

# UABCoRAL: A Preliminary study for Resolving the Scope of Negation

Binod Gyawali, Thamar Solorio

CoRAL Lab

Department of Computer and Information Sciences

University of Alabama at Birmingham

Birmingham, Alabama, USA

{bgyawali, solorio}@cis.uab.edu

## Abstract

This paper describes our participation in the closed track of the \*SEM 2012 Shared Task of finding the scope of negation. To perform the task, we propose a system that has three components: negation cue detection, scope of negation detection, and negated event detection. In the first phase, the system creates a lexicon of negation signals from the training data and uses the lexicon to identify the negation cues. Then, it applies machine learning approaches to detect the scope and negated event for each negation cue identified in the first phase. Using a preliminary approach, our system achieves a reasonably good accuracy in identifying the scope of negation.

## 1 Introduction

All human language samples, either written or spoken, contain some information in negated form. In tasks such as information retrieval, sometimes, we should consider only the positive information of an event and disregard its negation information, and vice versa. For example, while searching for the patients with diabetes, we should not include a patient who has a clinical report saying *No symptoms of diabetes were observed*. Thus, finding the negation and its scope is important in tasks where the negation and assertion information need to be treated differently. However, most of the systems developed for processing natural language data do not consider negations present in the sentences. Although various works (Morante et al., 2008; Morante and Daelemans, 2009; Li et al., 2010; Councill et al., 2010;

Apostolova et al., 2011) have dealt with the identification of negations and their scope in sentences, this is still a challenging task.

The first task in \*SEM 2012 Shared Task (Morante and Blanco, 2012) is concerned with finding the scope of negation. The task includes identifying: i) negation cues, ii) scope of negation, and iii) negated event for each negation present in the sentences. Negation cue is a word, part of a word, or a combination of words that carries the negation information. Scope of negation in a sentence is the longest group of words in the sentence that is influenced by the negation cue. Negated event is the shortest group of words that is actually affected by the negation cue. In Example (1) below, word *no* is a negation cue, the discontinuous word sequences ‘I gave him’ and ‘sign of my occupation’ are the scopes, and ‘gave’ is the negated event.

(1) I [gave] him *no* sign of my occupation.

In this paper, we propose a system to detect the scope of negation for the closed track of \*SEM 2012 Shared Task. Our system uses a combination of a rule based approach, and a machine learning approach. We use a rule based approach to create a lexicon of all the negation words present in the training data. Then we use this lexicon to detect the negation cues present in the test data. We do a preliminary analysis of finding the scope of negation and the negated events by applying a machine learning approach, and using basic features created from the words, lemmas, and parts-of-speech (POS) tags of words in the sentences. The F-measure scores

achieved by our system are about 85% for negation cue detection, 65% in full scope identification, 48% in negated event detection, and 39% in identifying full negation. Our error analysis shows that the use of lexicon is not very appropriate to detect the negation cues. We also describe the challenges in identifying the scope and the negated events.

## 2 Problem Description

The \*SEM 2012 shared task competition provided three data sets: training, development, and test data set. Each sentence in each data set is split into words. The dataset contains the information such as lemma, part of speech, and other syntactic information of each word. Each sentence of training and development data is annotated with negation cues, scopes and negated events. Using the training and the development data, the task is to identify negation cues, scopes and negated events in all unannotated sentences of the test data.

Sentence tokens	Negation cue	Scope	Negated event
I	-	I	-
am	-	am	-
not	not	-	-
sure	-	sure	sure
whether	-	whether	-
I	-	I	-
left	-	left	-
it	-	it	-
here	-	here	-

Table 1: An example of negation cue, scope and the negated event

A sentence can contain more than one negation cue. Negation cues in the data set can be i) a single word token such as *n't*, *nowhere*, ii) a continuous sequence of two or more words, such as *no more*, *by no means* or iii) two or more discontinuous words such as *..neither...nor...* A negation cue is either a part or same as its corresponding negation word. This corresponding negation word is referred as a negation signal in the remaining sections of the paper. For example, for a negation signal *unnecessary*, the negation cue is *un*, and similarly, for a negation signal *needless*, the negation cue is *less*.

