

Some Issues on Detecting Negation from Text

Eduardo Blanco and **Dan Moldovan**

Human Language Technology Research Institute

The University of Texas at Dallas

Richardson, TX 75080 USA

{eduardo,moldovan}@hlt.utdallas.edu

Abstract

Negation is present in all human languages and it is used to reverse the polarity of parts of a statement. It is a complex phenomenon that interacts with many other aspects of language. Besides the direct meaning, negated statements often carry a latent positive meaning. Negation can be interpreted in terms of its scope and focus. This paper explores the importance of both scope and focus to capture the meaning of negated statements. Some issues on detecting negation from text are outlined, the forms in which negation occurs are depicted and heuristics to detect its scope and focus are proposed.

Introduction

Capturing the semantics of text is a long term goal in the natural language processing community. Whereas philosophers and linguists have proposed several theories, along with models to represent the meaning of text, the field of computational linguistics is still far from doing so automatically. The ambiguity of natural language, the need to detect implicit knowledge, and the demand for commonsense knowledge and reasoning are only a few of the difficulties to overcome to understand text. We must say that significant progress has been made, especially on detection of semantic relations, ontologies and reasoning methods.

Negation is present in all languages and it is always the case that statements are affirmative by default. Negation is marked and it typically signals something unusual or an exception. It may be present in all units of language, e.g., words (*incredible*), clauses (*He doesn't have friends*). Negation and its correlates (truth values, lying, irony, false or contradictory statements) are defining and exclusive characteristics of the human species (Horn 1989; Horn and Kato 2000). Simply put, negation is a process that turns parts of a statement into its opposite.

Negation is fairly well-understood and described in grammars; the valid ways to express a negation are formulated and documented. However, there has not been much work on automatically detecting it, or more importantly, on representing the semantics of negations from natural language.

At first glance, negation might seem easy to deal with. One might think that the problem could be reduced to find

negative polarity items, determine their scope and reverse its polarity. Actually, it is much more problematic.

Negation plays a remarkable role towards understanding text and poses considerable challenges. Negation interacts with many other phenomena and it is used for so many different purposes that a deep analysis is needed. The following are some issues found when dealing with negation.

Detecting the scope of negation in itself is challenging: *All vegetarians do not eat meat* means that vegetarians do not eat meat and yet *All that glitters is not gold* means that it is not the case that all that glitters is gold (so out of all things that glitter, some are gold and some are not). In the former example, the universal quantifier *all* has scope over the negation; in the latter, the negation has scope over *all*.

In logic, two negatives always cancel each other out. On the other hand, in language that is only theoretically the case: *she is not unhappy* does not mean that *she is happy*; it means that *she is not fully unhappy, but she is not happy either*.

Some negated statements carry a positive implicit meaning. For example, *cows do not eat meat* implies that *cows eat something other than meat*. Otherwise, the speaker would have stated *cows do not eat*. A clearer example is the correct and yet puzzling statement *tables do not eat meat*. The sentence sounds unnatural because the underlying positive statement (i.e., *tables eat something other than meat*) contradicts the fact that *tables do not eat*.

Negation can express *less than* or *in between* when used in a scalar context. For example, *John does not have three children* probably means that he has either one or two children. Contrasts may use negation to disagree about a statement and not to negate it. For example, *That truck is not big, it is massive* defines the truck as *massive*, and therefore, *big*.

In this paper, we investigate the significance of negation in semantic representation of natural language. We illustrate the importance of detecting both its scope and focus and propose a semantic representation that benefits from focus detection.

Related Work

Negation has been widely studied outside of computational linguistics. In logic, it is usually the simplest unary operator and it reverses the truth value. The seminal work on negation by Horn (1989) presents the main thoughts in philosophy and psychology. We follow him in the next two paragraphs.

Two of the most basic philosophical laws put forth by Aristotle are the Law of Contradiction (LC, it is impossible to be and not be at the same time) and the Law of Excluded Middle (LEM, in every case we must either affirm or deny). LEM is not always applicable to statements involving negation of scalar values (e.g., one can deny *being cold* and *not being cold*). Philosophers also realized that a negative statement can have latent positive meaning, e.g. *Socrates is not well* implicitly states that *Socrates is alive*.

