Inferring the Scope of Negation in Biomedical Documents

Miguel Ballesteros†, Virginia Francisco‡,
Alberto Díaz†, Jesús Herrera†, Pablo Gervás‡

†Departamento de Ingeniería del Software e Inteligencia Artificial
‡Instituto de Tecnología del Conocimiento
Universidad Complutense de Madrid
C/ Profesor José García Santesmases, s/n
E–28040 Madrid, Spain
{miballes, virginia, albertodiaz, jesus.herrera}@fdi.ucm.es,
pgervas@sip.ucm.es

Abstract. In the last few years negation detection systems for biomedical texts have been developed successfully. In this paper we present a system that finds and annotates the scope of negation in English sentences. It infers which words are affected by negations by browsing dependency syntactic structures. Thus, firstly a greedy algorithm detects negation cues[1], like *no* or *not*. And secondly the scope of these negation cues is computed. We tested the system over the Bioscope corpus, annotated with negation, obtaining competitive results. The system presented in this paper can be accessed via web. [2]

1 Introduction

Generally speaking, negation turns an affirmative statement into negative (*I want / I do not want*) and indicates if it is a statement or not. It is also a complex phenomenon in natural languages and it has been an active research topic for decades. Researchers have approached this topic from both linguistic and philosophical perspectives [2]. In most cases, negation involves a negation cue and a negated syntagma containing one or more words that are within the scope of negation. In the following example: *"There is no detectable effect on leg segmentation"*, *'no'* is the negation cue used to denote that the following concept is negated, being *"detectable effect on leg segmentation"* the negated syntagma.

Nowadays negation detection is an emergent task in natural language processing. Detecting uncertain and negative assertions is essential in most text mining tasks where, in general, the aim is to derive factual knowledge from textual data. For instance, in text mining extracted information that is within the scope of negation should either be discarded or presented separately from factual information.

[1] A negation cue is defined as the lexical marker that expresses negation [1].
[2] http://minerva.fdi.ucm.es:8888/ScopeTagger

Negation is commonly seen in clinical documents and it is an important source of low precision in automated indexing systems [3]; as it is evidenced in Chapman's work when querying large medical free–text databases, the presence of negations can yield numerous false–positive matches, because the medical personnel is trained to include pertinent negatives in their reports. In a search for *fracture* in a certain radiology reports database, 95 to 99 percent of the returned reports would state *"no signs of fracture"* or words to that effect. Therefore, to increase the utility of indexing medical documents, it is necessary to acknowledge whether words have been negated or not.

In this paper we present a system that annotates the scope of negation, making use of a simple technique: first a manually-defined set of keywords is matched, then an algorithm marks the range of the scope within the sentence using unlabelled dependency syntactic structures. Finally, we evaluated the system with an established corpus annotated with the scope of negations, Bioscope [4], with three different collections: (i) the Papers collection, (ii) the Clinical reports collection and (iii) the Abstracts collection, containing 12.70%, 13.55% and 13.45% of sentences containing negations respectively.

In Section 2 we discuss previous works on processing negation. In Section 3 we describe the algorithms that we propose for inferring the scope of negation. In Section 4 we discuss the evaluation performed and, finally, in Section 5 we give our conclusions and suggestions for future work.

2 Previous Work

Nowadays there are two main kinds of systems that work with negation: systems that detect wordforms affected by negations, and (more recent) systems that classify the whole scope of negations, which is a more difficult task. Our system is classified in the second kind of systems. There is also a trend on works that try to extract negated events, such as [5].

