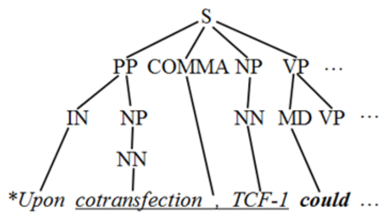

Fig. 1. The instance that needs to re-annotate.

We measured the consistency level of the annotation by inter-annotator agreement analysis. The Cohen's kappa statistic (Conger, 1980) for our annotation is 0.83. After the careful examination of the disagreements by linguists, they are resolved. The main conflict is whether the modifiers of the target (mentioned in Section 4.1) are involved.

5 Methodology

We propose a machine learning approach with various kinds of lexical and syntactic features for negation and speculation target identification.

Features. By taking the noun phrase in sentence as target candidate, we adopt ranking Support Vector Machine (rSVM) model for training. To capture more useful information, we propose various kinds of refined features from lexical and syntactic perspective. Table 2 lists these features.

Features (1-5) are the basic information of negative/speculative trigger and target candidate, including trigger itself and its part-of-speech (POS), candidate itself, the head word of candidate and its POS. For feature S1 and S2, if the trigger or its POS are different in two instances, the corresponding targets are likely to act different roles in sentence. Considering the triggers in sentence (11) and (12), for trigger *hypothesis* and its POS *NN*, target is more likely present in predicative; for trigger *assumes* and its POS *VBZ*, target is likely to be the subject of its clause. And likewise, the head word features are also informative for target.

Features (6-9) are the syntactic relationship between target candidate and negative/speculative trigger, including syntactic path, relative position, distance in paring tree, and distance of tokens. For feature S6, the syntactic path feature provides structural information of parsing tree, but it is sparse when candidate is far away from the trigger. Although feature S7 is simple, it relates the voice of a verbal trigger. For example, considering the verbal speculative triggers *be suggesting* and *be suggested*, the role of target is likely to be different.

Table 2. Full set of lexical and syntactic features on target identification.

No.	Feature	Explanation
S1	keyword	keyword itself
S2	keyword_POS	keyword's part-of-speech
S3	candidate	candidate itself
S4	headword	candidate's head word
S5	headword_POS	part-of-speech of candidate's head word
S6	syn_path_keyword	syntactic path between keyword and candidate
S7	PR_keyword	positional relationship of candidate with keyword
S8	syn_dis_keyword	syntactic distance from candidate to keyword
S9	word_dis_keyword	word distance from candidate to keyword
S10	left_phrase_type_tag	left sibling tag of candidate's syntactic category
S11	right_phrase_type_tag	right sibling tag of candidate's syntactic category
S12	left_phrase_type_seq	sequence of words govern by left sibling of candidate's syntactic category
S13	right_phrase_type_seq	sequence of words govern by right sibling of candidate's syntactic category
S14	nearest_verb	nearest verb with keyword in syntactic parsing tree
S15	PR_SF11	positional relationship of candidate with SF11 verb
S16	syn_dis_SF11	syntactic distance from candidate to SF11 verb
S17	word_dis_SF11	word distance from candidate to SF11 verb
S18	syn_path_SF11	syntactic path between SF11 verb and candidate
C1	S1 + S14	
C2	S7 + S15	
C3	S8 + S16	
C4	S9 + S17	

Features (10-13) are the adjacent syntactic features of target candidate, including the left and right syntactic categories, and the left and right chunks. Intuitively, these features are sparse and featureless on lexical level (S12 and S13), but not on syntactic level (S10 and S11).

Feature (14-18) are the syntactic information associated with the verb, which may have the directly relatedness between negative/speculative trigger and its corresponding target. Motivating in part by semantic role labeling (SRL), we infer that features related to verb in sentences are effective for target identification. That is because, in SRL, the predicate verb involves lots of dominated and modified relationship between itself and other semantic roles. Similarly, the negative or speculative triggers are closely connected with the verb on syntactic structures. For this reason, we explore the features from the verb for getting more supplementary syntactic information.