Scope of a negation in a sentence can be a continuous sequence of words or a discontinuous set of words in the sentence. Scope of negation sometimes includes the negation word. A negation word may not have a negated event. Presence of a negated event in a sentence depends upon the facts described by the sentence. Non-factual sentences such as interrogative, imperative, and conditional do not contain negated events. Morante and Daelemans (2012) describe the details of the negation cue, scope, and negated event, and the annotation guidelines. An example of the task is shown in Table 1.

## 3 System Description

We decompose the system to identify the scope of negation into three tasks. They are:

1. Finding the negation cue
2. Finding the scope of negation
3. Finding the negated event

The scope detection and the negated event detection tasks are dependent on the task of finding the negation cue. But the scope detection and the negated event detection tasks are independent of each other.

We identify the negation cues present in the test data based on a lexicon of negation signals that are present in the training and the development data. The tasks of identifying scope of negation and negated event are modeled as classification problems. To identify scope and negated event, we train classifiers with the instances created from the training data provided. We create test instances from the test data annotated with negation cues predicted by our cue detection component. Due to the use of test data annotated by our cue detection component, the false negative rate in predicting the negation cues is propagated to the scope detection as well as negated event detection components. The details of all the three components are described in the subsections below.

### 3.1 Identifying the negation cue

In this task, we identify all the negation cues present in the sentences. We group the negation cues under three types depending upon how they are present in the data. They are: single word cues, continuous

multiword cues, and discontinuous multiword cues. All the cues present in the training and development datasets are shown in Table 2.

Cue types	Cues
Single word cues	absence, dis, except, fail, im, in, ir, less, n't, neglected, neither, never, no, nobody, none, nor, not, nothing, nowhere, prevent, refused, save, un, without
Continuous multiword cues	no more, rather than, by no means, nothing at all, on the contrary, not for the world
Discontinuous multiword cues	neither nor, no nor, not not

Table 2: Negation cues present in training and development data

In the training and development data, multiword negation cues account for only 1.40% of the total negation cues. At this stage, we decided to focus on identifying the single word negation cues. The system first creates a lexicon that contains the pairs of negation cues and their corresponding negation signals for all the single word negation cues present in the training and the development datasets. In order to identify a negation cue in the test set, the system searches all the words in the sentences of the test data that match the negation signals of the lexicon. For each word that matches, it assigns the corresponding cue of the signal from the lexicon as its negation cue.

### 3.2 Identifying the scope of negation

We apply a machine learning technique to identify the scope of negation. For each negation cue present in a sentence, we create the problem instances as the tuple of the negation signal and each word present in the same sentence. To create the instances, we use only those sentences having at least one negation. For training, we create instances from the training data, but we consider only those words that are within a window of size 20 from the negation signal and within the sentence boundary. We restricted the words to be within the window in order to minimize the problem of imbalanced data. This window was chosen following our observation that only 1.26% of the scope tokens go beyond the 20 word window from the negation signal. Including the words

beyond this window causes a major increase in the negative instances resulting in a highly imbalanced training set. While creating test instances, we do not restrict the words by window size. This restriction is not done in order to include all the words of the sentences in the test instances. An instance is labeled as positive if the word used to create the instance is the scope of the negation signal; else it is labeled as negative.