For the biomedical domain there is plenty of research studying negation and finding how to detect it. For instance, Chapman et al. [6] detected negations and identified medical terms affected, by means of a simple regular expression algorithm called NegEx. It achieves 84.5% precision and 77.8% overall recall over 400 randomly selected sentences. In a similar way Mutalik et al. [7] recognized negated patterns in biomedical texts by using a training set of 40 medical documents; the set was manually inspected and used to develop a rule–based system (Negfinder), able to recognize a set of negated patterns in texts. They showed very good evaluation results, verified by human interaction, yielding 95.7% recall and 91.8% precision. Also, Huang and Lowe [8] implemented a hybrid approach to an automated negation detection system. They combined regular expression matching with grammatical parsing, to check the limits in automatically detecting negations in clinical reports. Their approach identified negated phrases with 98.6% precision and 92.6% recall in a a test set of 120 reports.

On the other hand, there are systems that infer the scope of negation. This is a more difficult problem, because it involves determining the words that are within the scope of a negation cue, where to open the scope, and finally, where to close it. One of the main works has been carried out by Morante's team [9, 10] in which a machine learning approach for the biomedical domain is shown. The system was evaluated with the Bioscope corpus and their results were: 80.11% overall precision and 78.44% recall in finding scopes of negation.

In 2010, a Workshop on Negation and Speculation in Natural Language Processing [11] was held in Uppsala, Sweden, bringing together researchers working on negation and speculation from any field related to computational language learning and processing. A specific goal was to describe the lexical aspects of negation to define how this phenomenon could be modelled for computational purposes, to explore techniques and to analyze how its treatment affects the efficiency of Natural Language Processing applications. Most of the approaches presented in the Workshop were in the biomedical domain, which is probably the most studied one in negation detection. An interesting paper for our work was presented in that workshop by Councill et al. [12]. They used the Bioscope corpus to evaluate their scope finding system, that is based on dependency parsing, and their results were 78.2% recall and 81.9% precision.

Later, Zhu et al. [13] presented a unified framework for scope learning by means of shallow semantic parsing, evaluating it with the Bioscope corpus. They divided the process in three main steps and they carried out the evaluation considering *golden cues* (which means that their system does not need to find where the cue is and which one it is), and *golden trees* (which means that their system does not need to find how the correct tree is, because it is given). They also reported their results when the system should predict correctly the tree and the cue. Their work was focused on inferring the scope of negation, but also on speculation. Their results, without using golden cues were 72.53% recall and 72.24% precision. When the *golden cue* was given, they were notably higher. This means that such kind of system in sinergy with an accurate cue classifier could be an interesting option to achieve very high results.

And finally, another interesting approach is presented by Agarwal and Yu [14]. They achieved an F1-score of 98% and 95% on detecting negation cue phrases and their scope in clinical notes, and an F1-score of 97% and 85% on detecting negation cue phrases and their scope in biological literature, they also included a website to test the system. [3]

3 Negation Scope Finding

Our system consists of two algorithms: the first one is capable of inferring words affected by the negative operators (cues) traversing dependency trees and the second one is capable of annotating sentences within the scope of negations. This second algorithm consists of a set of rules that have been built making use of a developing set, which was extracted from the Papers collection of the Bioscope corpus, more concretely, this developing set was formed by the first 10% of the

[3] http://snake.ims.uwm.edu/negscope/index.php

sentences that appear in the Bioscope Scientific Papers Collection (this set of sentences is removed in the Evaluation presented in Section 5).

In the first algorithm our system traverses a dependency tree, searching for negation cues to determine the correct scope over the tree. Our contribution lies in the identification of the scope, which is not explicit in the dependency tree. We selected Minipar parser [15] to develop the present experiment, which is a rule–based dependency parser capable of perform high accuracy unlabelled parsing (which is what we need for the present experiment). We selected Minipar because we only need unlabelled parsing and there is no need to train the system collecting annotated corpora.

Therefore, our system works as follows: a parse given by Minipar and a negation cue lexicon (as the one described in Section 3.1), are the inputs for the Affected Wordforms Detection Algorithm described in Section 3.2. Then, the Scope Finding Algorithm described in Section 3.3 acts on the set of negated nodes returning an annotated sentence with the scope of negation.