Since above features may not work on target identification of both negation and speculation with equal effectiveness, we adopt a greedy feature selection algorithm as described in Jiang et al. (2006) to pick up positive features incrementally according to their contributions on the performance of our system. 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.

Post-Processing. As mentioned in Section 4.1, a target is the object described by a negative or speculative trigger. According to the annotation guidelines, we adopt the maximal principle to label targets, which involve some modified structures, such as prepositional phrase and attributive clause. Our classification system takes noun phrases as instances, but in fact, some syntactic structures of targets are NP+PP but not NP. In that case, we cannot get correct results. For this reason, we propose a post-processing step to improve performance, described in Algorithm 1 below. In post-processing, if the syntactic category of prediction has a right sibling of PP or SBAR, we connect the sibling to the prediction and continue to check the rest.

Input:
    syntactic parsing tree: T,
    prediction node: N_{pred},
Output:
    word sequence of N_{pred}: W_{pred}

Initialize:
    W_{pred} = get_sequence(N_{pred});
    $N_{sibling} \leftarrow$ get_right_sibling(N_{pred});
while $N_{sibling} \neq$ NULL
    if $N_{sibling}$ = "PP" or "SBAR" **then**
        $W_{pred} \leftarrow W_{pred}$ + get_sequence($N_{sibling}$);
    end
    $N_{sibling} \leftarrow$ get_right_sibling($N_{sibling}$);
End

Algorithm 1. Post-processing algorithm.

6 Experimentation

6.1 Experimental Settings

Dataset: In consideration of the features selection, we have split the corpus into 5 equal parts, within which 2 parts are used for feature selection and the rest for experiments. On the one hand, in feature selection, we divide the data into 5 equal parts, within which 4 parts for training and the rest for developing. We divide the experimental data into 10 folds randomly, so as to perform 10-fold cross validation.

Syntactic parser: All sentences in our corpus are tokenized and parsed using the Berkeley Parser (Petrov et al, 2007)[2] which have been trained on the GENIA TreeBank 1.0 (abbr., GTB; Tateisi et al., 2005)[3], a bracketed corpus with PTB style in

[2] http://code.google.com/p/berkeleyparser
[3] http://www-tsujii.is.s.u-tokyo.ac.jp/GENIA

biomedical field. 10-fold cross-validation on GTB1.0 shows that the syntactic parser achieves 87.12% in F1-measure.

Classifier: We selected the SVM^Light4 with the default parameters configuration as our classifier.

Evaluation metrics: Exact match is used to evaluate the correctness of a target (Accuracy, abbr., Acc). That is to say, an extracted target is considered as correct only if it has exactly the same span boundaries as the annotated ones in gold standard. Additionally, we adopt Precision (P), Recall (R), and F1-measure (F1) as evaluation metrics. The accuracy which takes sentence as a unit measures the performance of our system. The PRF-measure which takes target candidate as a unit reports the performance of the binary classifier by which every instance has been classified.

6.2 Results of Target Identification

Performance of baselines. We implemented four baselines to measure the difficulty of the target identification task:

- ***Baseline_First***: select the first noun phrase in sentence as target.
- ***Baseline_Last***: select the last noun phrase in sentence as target.
- ***Baseline_Longest***: select the longest noun phrase in sentence as target.
- ***Baseline_Nearest:*** select the noun phrase which is nearest to the trigger as target. The distance is measured by syntactic path. For example, in *NP>S<VP<VBN*, the distance from *NP* to *VBN* is 3.

Table 3 lists performances of baseline systems without post-processing. It shows that the performances of *Baseline_Longest* and *Baseline_Nearest* are higher than the other two systems. The two former baselines do not consider the relationship between trigger and target, which is direct clue for target identification. However, the *Baseline_Longest* system adopts no information involving trigger either, but its performance improves a little. We infer that the longest noun phrase in a sentence involves many modifiers which are always the object most impacted by the trigger. For both negation and speculation target identification, the *Baseline_Nearest* system achieves the best performance. It indicates that the syntactic path characteristics are effective to detect the target dominated by trigger. Inspired by the *Baseline_Nearest* system, we employ some syntactic path features in our classification.