Psychology has studied the constructs, usage and cognitive processing of negation. They note that negated statements are not on equal footing with a positive statement; they are a different and subordinate kind of statement. Actually, there is evidence that children acquire negation later in life than the power to communicate. Psychology also confirms the intuitive thought that humans normally communicate in positive terms and reserve negation mostly to describe unusual or unexpected situations.

Linguists have found negation a highly complex phenomenon. The Cambridge Grammar of the English Language (Huddleston and Pullum 2002) dedicates over 60 pages to negation, covering verbal (e.g., *he doesn't agree*) and non-verbal (e.g., *Not all of them agree*), polarity items (e.g., *already*, *any*) and multiple negation. Among others, negation interacts with quantifiers and anaphora (Hintikka 2002). For example, the reference to *They* is easier to solve in (1) than in (2): (1) *Some of the students passed the exam. They must have studied hard*; (2) *Not all the students failed the examination. They must have studied hard*. Negation also influences reasoning (Dowty 1994; Sánchez Valencia 1991). Zeijlstra (2007) analyzes the way different languages position and form negative elements as well as the interpretation of more than one negative element.

Within natural language processing, negation has drawn attention mainly in sentiment analysis (Wilson, Wiebe, and Hoffmann 2009) and the biomedical domain. The Negation and Speculation in Natural Language Processing Workshop (Morante and Sporleder 2010) targeted negation and speculation; most contributions specialize in the above subfields. The CoNLL-2010 Shared Task (Farkas et al. 2010) targeted the detection of hedges and their scope. Councilill, McDonald, and Velikovich (2010) created their own corpus to detect explicit negations and their scope in a supervised manner. Using the BioScope corpus, Morante and Daelemans (2009) propose a supervised scope detector and Özgür and Radev (2009) offer syntactic rules to detect scopes. Wiegand et al. (2010) survey the role of negation in sentiment analysis.

Some applications in natural language processing deal indirectly with negation. Among many others, van Munster (1988) considers negation for machine translation, Rose et al. (2003) for text classification and Bos and Markert (2005) for recognizing entailments.

As far as we are concerned, there is not a corpus with scope and focus annotation. The BioScope corpus (Vincze et al. 2008) annotates negation and linguistic scopes for the biomedical domain. PropBank (Palmer, Gildea, and Kingsbury 2005) annotates verbal negation, but it considers neither scope nor focus, it just indicates the verb to which the negation mark attaches. FactBank (Saurí and Pustejovsky

2009) annotates event mentions exclusively with their degree of factuality.

None of the above references aim at detecting or annotating the focus of negation in natural language. To the best of our knowledge, researches have not yet targeted this topic.

Negation in Natural Language

In this Section we follow Huddleston and Pullum (2002) to describe the characteristics of negation in natural language with a focus on English language.

Unlike affirmative statements, negation is marked by words (e.g., *not*, *no*, *never*) or affixes (e.g., *-n't*, *un-*). Negation can interact with other words in special ways. For example, negated clauses use different connective adjuncts than positive clauses do: *neither*, *nor* instead of *either*, *or*. The so-called *negatively-oriented polarity-sensitive items* (Huddleston and Pullum 2002) include, among many others, words starting with *any-* (*anybody*, *anyone*, *anywhere*, etc.), the modal auxiliaries *dare* and *need* and the grammatical units *at all*, *much* and *till*. Negation in verbs usually requires an auxiliary; if none is present, the auxiliary *do* is inserted (*I read the paper* vs. *I didn't read the paper*).

We can distinguish four contrasts for negation:

Verbal vs. Non-verbal. Verbal if the marker of negation is grammatically associated with the verb (e.g., *I did not see anything at all*); non-verbal if it is associated with a dependent of the verb (e.g., *I saw nothing at all*).

Analytic vs. Synthetic. Analytic if the negation is marked by words whose sole syntactic function is to mark negation (e.g. *Bill did not go*); synthetic if the words have some other function as well (e.g., *Nobody went to the meeting*). In the latter example, *Nobody* marks the negation and plays the role of AGENT.

Clausal vs. Subclausal. Clausal if the negation yields a negative clause (*She didn't have a large income*); subclausal otherwise (*She had a not inconsiderable income*)

Ordinary vs. Metalinguistic. A negation is ordinary if it indicates that something is not the case, e.g., (1) *She didn't have lunch with my old man: he couldn't make it*. On the other hand, a negation is metalinguistic if it does not dispute the truth but rather reformulates a statement, e.g., (2) *She didn't have lunch with your 'old man': she had lunch with your father*. Note that in (1) the lunch never took place, whereas in (2) a lunch did take place.