3.1 Negation Cue Lexicon

To determine the scope of negation, first of all a set of negation cues must be established. We considered the guidelines presented in [16] to set up our negation cue lexicon and also the work done by Mutalik et al. [7] in which some of the negation cues are presented.

The lexicon used to configure our system is shown in Table 1. This lexicon only contains the lemmas of each wordform, but our system is able to parse not only the lemma but all kind of verb forms.

Table 1. Lemmas of the Bioscope negation cues contained in our Negation Cue lexicon.

not	no	neither..nor	none
discard	rule out	fail	avoid
absence	lack (v)	lack (n)	without
unable	rather than	absent	cannot

These negation cues are the ones selected to develop the present work, but similar analyses can be accomplished with a different set.

3.2 Affected Wordforms Detection Algorithm

We implemented an algorithm that takes the dependency tree for a sentence returned by the dependency parser and returns for each negation cue a set of words affected by the cue. It uses the lexicon of negation cues presented in the previous section.

The algorithm traverses the dependency tree of a sentence, and it carries out the following steps:

1. It detects all the nodes that are contained in the lexicon of negation cues.
   - If the negation cue is a verb, it is marked as a negation cue.
   - If the negation cue is not a verb, the algorithm marks the verb (if exists) affected by it as a negation cue. In this way, the words that depends on the verb are affected by the negation cue.
2. For the rest of nodes, if a node depends directly on any of the ones previously marked as a negation cue, the system marked it as negated. Moreover, the negation is propagated from the cue word through the dependency graph until finding terminals, so wordforms that are not directly related with the cue are detected too.

The algorithm generates finally a set of nodes containing the wordforms within the scope of negation cues involved in the sentence. It is worth to emphasize that our system uses the same process for all the different negation cues, that makes the system easy to adapt to different domains.

3.3 Scope Finding Algorithm

This second algorithm is implemented using as an example a subset of the Scientific Papers collection of the Bioscope corpus (first 10% of the sentences containing negations) in order to give shape to the rules involved in this algorithm. We selected the Scientific papers collection to come up with the rules because it contains much more variety of sentences.

It works with the set of words returned by the Affected Wordforms Detection Algorithm, described in the previous Subsection, and the dependency tree given by the dependency parser in order to annotate sentences with the scope of negation, inferring where the scope must be opened and where it must be closed.

It is worth to emphasize that when the scope is opened to the right of the negation cue, the scope of negation leaves the subject out. This correspond to sentences in active voice and they are the most frequent case. Additionally, there are some cases in which the scope is opened to the left of the negation cues. The most frequent one is the passive voice. As shown in [17], passive voice is an exception in the way of tagging sentences in Bioscope. In this case the subject is marked within the scope of negation, because if the sentence had been written in active voice, it would be the object of a transitive verb. Therefore, the first step is to decide where to open the scope of negation, which is related to the voice of the sentence: if the sentence is in passive voice the scope must be opened to the left of the negation cue and if the sentence is in active voice, the scope must be generally opened to the right of the cue.

Therefore, we considered that the Scope Finding Algorithm must be divided in two main processes: first, to detect if the sentence follows a negated passive voice structure or not, and second, to annotate the sentence with the scope of negation, or scopes if there is more than one (which is an usual situation).

A sentence is in passive voice if:

- It contains a transitive verb, such as, *show, consider, see, use, detect*, etc.

– It follows one of the patterns shown below:
  1. *modal verb + <u>not</u> + be + past participle.*
  2. *am/is/are/was/were + <u>not</u> + past participle.*
  3. *have/has been + <u>not</u> + past participle.*

Once our system has decided if the sentence is in passive voice or not, the Scope Finding Algorithm iterates the sentence, token by token and applies a set of rules about the scope opening and closing. The rules are applied in the order presented below and it only applies one rule for each token.