Table 4 shows the effectiveness of our post-processing algorithm described in Section 5.3. All of baselines greatly improve by the post-processing algorithm. Besides, it is worth noting that the Baseline_Longest system only improves of less than 2 and 3 in accuracy on negation and speculation, respectively, largely due to the completeness of the longest noun phrase in a sentence.

---

[4] http://svmlight.joachims.org

Table 3. Accuracy of baselines.

System	Neg	Spe
Baseline_First	17.68	23.37
Baseline_Last	21.04	18.29
Baseline_Longest	30.70	32.64
Baseline_Nearest	31.95	36.63

Table 4. Accuracy of post-processing.

System	Neg	Spe
Baseline_First	21.44	25.37
Baseline_Last	27.41	24.19
Baseline_Longest	32.13	35.59
Baseline_Nearest	39.93	43.77

Our Performance. We perform a greedy algorithm as described in Section 5 to select a set of effective syntactic features on Feature Selection dataset. Table 5 and 6 show the effects of selected features in an incremental way for negation and speculation respectively. We also employ all of the features described in Table 2 for target classification.

Table 5. Performance improvement of features incrementally on negation.

Features	P	R	F	Acc
S6	45.29	33.64	38.61	40.31
+ S3	51.77	46.98	49.26	50.02
+ C2	60.40	51.19	55.41	57.06
+ S11	66.58	56.79	61.30	61.21
+ S8	71.18	58.44	64.18	63.92
+ C1	73.51	60.08	66.12	64.39
ALL	68.01	55.37	61.04	59.23

Table 5 shows that the system with feature set of {S6, S3, C2, S11, S8, and C1} achieves the best performance. Table 6 shows that the system the feature set of {S6, S3, S11, S9, S8, S15, and S7} achieves the best performance.

Table 6. Performance improvement of features incrementally on speculation.

Features	P	R	F	Acc
S6	59.38	48.86	53.16	54.19
+ S3	68.41	55.91	61.53	64.23
+ S11	73.28	59.50	65.67	67.36
+ S9	75.42	62.64	68.42	68.84
+ S8	76.63	62.98	69.14	69.09
+ S15	77.21	63.26	69.54	69.35
+ S7	78.03	63.55	70.05	69.37
ALL	76.16	59.48	66.79	67.99

It is worth noting that the features S6, S3, S11 and S8 are effective both on negation and speculation. Feature S6 directly represents the connecting pathway between trigger and target. For instance, on negation, corresponding a preposition trigger (e.g., “*without*”), the path is usually “*PP<NP*”. On speculation, if the trigger is a modal verb (e.g., “*might*”), the syntactic path between trigger and target probably is “*NP>S<VP*” or “*VP<NP*”. Feature S3 is the target candidate itself. In the same topic or discourse of a literature, the depicted target is likely to be concentrated. If a target is negated or speculated, the one in other sentences may also have the same semantic representation (negation or speculation).

adfa, p. 10, 2011.

Additionally, features S11 and S8 also have a little effect for target classification on both negation and speculation. In our corpus, 11.8% of negative triggers of the total are "*no* (det)" and 4.7% of negative triggers are "*without*". This kind of triggers usually takes a right sibling as their targets. Thus, feature S11 can dig the characteristics in this situation. Similar to feature S6, feature S8 also represents the syntactic relatedness between trigger and target.

Features C2 and C1 are the particular features on negation target identification. They are related to trigger and its nearest verb in syntactic parsing tree. It shows that the position of target suffers from the combined impact of trigger and its corresponding verb.

Features S9, S15, and S7 are the particular features on speculation target identification. Similar to feature S8, feature S9 is another kind of distance between trigger and target. Features S15 and S7 are the target candidate's position to verb and trigger respectively.

Table 7. Performance of target identification system on negation and speculation.