Scope and Focus

Negation has both scope and focus and they are extremely important to capture its semantics. The scope of negation is the part of the meaning that is negated. The focus is that part of the scope that is most prominently or explicitly negated (Huddleston and Pullum 2002).

The two concepts are interconnected. Scope refers to all elements whose individual falsity would make the negated statement strictly true. Focus is the element of the scope that is *intended* to be interpreted as false to make the overall negative true.

For example, consider (1) *Your children don't hate school* and its positive counterpart (2) *Your children hate school*.

The truth conditions of (2) are: (a) somebody hates something; (b) your children are the ones that hate; and (c) the hatred is felt towards school.

Each of the above statements has to be true in order for (2) to be true. And the falsity of any of them is sufficient to make (1) true. In other words, (1) would be true if nobody hates, your children are not the ones that hate or the school is not the thing being hated. Therefore, all three statements are inside the scope of *Your children don't hate school*.

The focus is usually more difficult to identify than the scope, especially without knowing the stress or intonation. With only taking into account single words, there are four possible focuses (wavy underline) for *Your children don't hate school*. Each of them encodes different meanings:

- *Your children don't hate school*; some children hate school, but not yours.
- *Your children don't hate school*; some of your relatives hate school, but not your children.
- *Your children don't hate school*; your children do not hate school, but they harbor a negative attitude towards school.
- *Your children don't hate school*; your children hate something, but they do not hate school.

As the above examples show, identifying the focus can be highly ambiguous. Other examples are more clear. For example, *Cows don't eat meat* has as plausible focuses the underlined words, with the interpretations *there are things that eat meat, but cows do not and cows eat, but they do not eat meat* respectively. Another example is *Liz didn't intentionally delete the file*, which means that *Liz deleted a file by mistake*.

Even though scope and focus are primarily semantic, syntactic information helps in detecting them.

Interpretation

As illustrated in the previous section, interpreting a negation is not straightforward and detecting its focus plays an important role. Other aspects must also be considered.

First, there is the direct literal meaning. In addition, negated statements often contain a positive statement or subtext beneath the direct meaning; a meaning that is suggested or implied by the statement. Consider the following example: *Mark could not finish his meal*.

This statement carries the direct literal meaning *Mark did not finish his meal*. Additionally, *Mark tried to finish his meal (but failed)* is implicitly stated although from a strictly logical point of view it is not entailed. Had the speaker not intended the implicit meaning, he would have stated *Mark did not finish his meal*.

Modals like *can* and *could* are usually combined with *not* when the speaker intends to communicate a failed attempt, not just lack of ability. On the other hand, a negation involving *would* usually conveys a refusal: *Mark wouldn't finish his meal* carries the implicature *He refused to eat his meal*.

Metalinguistic negation is not frequently seen in text, but it occurs fairly often in speech. This kind of negation is used to reformulate or correct a statement, but not to negate it. For example, *Max hasn't got four children: he's got five* does not really negate anything and should be interpreted as *Max has*

got five children (therefore, he also has four children). As Huddleston and Pullum (2002) state, if we only state *Max hasn't got four children*, the interpretation *Max has less than four children* would be correct.

Double negation is not common in text either. In language, two negatives cancel each other or emphasize the negativeness of a statement. For example, *Never before had no one been nominated for the position* carries the affirmative assertion *Always before someone had been nominated for the position*. On the other hand, *I haven't eaten nothing at all* carries a strong negative meaning. A third possibility is to use double negation to make a weaker claim about something: *it is not uncommon* means that *it is somewhere in between common and uncommon*.

Semantic Representation of Negation

Negation does not stand on its own. To be useful, it should be added as part of another existing knowledge representation. In this section, we outline how to incorporate negation into semantic relations. Rather than working with a specific set, we show how negation can be included into any set using well-known relations.

Semantic relations capture connections between concepts and labels them according to their nature. It is out of the scope of this paper to define them in depth or to establish a possible set of relations to consider. Given the sentence *The cow ate grass with her tongue*, a possible semantic representation using semantic relations is AGENT(*the cow, ate*), THEME(*grass, ate*), INSTRUMENT(*with her tongue, ate*) and PART-WHOLE(*her tongue, the cow*).