1. Scope opening:
   a. If the token is contained in the set of nodes marked as negated by the Affected Wordforms Detection Algorithm and the scope for the cue involved is not open: the system opens the scope at the token and establishes that the scope for the cue involved is already opened.
   b. If the token is a negation cue and the sentence is in passive voice: the system goes backward and opens the scope just before the subject of the sentence. The system opens and closes the cue at the token.
   c. If the token is a negation cue and the sentence is not in passive voice: the system opens the scope just before the token. The system opens and closes the cue at the token.
2. Scope closing:
   a. If the token is a punctuation symbol, followed by some wordforms that indicates another statement, such as *but*: the system closes the scope just after the token.
   b. If the token is any wordform and all the nodes that are marked as negated for the negation cue are already included in the scope: the system closes the scope just before the token.
   c. If the token is the end of sentence: the system closes the scope at the end of the sentence.
3. Adding words to the sentence: if none of the previous rules has been applied the token is added to the annotated sentence.

At this point, the system has computed the scope (scopes) of the negation (negations) for a given sentence, by inferring which nodes pertain to that scope (scopes) from the node (nodes) marked as negated. Figure 1 illustrates the processing of the following sentence: *The reason why the two other families were not detected is more complex.* In the figure, the potential usefulness of the syntactic structure to infer the scope of negation is evidenced.

4 Evaluation

In this Section we present the evaluation performed to test how good our system is. In Section 4.1 we present the design of the evaluation as well as the evaluation metrics, in Section 4.2 we show the results considering these metrics, and finally, in Section 4.3 we compared our results with other state of the art approaches.

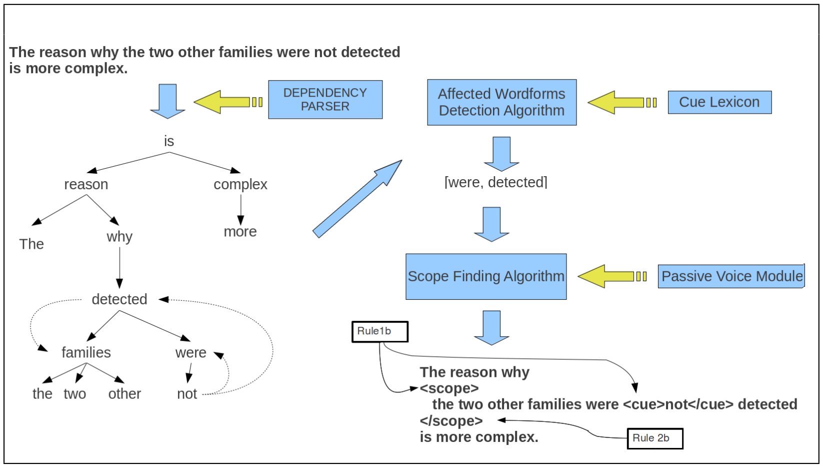

Fig. 1. The processing of a sentence by our system. As shown in the Figure, the rule applied to open the scope and open and closes the cue is *1b*. Finally, the rule applied to close the scope is *2b*, wherein the system closes the scope because there are no more wordforms marked as affected by the algorithm described in Section 3.2.

4.1 Evaluation Design

We selected Bioscope as the corpus for our evaluation. It is worth to emphasize that the evaluation is carried out over the output of the Scope Finding Algorithm, which is the output of the whole system and uses the Affected Wordforms Detection Algorithm.

Our first step in order to get the results was to select the sentences containing negations in the three collections of Bioscope, considering that for the Scientific Papers collection we must remove 10% of these sentences because this is the data set used to develop the rules of the algorithm presented in Section 3.3. Therefore, we evaluated our system with 100.0% of the clinical sentences containing negations, 90.0% of the papers sentences containing negations and 100% of the abstracts sentences containing negations.