We extract 10 features to identify the scope of negation as follows:

1. Negation signal in the tuple
2. Lemma of the negation signal
3. POS tag of the negation signal
4. Word in the tuple
5. Lemma of the word in the tuple
6. POS tag of the word in the tuple
7. Distance between the negation signal and the word in terms of number of words
8. Position of the word from the negation signal (left, right)
9. Whether a punctuation character (‘,’; ‘:’; ‘;’) exists between the word and the negation signal
10. Sequence of POS tags in between the negation signal and the word

After the classification, if an instance is predicted as positive, the word used to create the instance is considered as the scope of the negation signal. If a negation signal has prefix such as ‘dis’, ‘un’, ‘in’, ‘ir’, or ‘im’, the scope of negation includes only the part of word (signal) excluding the prefix. Thus, for each negation signal having these prefix, we remove the prefix from the signal and consider the remaining part of it as the scope, regardless of whether the classifier classifies the instance pair as positive or negative.

### 3.3 Identifying the negated event

The task of identifying the negated event is similar to the task of identifying the scope of negation. The process of creating the instances for this task is almost the same to that of finding the scope of negation, except that, we limit the window size to 4 words from the negation signal. 4.24% of the negated events lie away from the 4 word window. Beyond this window, the events are very sparse and a small increment in the window size leads to abrupt increase in negative instances and creates an imbalance in the data. The 4 word window size was selected based on the best result obtained among various experiments performed with different window sizes greater than and equal to 4. The same rule applies while creating instances for training data as well as test data. We use only nine features in this step, excluding the 9<sup>th</sup> feature used in the scope detection. We also apply the same rule of mapping the negation signals starting with ‘dis’, ‘un’, ‘in’, ‘ir’, and ‘im’ to the negated event as in the previous step.

## 4 Experimental Settings

We evaluated our system only on the test data of the shared task. For the machine learning tasks, we used the SVM light classifier (Joachims, 1999) with 4<sup>th</sup> degree polynomial kernel and other default parameters. The identification of cues, scopes, negated events, and full negation are evaluated on the basis of the F-measures. We also use ‘B’ variant for cues, scopes, negated events and the full negation for evaluation. The precision of ‘B’ variant is calculated as the ratio of true positives to the system count. Identification of cues and negated events are measured independent of any other steps. But the identification of the scopes is measured depending upon the correct identification of cues in three different ways as follows:

- i) scopes (cue match): the cue has to be correct for the scope to be correct
- ii) scopes (no cue match): the system must identify part of the cue for the scope to be correct
- iii) scope tokens (no cue match): a part of the system identified cue must overlap with the gold standard cue for the scope tokens to be correct

The F1 score of the full negation detection was

used to rank the systems of the participants. The details about the evaluation measures can be found in Morante and Blanco (2012).

## 5 Results Analysis

The results obtained by our system over the test data are shown in Table 3. The results obtained by each component, and their analysis are described in the subsections below.

### 5.1 Identifying the negation cues

The system is able to achieve an 85.77% F1 score in the task of identifying the negation cues using a simple approach based on the lexicon of the negation signals. Because of the system’s inability to identify multiword negation cues, it could not detect the multiword cues such as *..neither..nor..*, *..absolutely nothing..*, *..far from..*, *..never more..*, that account for 3.5% of the total negation cues present in the test data.

The accuracy of the system is limited by the coverage of the lexicon. Due to the low coverage of the lexicon, the system fails to identify signals such as *ceaseless*, *discoloured*, *incredulity*, *senseless*, and *unframed* that are present only in the test data. These signals account for 4.5% of the total negation signals present in the test data. Some words such as *never*, *nothing*, *not*, *n't*, *no*, and *without* are mostly present as the negation signals in the data. But these words are not always the negation signals. The phrase *no doubt* is present nine times in the test data, but the word *no* is a negation signal in only four of them. This accounts for 1.89% error in the negation cue detection. The word *save* is present once as a negation signal in the training data, but it is never a negation signal in the test data. Therefore, our lexicon based system invariably predicts two occurrences of *save* in the test data as negation signals.