Negation has been mostly ignored within semantic relations. If we want to represent the meaning of *The cow didn't eat grass with a fork*, several options arise. First, we find it useful to consider the semantic representation of the affirmative counterpart: AGENT(*the cow, ate*), THEME(*grass, ate*), and INSTRUMENT(*with a fork, ate*).

Second, we believe detecting scope and focus helps to incorporate negation. The scope of the above statement corresponds to the three semantic relations:

- AGENT(*the cow, ate*); the cow ate
- THEME(*grass, ate*); grass was eaten
- INSTRUMENT(*with a fork, ate*); a fork was used to eat

The falsity of any of the above items would make the negated statement true. The focus corresponds to INSTRUMENT(*with a fork, ate*). Even though it is open to discussion, it seems that this is the most likely focus. Thus, the negated statement should be interpreted as *the cow ate grass, but it did not do so using a fork*.

Table 1 depicts 5 different possible semantic representation for the above example. Option (1) does not incorporate any explicit representation of negation. It attaches the negated mark and the required auxiliary to the event *eat*; the negation is part of the relation arguments. This option fails to detect any underlying positive meaning and corresponds to the interpretation *the cow did not eat, grass was not eaten and a fork was not used to eat*.

Options 2–5 embody negation into the representation with the pseudo-relation NOT. NOT takes as its argument an in-

1	AGENT(<i>the cow, didn't eat</i>)	THEME(<i>grass, didn't eat</i>)	INSTRUMENT(<i>with a fork, didn't eat</i>)
2	NOT[AGENT(<i>the cow, ate</i>)]	THEME(<i>grass, ate</i>)	INSTRUMENT(<i>with a fork, ate</i>)
3	NOT[AGENT(<i>the cow, ate</i>)]	THEME(<i>grass, ate</i>)	INSTRUMENT(<i>with a fork, ate</i>)
4	AGENT(<i>the cow, ate</i>)	NOT[THEME(<i>grass, ate</i>)]	INSTRUMENT(<i>with a fork, ate</i>)
5	AGENT(<i>the cow, ate</i>)	THEME(<i>grass, ate</i>)	NOT(INSTRUMENT(<i>with a fork, ate</i>))

Table 1: Possible semantic representations for *The cow didn't eat grass with a fork.*

		#Occ.	%Occ.			#Occ.	%Occ.	
-n't	RB	4,005	55.87	never	RB	226	3.15	
	All	4,005	55.87		NNP	2	0.03	
not	RB	1,702	23.74		RBR	1	0.01	
	All	1,702	23.74		All	229	3.19	
no	DT	870	12.14		neither	DT	45	0.63
	RB	113	1.58			CC	29	0.40
	UH	14	0.20	RB		6	0.08	
	NNP	3	0.04	IN		1	0.01	
	All	1,000	13.95	All		81	1.13	
nor	CC	88	1.23	none	NN	63	0.88	
	All	88	1.23		NNP	1	0.01	
					All	64	0.89	

Table 2: Negation-bearing words and affixes, POS tags and their occurrences. Total number of occurrences is 7,169.

stantiated relation or set of relations and indicates that the argument does not hold. NOT[R(x, y)] should be read *it is not the case that R holds between x and y.*

Option 2 includes all the scope as the argument of NOT and corresponds to the interpretation *it is not the case that* (a) *the cow ate*, (b) *grass was eaten*, and (c) *a fork was used to eat*. Options (2) and (1) are very similar, the only difference is the fact that (2) removes the negative marks from the relation arguments and instead explicitly represents the negation using NOT.

The remaining options encode different meanings:

- Option (3) negates the AGENT; it corresponds to *the cow didn't eat, but grass was eaten with a fork.*
- Option (4) applies NOT to the THEME; it corresponds to *the cow ate something with a fork, but it was not grass.*
- Option (5) denies the INSTRUMENT, encoding the meaning *the cow ate grass, but it did not use a fork to do so.*

We prefer option (5) since it correctly captures the semantics of *The cow didn't eat grass with a fork* by revealing the hidden positive meaning. This representation corresponds to the semantic representation of the affirmative counterpart (*The cow ate grass with a fork*) after applying NOT over the focus of the negation. This fact justifies and motivates the importance of detecting both the scope and especially the focus of a negation.