The following evaluation measures were used:

$$P(Precision) = \frac{Tokens \quad correctly \quad negated \quad by \quad our \quad system}{Tokens \quad negated \quad by \quad the \quad system}$$

$$R(Recall) = \frac{Tokens \quad correctly \quad negated \quad by \quad our \quad system}{Tokens \quad negated \quad in \quad the \quad collection}$$

To balance the results in recall and precision, we used micro F1.

$$F1 = \frac{2PR}{P + R}$$

Additionally, we evaluate our system with the percentage of correct scopes (PCS), and the percentage of correct negation cues (PCNC).

$$PCS = \frac{Correct \quad Scopes \quad annotated \quad by \quad our \quad system}{Scopes \quad annotated \quad in \quad the \quad collection}$$

$$PCNC = \frac{Correct \quad negation \quad Cues \quad annotated \quad by \quad our \quad system}{Negation \quad Cues \quad annotated \quad in \quad the \quad collection}$$

By using all these measures we are considering not only a token–based evaluation but a whole scope classification measure, that really shows how good is a system that classifies the scope of negation in sentences.

4.2 Results and Discussion

When parsing the three collections of Bioscope, our system obtained the results given in Table 2.

It is worth to emphasize that we did scope identification with automatic cue recognition, so the input of our program, as shown in Section 3, is the sentence without any extra information.

Table 2. Results of our work, when evaluating it with the three collections of Bioscope.

Collection	**Precision**	**Recall**	**F1**	**PCS**	**PCNC**
Papers	73.49%	80.70%	76.93%	56.43%	91.15%
Abstracts	84.92%	84.03%	84.48%	68.92%	95.56%
Clinical	95.83%	90.58%	93.13%	89.06%	94.82%

The different results could be explained as follows. The main reason for the better results that our system achieves when annotating the sentences from the Clinical Reports collection is that the sentences contained in the clinical reports collection followed very easy syntactic structures and most of them contain the scope to the right. The average sentence length in the clinical reports collection is 7.73 wordforms, while in the papers and abstracts collections is more than 26 wordforms [10]. Moreover, a lot of sentences in the abstracts and papers collections contain more than one negation cue, which are more difficult to parse than those having only one. Finding the scope of negation in the simpler sentences of the clinical reports collection is easier than finding it in the sentences of the abstracts and papers collections.

In a similar way, the results for the abstracts collection are better than the ones for the scientific papers collection. One possible reason of this is the simplicity that usually characterizes abstract sentences. Sentences in an abstract are usually easy to understand, the writer commonly shows the main ideas of what is explained below with more simple syntactic structures.

Analyzing the sentences with mistakes, we found that there are two main reasons for these errors:

– Minipar, as any other dependency parser, is not error free, being able to cover about 79% of the dependency relations. If the dependency tree returned by Minipar is not correct, our system is not able to infer accurately the scope of negation in the sentence. In some cases we found that the tree is incomplete, and some information is missed. In these rare cases our system does not have enough information to decide how far the scope must be annotated. In most of these cases the cue is not correctly found and if it is, the scope probably closes incorrectly. As a suggestion for further work, we could replace Minipar with other dependency parsers that perform better the task of analyzing negation.
– There are some negation cues that are not always considered as negation cues, such as *negative*. This fact is evidenced in Morante's work [16]. Due to the characteristics of our system, we must define the negation cue lexicon at the beginning. A semantic module is the obvious suggestion to tackle these special cases, because we could be able to infer the negative semantics of a single word that is not always acting as a negation signal avoiding the noise produced by an initial static decision.

4.3 Comparison with other State–of–the–Art Systems

In this Section, we show an approximate comparison with some of the systems of the state of the art. We compare our results with the machine learning approach of Morante and Daelemans [10], the shallow semantic parsing approach of Zhu et al. [13] and the dependency system of Councill et al. [12]. The main comparison is shown in Table 3 where we show the precision, recall, F1, PCS and PCNC with other state of the art approaches. As evidenced in the results, our system performs very good for all the results with the exception of PCNC, in which the effect of our static and small lexicon of cues causes noise in the performance.