### 5.2 Identifying the scope of negation

The system achieves 63.46% F1 score in identifying scopes with cue match, 64.76% F1 score in identifying scopes with no cue match, and 76.23% F1 score in identifying scope tokens with no cue match. The results show that our system has a higher precision than recall in identifying the scope. As mentioned

	gold	system	tp	fp	fn	precision (%)	recall (%)	F1 (%)
Cues	264	284	226	37	38	85.93	85.61	85.77
Scopes (cue match)	249	239	132	35	117	79.04	53.01	63.46
Scopes (no cue match)	249	239	132	35	113	79.53	54.62	64.76
Scope tokens (no cue match)	1805	1456	1243	213	562	85.37	68.86	76.23
Negated (no cue match)	173	104	65	35	104	65.00	38.46	48.33
Full negation	264	284	73	37	191	66.36	27.65	39.04
Cues B	264	284	226	37	38	79.58	85.61	82.48
Scopes B (cue match)	249	239	132	35	117	55.23	53.01	54.10
Scopes B (no cue match)	249	239	132	35	113	56.90	54.62	55.74
Negated B (no cue match)	173	104	65	35	104	62.50	38.46	47.62
Full negation B	264	284	73	37	191	25.70	27.65	26.64
Total sentences: 1089								
Negation sentences: 235								
Negation sentences with errors: 172								
% Correct sentences: 81.73								
% Correct negation sentences: 26.81								

Table 3: Results of the system

earlier, the negation cues identified in the first task are used to identify the scope of negation and the negated events. Using the test data with 15% error in negation cues as the input to this component and some of the wrong predictions of the scope by this component led to a low recall value in the scope detection.

The results show that the system works well when a negation signal has fewer scope tokens and when the scope tokens are closer to the negation signal. There are some cases where the system could not identify the scope tokens properly. It is unable to detect the scope tokens that are farther in distance from the negation signals. The system is not performing well in predicting the discontinuous scopes. When a negation cue has discontinuous scope, mostly the system predicts one sequence of words correctly but could not identify the next sequence. In sentence (2) in the example below, the underlined word sequences are the discontinuous scopes of the negation cue *not*. In the sentence, our system predicts only the second sequence of scope, but not the first sequence. In some cases, our system does not have a good coverage of scope tokens. In sentence (3), the underlined word sequence is the scope of the signal *no*, but our system detects only *at ninety was hardship* as its scope. These inability to detect the full scope have led to have a higher accuracy in predicting the partial scope tokens (76.23%) than predicting the full scope (64.76%).

(2) the box is a half pound box of honeydew tobacco and does not help us in any way

(3) ...a thermometer at ninety was no hardship

(4) ...I cannot see anything save very vague indications

Analyzing the results, we see that the error in predicting the scope of the negation is high when the scope is distributed in two different phrases. In the example (2) above, *does not help us in any way* is a single verb phrase and all the scope within the phrase is correctly identified by our system. *The box* being a separate phrase, it is unable to identify it. However, in some cases such as example (4), the system could not identify any scope tokens for negation cue *not*.

Some of the findings of previous works have shown that the features related to syntactic path are helpful in identifying the scope of negation. Li et al. (2010) used the syntactic path from the word to the negation signal and showed that this helped to improve the accuracy of scope detection. Similarly, work by Councill et al. (2010) showed that the accuracy of scope detection could be increased using the features from the dependency parse tree. In our experiment, there was a good improvement in the scope detection rate when we included “sequence of POS tags” between the negation signal and the word as a feature. This improvement after including the sequence of POS tags feature and its consistency

with the previous works implies that adding path related features might help to improve the accuracy in scope detection.

### 5.3 Identifying the negated event

We are able to achieve an F1 score of 48.33% in predicting the negated events, which is the lowest score among all three components. As in the scope detection task, error in negation cue detection led to lower the recall rate of the negated event detection system. The accuracy of full negation is based on the correct identification of the negation cues, scope and the negated events of all the negations present in the sentences. The output shows that there are many cases where negation cues and the scope are correctly identified but there is an error in identifying the negated events. The higher error in predicting the negated events led to reduce the score of full negation and achieve an F1 score of 39.04%.