Automatic Discovery of Scope and Focus

We believe that in order to detect negation in natural language we must first analyze how often and in which form negation occurs. We do so by using the WSJ section of the Penn TreeBank, which contains full syntactic information. Even though grammars offer the correct ways of forming a negation, they neither take into account frequency nor the

VP-be-not-PRD	VP-be-not-VP	VP-aux-not-VP	VP-MD-not-VP
(VP (* <be>))	(VP (* <be>))	(VP (* <aux>))	(VP (MD *))
(...)	(...)	(...)	(...)
(RB n[o']t)	(RB n[o']t)	(RB n[o']t)	(RB n[o']t)
(...)	(...)	(...)	(...)
(*-PRD *)	(VP *)	(VP *)	(VP *)

Table 3: Syntactic structure for each pattern.

actual usage of negation constructs. That said, we must recognize that work on theoretical linguistics has been invaluable in understanding the intrications of negation.

First, we counted the occurrences of negation-bearing words in the corpus. Table 2 shows the counts for each word considering part-of-speech tags. The total number of occurrences is 7,169, and 5,707 correspond to *not* and its affixed form *n't* in a RB tag. The rest of words occurs significantly less, between 1,000 and 64 times.

Because *not* and *n't* correspond to 79.61% of the negative-bearing words occurrences, we focus exclusively on them. From now on, we use *not* to refer to both *not* and *n't*.

Syntactic Patterns of Negation

The syntactic patterns which contain *not* are depicted in Tables 3 and 4. Table 3 contains the syntactic structure for the most common patterns and Table 4 their number of occurrences and examples from the Penn TreeBank. The 1,025 occurrences of pattern VP-be-not-PRD occur with the following distribution for the PRD node: ADJP (508 times), NP (380), PP (83), ADVP (23), SBAR (21) and S (10). In Table 3, pattern VP-aux-not-VP, *aux* can be either *do* or *have*. All five patterns, which account for 94% of occurrences of *not*, start with a VP.

Pattern VP-aux-not-VP (patterns 3 and 4 in Table 4) accounts for 40% of occurrences. Patterns involving the verb *be* (patterns 1 and 2) account for 30% of occurrences, and pattern 5, which includes a modal, for 24%. In the remainder of this paper, we present heuristics and results for detecting the scope and focus of negations that follow patterns 1–4. These patterns account for 69.90% of the total number of occurrences of *not* in the Penn TreeBank.

Determining Scope and Focus

In order to determine scope and focus of a negation, we propose two simple heuristics based exclusively in syntactic information. This simple method should be interpreted as a first approach for the task. The heuristics were defined after examining examples and taking into account theoretical works on negation.

	Pattern	Occurr.		Example	
		#	%	File, #sent	Sentence
1	VP-be-not-PRD	1,025	17.96	wsj_0039, 0	As an actor, Charles Lane [isn't [the inheritor of Charlie Chaplin's spirit] _{NP}] _{VP} .
2	VP-be-not-VP	673	11.79	wsj_0062, 13	He says Campbell [wasn't even [contacted by the magazine for the opportunity to comment] _{VP}] _{VP} .
3	VP-do-not-VP	1,767	30.96	wsj_0039, 38	But she [didn't [deserve to have her head chopped off] _{VP}] _{VP} .
4	VP-have-not-VP	524	9.18	wsj_0008, 0	The federal government suspended sales of U.S. savings bonds because Congress [hasn't [lifted the ceiling on government debt] _{VP}] _{VP} .
5	VP-MD-not-VP	1,370	24.00	wsj_0155, 0	World sugar futures prices soared on rumors that [...] Brazil, a major grower and exporter, [might not [ship sugar this crop year and next] _{VP}] _{VP} .
6	Other	348	6.10	wsj_0203, 12	[Not [all those who wrote] _{NP}] _{NP} [oppose the changes] _{VP} .

Table 4: Most productive syntactic patterns containing *not* or *-n't*, number of occurrences and examples from the Penn Tree-Bank. Details for each pattern can be found in Table 3.

Pattern	Accuracy Detecting	
	Scope	Focus
VP-be-not-PRD	0.65	0.15
VP-be-not-VP	0.71	0.64
VP-aux-not-VP	0.67	0.46
All	0.66	0.41

Table 5: Accuracies for scope and focus detection.