It is important to notice that in this table we show the results of Zhu's and Morante's systems when using automatic cue recognition, as we did in our system. Therefore, we are not reporting their results when using neither *golden cues* nor *golden trees*, which are much higher. In addition, for Councill's system only results for the scientific papers collection are shown because it is the only collection in which they published results.

Morante's system is based on machine–learning. In contrast, our system was constructed using as development set a subset of the sentences presented in the papers collection, as it is described in Section 3.3. Thus, while we tested our system with the Bioscope corpus (with the exception of the first 10% of the developing set of Papers), Morante et al. performed 10–fold cross validation experiments with the abstracts collection. And, for the other 2 collections, they trained with the abstracts set and tested with the corresponding collection. This fact affect the results, but we tried to make the results as comparable as possible.

This is why the Morante et al. results are much more comparable to ours in the case of the Abstracts collection. In the same way, Councill et al. results carrying out the experiment only with papers is much more directly comparable to our results because they get papers sentences to carry out the training. As

Table 3. Results of our work, evaluated with the three collections of Bioscope and compared with the systems of Morante et Al., Zhu et al. and Councill et Al.

Collection	System	Precision	Recall	F1	PCS	PCNC
Papers	Our Results	73.49%	80.70%	76.93%	56.43%	91.15%
	Morante et Al.	72.21%	69.72%	70.94%	41.00%	92.15%
	Zhu et Al.	56.27%	58.20%	57.22%	–	–
	Councill et Al.	80.80%	70.80%	75.50%	53.70%	–
Abstracts	Our Results	84.92%	84.03%	84.48%	68.92%	95.56%
	Morante et Al.	81.76%	83.45%	82.60%	66.07%	95.09%
	Zhu et Al.	78.24%	78.77%	78.50%	–	–
Clinical	Our Results	95.83%	90.58%	93.13%	89.06%	94.82%
	Morante et Al.	86.38%	82.14%	84.20%	70.75%	97.72%
	Zhu et Al.	82.22%	80.62%	81.41%	–	–

shown in Table 3, we can observe how these 2 cases produced similar results to ours. Councill et al. system seems to be very competitive because the original task of this work was to annotate sentences to solve a sentiment analysis task, nevertheless, their results with Bioscope were very good.

Table 4. PCS per negation cue for negation cues that occur 10 or more than 10 times in one of the subcorpus and appear in our lexicon of negation cues. The column # shows the number of appearances for each case; the column **Our** shows our system values and Morante's system values are given in column **Mor.**

	Abstracts			Papers			Clinical		
	#	**Mor.**	**Our**	#	**Mor.**	**Our**	#	**Mor.**	**Our**
absence	57	56.14	**71.93**	–	–	–	–	–	–
absent	13	15.38	**38.46**	–	–	–	–	–	–
cannot	28	**42.85**	28.57	16	50.00	50.00	–	–	–
fail	57	63.15	**85.97**	13	38.46	**53.84**	–	–	–
lack	85	**57.64**	52.94	20	45.00	**50.00**	–	–	–
neither	33	51.51	**72.72**	–	–	–	–	–	–
no	207	73.42	**81.64**	44	50.00	**54.54**	673	73.10	**89.60**
none	–	–	–	10	0.00	**71.42**	–	–	–
not	1036	**69.40**	66.41	200	39.50	**64.50**	57	50.87	**66.66**
rather than	20	65.00	65.00	12	**41.66**	25.00	–	–	–
unable	30	40.00	**73.33**	–	–	–	–	–	–
without	82	**89.02**	79.27	24	58.33	**70.83**	–	–	–

In Table 4 we show the percentage of correct scopes (PCS) per negation cue, for negation cues that occur 10 or more than 10 times in each collection present in Bioscope. We compare our results with the ones published by Morante and Daelemans [10], which is the same system studied in Table 3. Negation cues with a lower PCS have a higher percentage of scopes to the left (*absent, unable*). In this case we consider all the test set including the data seen (in the algorithm construction) in the Papers collection, that conforms the 100.0% of the papers sentences containing negations in order to obtain the same numbers as Morante et al. system, therefore, it is worth to mention that this column of data should be observed under this perspective.