	P	R	F	Acc
Negation	76.27	63.53	69.32	70.13
Speculation	84.32	69.85	76.41	74.46

Table 7 shows the performance of our target identification system with post-processing. It significantly improves the accuracy by 5.74 from 64.39 to 70.13 on negation (p<0.05) and by 5.09 from 69.37 to 74.46 on speculation (p<0.05). It indicates that not all targets are noun phrases and the post-processing algorithm we proposed is instrumental.

7 Conclusion

In this paper, we propose target identification on negation and speculation, a novel task on negation and speculation in biomedical texts. Due to the lack of corpus, we add a new layer of the target information over the BioScope corpus. On the basis, a set of features are depicted and a supervised model is proposed to implement target identification on negation and speculation. The experimental results show that syntactic features play a critical role in capturing the domination relationship between a negative or speculative trigger and its target.

In future work, we will finalize and release the corpus and explore more useful features for target identification on negation and speculation. Moreover, we will systematically explore its application in other domains, e.g., legal or socio-political genre.

Acknowledgments

This research is supported by the National Natural Science Foundation of China, No.61272260, No.61331011, No.61273320, the Natural Science Foundation of Jiangsu Province, No. BK2011282, and the Major Project of College Natural Science

Foundation of Jiangsu Province, No.11KIJ520003. The authors would like to thank the anonymous reviewers for their insightful comments and suggestions.

References

1. Xavier Carreras and Lluís Màrquez. 2005. Introduction to the CoNLL-2005 Shared Task: Semantic Role Labeling. *In Proceedings of the 9th Conference on Computational Natural Language Learning*, pages 152-164.
2. Anthony J. Conger. 1980. Integration and generalization of kappas for multiple raters. *In Psychological Bulletin*, 88(2), pages 322-328.
3. Richárd Farkas, Veronika Vincze, György Móra, János Csirik, and György Szarvas. 2010. The CoNLL-2010 Shared Task: Learning to Detect Hedges and their Scope in Natural Language Text. *In Proceedings of the Fourteenth Conference on Computational Natural Language Learning*, pages 1-12, Uppsala, Sweden.
4. Zhengping Jiang and Hwee T. Ng. 2006. Semantic Role Labeling of NomBank: A Maximum Entropy Approach. *In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pages 138-145.
5. Marc Light, Xinying Qiu, Padmini Srinivasan. 2004. The Language of Bioscience: Facts, Peculations, and Statements in Between. *In Proceedings of the HLT BioLINK'2004*, pages 17-24.
6. Ben Medlock and Ted Briscoe. 2007. Weakly supervised learning for hedge classification in scientific literature. *In Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 992-999, Prague.
7. Roser Morante, Anthony Liekens, and Walter Daelemans. 2008. Learning the scope of negation in biomedical texts. *In Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pages 715-724, Honolulu, HI.
8. Roser Morante and Caroline Sporleder, editors. 2010. Proceedings of the Workshop on Negation and Speculation in Natural Language Processing. University of Antwerp, Uppsala, Sweden.
9. Roser Morante and Caroline Sporleder. 2012. Modality and Negation: An Introduction to the Special Issue. *Computational Linguistics*, 38(2), pages 223-260.
10. Bo Pang and Lillian Lee. 2008. Opinion Mining and Sentiment Analysis: Foundations and Trends. *Information Retrieval*, 2(12), pages 1-135.
11. Slav Petrov and Dan Klein. 2007. Improved Inference for Unlexicalized Parsing. *In Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics*, pages 404-411.
12. Gyorgy Szarvas. 2008. Hedge classification in biomedical texts with a weakly supervised selection of keywords. *In Proceedings of the 46th Annual Meeting of the Association of Computational Linguistics*, pages 281-289.
13. Yuka Tateisi, Akane Yakushiji, Tomoko Ohta, and Jun'ichi Tsujii. 2005. Syntax Annotation for the GENIA Corpus. *In Proceedings of IJCNLP, Companion volume*, pages 222-227.
14. Veronika Vincze, Gyorgy Szarvas, Richard Farkas, Gyorgy Mora, and Janos Csirik. 2008. The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes. *BMC Bioinformatics*, 9(Suppl 11):S9+.