For scope detection, the primary constituent following the negation is selected as the beginning of the scope. This corresponds to the $*-PRD$ node in pattern 1 and the VP in the other patterns (Table 3). However, we do not select everything within that constituent to be in the scope. We exclude subclauses (S or $SBAR$) and prepositional phrases (PP).

Detecting the focus is a challenging task and defining syntactic heuristics has proved harder. We experimented with selecting the first major child of the scope constituent. This heuristic is a fair approximation in some cases, but in far too many instances it is wrong.

We have evaluated the accuracy of the method using the first 100 sentences containing patterns $VP-be-not-PRD$, $VP-be-not-VP$ or $VP-aux-not-VP$, which correspond to patterns 1–4 from Table 4 (*aux* can be either *do* or *have*). Table 5 show the accuracies for scope and focus detection.

Detecting the scope is done fairly well for the four patterns; the accuracies range from 0.65 to 0.71. On the other hand, the accuracies of detecting the focus, ranging from 0.15 to 0.64, are not satisfactory. We must note that focus detection partially relies on correct scope detection.

Error Analysis The scope of the negation was detected with an overall accuracy of 0.66. The errors are due to the following causes:

- excluding PP is responsible for 59% of the errors;
- excluding S and $SBAR$ for 19%;
- not handling the inversion of scope with another scope in the sentence (e.g., context) is responsible for 14%; and
- the remainder 8% is due to miscellaneous causes.

This error analysis suggests that excluding subclauses (S and $SBAR$) for the scope is an adequate heuristic, but also reveals that excluding PP is less sound. For example, given *The cow did not eat grass [with a fork]_{PP}*, the proposed

heuristic excludes the PP from the scope. This mistake implies that the focus is not correctly identified either.

Detecting focus is more problematic. Our error analysis suggests that selecting the *last* major child of the scope might be a better heuristic. Several examples support this idea, e.g., *He didn't go to Cancun to relax* and *The cow did not eat grass with a fork*. As we have already argued, detecting the focus of negation is an ambiguous task and sometimes unclear even for humans.

Enhancements

The above method can be extended in several ways. First, we only take into account negation-bearing words, ignoring negation-bearing prefixes (*in-*, *im-*, *un-*, *dis-*, etc.).

No simple search could unequivocally distinguish between a negated word such as *ineffective* and the words that just happen to begin with the letters of a negative prefix, such as *invite*. The problem could be partially solved by checking if, after removal of the prefix, the word is still valid. This method mismarks *inform* as negation because *form* is a valid word. To complicate matters further, some words are valid both as negated base words and as words in their own right: the adjective *invalid* means *not valid*, while the noun *invalid* describes a disabled person. A dictionary of antonyms would be most useful for this task.

For detecting scope, determining if the negation is clausal (i.e., it applies to the whole sentence or clause) or subclausal (i.e., it applies only to a constituent or a word) could be beneficial. Taking into account the presence of negatively-oriented polarity-sensitive items (NPI) could aid in establishing the boundaries of the negation scope. Current heuristics discard S , $SBAR$ and PP from the scope, but it seems reasonable not to do so if they include a NPI. Focus detection is more complicated and highly ambiguous. Advanced heuristics are needed as well as text understanding.

The proposed heuristics use exclusively syntactic information, even though examples show that the tasks are mostly semantic. Using these heuristics as features for a machine learning algorithm is the next natural step to take. Syntactic and semantic features used for semantic role labeling might be suitable. The presence of semantic relations is likely to help detecting the focus, as well as semantic classes and the position of a potential focus with respect to the VP .

Conclusions

Capturing the meaning of negated statements is a complicated task, primarily due to the interaction of negation with other linguistic phenomena. Negation is exclusive to humans and can be used for different purposes: reverse the polarity of a statement (e.g., *She didn't like it*), emphasize the negativeness (e.g., *She hasn't eaten nothing*) or make weaker claims (e.g. *She is not dumb*).

The valid ways in which negation can be expressed have been widely studied. Negated statements often carry positive meaning beneath the direct meaning and detecting them yields precious knowledge. For example, given *John didn't go to Cancun to relax*, a reader will interpret that *John went to Cancun, but his purpose was not to relax*.

Negation can be analyzed in terms of scope and focus. Broadly speaking, the scope consists of the part of the meaning that is being negated and the focus is that part of the scope that is most prominently or explicitly negated (Huddleston and Pullum 2002). Both scope and focus are primarily semantic, especially the latter.