In this depicted table of results we can see how the Morante's machine learning approach is not able to cover negation signals with a very low frequency in the training set but a higher frequency in the test, as we can see in the results for the negation cue *none* in the papers collection (it is worth to remind that they used the abstract collection for training and carried out the testing with the papers collection, in this case). Nonetheless, our system classifies the scope of this signal with higher accuracy.

Considering the most frequent negation cues, *not* and *no* (this cues are the ones with the strongest effect on the accuracy at the end), our system beats the results of Morante et Al. in clinical reports and papers collection. However, they beat our results for the cue *not* in the abstracts collection.

Finally, we did not include the Agarwal and Yu's work [14] in the comparison, which achieves an F1-score of 98% and 95% on detecting negation cue phrases and their scope in clinical notes, and an F1-score of 97% and 85% on detecting negation cue phrases and their scope in biological literature. This approach using conditional random fields present very high results, but as discussed in the Tutorial Given at the IJCNLP 2011 conference at Chiang Mai, Thailand,[4] the corpus partitions and the evaluation measures are different. Thus, the systems are, at least, not directly comparable, which shows that there are other ways to evaluate this task. Nevertheless, we consider important to mention this work, that includes a website in which is possible to test the system and shows a very interesting approach, as we described in Section 2.

5 Conclusions and Future Work

In this paper we presented a high performance system able to infer the scope of negations. From the results of our experiments we can conclude that dependency parsing is a valuable auxiliary technique for negation detection, at least in the particular case of English.

As a suggestion for future work, we consider that the scope of negation must not always be annotated as continuous. In Bioscope, the scope of negation leaves the subject out, with the exception of passive voice sentences. Nonetheless we consider that the subject must always be considered as a part of the scope. Moreover, when there is an affirmative sentence that affects the subject of a negative passive voice sentence, it is difficult to infer automatically which subject is considered. For instance, in the following sentence *"Therefore, TNF-alpha mRNA induction by PMA, like its induction by virus and LPS, [is* ***not*** *primarily mediated by NF-kappa B], but rather is mediated through other sequences and protein factors."*, the scope of negation is in passive voice but the subject is

[4] http://www.ijcnlp2011.org/ijcnlp2011/downloads/tutorial/tu3_present.pdf

implicit by the word *is*, and it is not directly included in the scope of negation if we follow the annotation guidelines of Bioscope, as it is done in the present work. Thus, we suggest that the scope must be discontinuous in the way of considering other wordforms that in Bioscope are out of the scope, but are directly affected by the negation cue. It can be achieved using a tabular format for the corpus, instead of plain sentences annotated with XML language. As we show in the present work, we decided to evaluate our system with Bioscope, thus our system annotates the sentences in the same way as it is done in that corpus. For further work we are going to consider other approaches to evaluate our system, some of them are introduced in [1].

Observing the results shown in Table 4, where we show the PCS per negation cue, an interesting idea for future work could be a system that uses in sinergy the results of both systems, our system and Morante et Al. Combining what is presented in this paper with the fact that a common approach to fault–tolerant systems is the implementation of a voting system that select the best output for two systems that perform the same computational task, in this case depending on the cue involved.

Finally, for future work, we are going to study different dependency syntactic parsers in order to find which one is the best for inferring the scope of negation. We are also thinking to test our system into the opinion mining domain, as it is done in [18], in which the effect of negation is studied for sentiment analysis in product reviewing.

Acknowledgments

This research is funded by the Spanish Ministry of Education and Science (TIN2009-14659-C03-01 Project), Universidad Complutense de Madrid and Banco Santander Central Hispano (GR58/08 Research Group Grant).