Our system is unable to detect some negated events even though they are adjacent to the negation signal. This shows that the use of simple features extracted from words, lemmas, and POS tags is not enough to predict the negated events properly. Adding features related to words in left and right of the negation signal and the path feature may help to improve the detection of negated events.

In order to analyze the impact of error in the negation cue detection component upon the scope and negated event detection components, we performed an experiment using the gold standard negation cues to detect the scope and the negated events. F1 scores achieved by this system are 73.1% in full scope detection, 54.87% in negated event detection, 81.46% in scope tokens detection, and 49.57% in full negation detection. The result shows that there is almost 10% increment in the F1 score in all the components. Thus, having an improved cue detection component greatly helps to improve the accuracy of scope and negated event detection components.

## 6 Discussion and Conclusion

In this paper we outline a combination of a rule based approach and a machine learning approach to identify the negation cue, scope of negation, and the negated event. We show that applying a basic approach of using a lexicon to predict the negation cues

achieves a considerable accuracy. However, our system is unable to identify the negation cues such as *never*, *not*, *nothing*, *n't*, and *save* that can appear as a negation signal as well as in other non-negated contexts. It also cannot cover the negation cues of the signals that are not present in the training data. Moreover, in order to improve the overall accuracy of the scope and negated event detection, we need an accurate system to detect the negation cues since the error in the negation cue detection propagates to the next steps of identifying the scope and the negated event. It is difficult to identify the scope of negations that are farther in distance from the negation signal. Detecting the tokens of the scope that are discontinuous is also challenging.

As future work, we would like to extend our task to use a machine learning approach instead of the lexicon of negation signals to better predict the negation cues. The system we presented here uses a preliminary approach without including any syntactic information to detect the scope and negated events. We would also incorporate syntactic information to identify the scope and negated events in our future work. To improve the accuracy of identifying the scope and the negated events, adding other features related to the neighbor words of the negation signal might be helpful. In our tasks, we limit the scope and negated event instances by the window size in order to avoid imbalance data problem. Another interesting work to achieve better accuracy could be to use other approaches of imbalanced dataset classification instead of limiting the training instances by the window size.

## References

- Emilia Apostolova, Noriko Tomuro, and Dina Demner-Fushman. 2011. Automatic extraction of lexico-syntactic patterns for detection of negation and speculation scopes. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers - Volume 2*, HLT '11, pages 283–287, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Isaac G. Councill, Ryan McDonald, and Leonid Velekovich. 2010. What's great and what's not: learning to classify the scope of negation for improved sentiment analysis. In *Proceedings of the Workshop on Negation and Speculation in Natural Language Pro-*

- cessing, NeSp-NLP '10, pages 51–59, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Thorsten Joachims. 1999. Making large-scale support vector machine learning practical. In *Advances in kernel methods: support vector learning*, pages 169–184. MIT Press, Cambridge, MA, USA.
- Junhui Li, Guodong Zhou, Hongling Wang, and Qiaoming Zhu. 2010. Learning the scope of negation via shallow semantic parsing. In *Proceedings of the 23rd International Conference on Computational Linguistics, COLING '10*, pages 671–679, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Roser Morante and Eduardo Blanco. 2012. \*SEM 2012 Shared Task: Resolving the Scope and Focus of Negation. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics (\*SEM 2012)*, Montreal, Canada.
- Roser Morante and Walter Daelemans. 2009. A meta-learning approach to processing the scope of negation. In *Proceedings of the Thirteenth Conference on Computational Natural Language Learning, CoNLL '09*, pages 21–29, Stroudsburg, PA, USA.
- Roser Morante and Walter Daelemans. 2012. ConanDoyle-neg: Annotation of negation in Conan Doyle stories. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC)*, Istanbul.
- Roser Morante, Anthony Liekens, and Walter Daelemans. 2008. Learning the scope of negation in biomedical texts. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP '08*, pages 715–724. Association for Computational Linguistics.