We have proposed a way to add negation to semantic representation of text using the pseudo-relation NOT[x]. NOT takes as its argument x one or more semantic relations and should be read *it is not the case that x*. The semantics of a negated statement can be captured by first obtaining the semantic representation of the positive counterpart and then applying NOT to the focus of the negation.

Given *John did not go to Cancun to relax*, following the above steps we obtain AGENT(*John*, *went*), AT-LOCATION(*went*, *to Cancun*) and NOT[PURPOSE(*to relax*, *went*)]. The accurate determination of scope and focus is the key to detecting the latent positive meaning and obtaining a careful representation. The semantic representation of the above example explicitly states which part of the meaning is negated (*the purpose of going was not to relax*) and which one is affirmative (*John went to Cancun*).

The forms of negation have been analyzed using the Penn TreeBank. The most frequently occurring negation-bearing words have been analyzed. The syntactic patterns these words occur in have been identified based on productivity. Heuristics have been proposed for detecting scope and focus of negations following the most occurring patterns. Results show acceptable performance for scope detection and opportunity to improve on focus detection for future research.

References

Bos, J., and Markert, K. 2005. Recognising Textual Entailment with Logical Inference. In *Proceedings of HLT-EMNLP conference*.

Councill, I.; McDonald, R.; and Velikovich, L. 2010. What's Great and What's Not: Learning to Classify the Scope of Negation for Improved Sentiment Analysis. In *Proceedings of the NeSp-NLP Workshop*.

Dowty, D. 1994. The Role of Negative Polarity and Concord Marking in Natural Language Reasoning. In *Proceedings of Semantics and Linguistics Theory (SALT) 4*.

Farkas, R.; Vincze, V.; Móra, G.; Csirik, J.; and Szarvas, G.

2010. The CoNLL-2010 Shared Task: Learning to Detect Hedges and their Scope in Natural Language Text.

Hintikka, J. 2002. Negation in Logic and in Natural Language. *Linguistics and Philosophy* 25(5/6).

Horn, L. R., and Kato, Y., eds. 2000. *Negation and Polarity - Syntactic and Semantic Perspectives (Oxford Linguistics)*. Oxford University Press, USA.

Horn, L. R. 1989. *A Natural History of Negation*. University Of Chicago Press.

Huddleston, R. D., and Pullum, G. K. 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press.

Morante, R., and Daelemans, W. 2009. Learning the Scope of Hedge Cues in Biomedical Texts. In *Proceedings of the BioNLP 2009 Workshop*.

Morante, R., and Sporleder, C., eds. 2010. *Proceedings of the Workshop on Negation and Speculation in Natural Language Processing*.

Özgür, A., and Radev, D. R. 2009. Detecting Speculations and their Scopes in Scientific Text. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*.

Palmer, M.; Gildea, D.; and Kingsbury, P. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics* 31(1):71–106.

Rose, C. P.; Roque, A.; Bhembe, D.; and Vanlehn, K. 2003. A Hybrid Text Classification Approach for Analysis of Student Essays. In *Building Educational Applications Using NLP*.

Sánchez Valencia, V. 1991. *Studies on Natural Logic and Categorical Grammar*. Ph.D. Dissertation, University of Amsterdam.

Saurí, R., and Pustejovsky, J. 2009. FactBank: a corpus annotated with event factuality. *Language Resources and Evaluation* 43(3):227–268–268.

van Munster, E. 1988. The treatment of Scope and Negation in Rosetta. In *Proceedings of the 12th International Conference on Computational Linguistics*.

Vincze, V.; Szarvas, G.; Farkas, R.; Mora, G.; and Csirik, J. 2008. The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes. *BMC Bioinformatics* 9(Suppl 11).

Wiegand, M.; Balahur, A.; Roth, B.; Klakow, D.; and Montoyo, A. 2010. A Survey on the Role of Negation in Sentiment Analysis. In *Proceedings of the NeSp-NLP Workshop*.

Wilson, T.; Wiebe, J.; and Hoffmann, P. 2009. Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational Linguistics* 35(3):399–433.

Zeijlstra, H. 2007. Negation in Natural Language: On the Form and Meaning of Negative Elements. *Language and Linguistics Compass* 1(5):498–518.