References

1. Morante, R., Schrauwen, S., Daelemans, W.: Corpus-based approaches to processing the scope of negation cues: an evaluation of the state of the art. In Bos, J., Pulman, S., eds.: Proceedings of the Ninth International Conference on Computational Semantics (IWCS 2011). (2011) 350–354
2. Horn, L.R.: A natural history of negation / Laurence R. Horn. University of Chicago Press, Chicago : (1989)
3. Chapman, W.W., Bridewell, W., Hanbury, P., Cooper, G.F., Buchanan, B.G., Phd, W.W.C., Bs, W.B., Bs, P.H., Phd, G.F.C.M., Phd, B.G.B.: Evaluation of negation phrases in narrative clinical reports (2002)
4. Vincze, V., Szarvas, G., Farkas, R., Mora, G., Csirik, J.: The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes. BMC Bioinformatics **9** (2008) S9+
5. Sarafraz, F., Nenadic, G.: Identification of negated regulation events in the literature: exploring the feature space. In: Semantic Mining in Biomedicine. (2010)

6. Chapman, W.W., Bridewell, W., Hanbury, P., Cooper, G.F., Buchanan, B.G.: A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform **2001** (2001) 34–301
7. Mutalik, P.G., Deshpande, A., Nadkarni, P.M.: Use of General-purpose Negation Detection to Augment Concept Indexing of Medical Documents: A Quantitative Study using the UMLS. Journal of the American Medical Informatics Association **8** (2001) 598–609
8. Huang, Y., Lowe, H.J.: A Novel Hybrid Approach to Automated Negation Detection in Clinical Radiology Reports. Journal of the American Medical Informatics Association **14** (2007) 304–311
9. Morante, R., Liekens, A., Daelemans, W.: Learning the scope of negation in biomedical texts. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. EMNLP '08, Stroudsburg, PA, USA, Association for Computational Linguistics (2008) 715–724
10. Morante, R., Daelemans, W.: A metalearning approach to processing the scope of negation. In: Proceedings of the Thirteenth Conference on Computational Natural Language Learning. CoNLL '09, Stroudsburg, PA, USA, Association for Computational Linguistics (2009) 21–29
11. Morante, R., Sporleder, C., eds.: Proceedings of the Workshop on Negation and Speculation in Natural Language Processing, Uppsala, Sweden, University of Antwerp (2010)
12. Councill, I.G., McDonald, R., Velikovich, L.: 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 Processing. NeSp-NLP '10, Stroudsburg, PA, USA, Association for Computational Linguistics (2010) 51–59
13. Zhu, Q., Li, J., Wang, H., Zhou, G.: A unified framework for scope learning via simplified shallow semantic parsing. In: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. EMNLP '10, Stroudsburg, PA, USA, Association for Computational Linguistics (2010) 714–724
14. Agarwal, S., Yu, H.: Biomedical negation scope detection with conditional random fields.j am med inform assoc. Journal of the American Medical Informatics Associations (2010)
15. Lin, D.: Dependency-based evaluation of MINIPAR. In: Proc. Workshop on the Evaluation of Parsing Systems. Granada (1998)
16. Morante, R.: Descriptive analysis of negation cues in biomedical texts. In Chair), N.C.C., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., Tapias, D., eds.: Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), Valletta, Malta, European Language Resources Association (ELRA) (2010)
17. Szarvas, G., Vincze, V., Farkas, R., Csirik, J.: The bioscope corpus: annotation for negation, uncertainty and their scope in biomedical texts. In: Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing. BioNLP '08, Stroudsburg, PA, USA, Association for Computational Linguistics (2008) 38–45
18. Carrillo de Albornoz, J., Plaza, L., Gervás, P., Díaz, A.: A joint model of feature mining and sentiment analysis for product review rating. In: Proceedings of the 33rd European Conference on Information Retrieval (ECIR 2011). (2011)