

# Understanding Negation in Positive Terms Using Syntactic Dependencies

Zahra Sarabi and Eduardo Blanco

Human Intelligence and Language Technologies Lab  
University of North Texas  
Denton, TX, 76203

zahrasarabi@my.unt.edu, eduardo.blanco@unt.edu

## Abstract

This paper presents a two-step procedure to extract positive meaning from verbal negation. We first generate potential positive interpretations manipulating syntactic dependencies. Then, we score them according to their likelihood. Manual annotations show that positive interpretations are ubiquitous and intuitive to humans. Experimental results show that dependencies are better suited than semantic roles for this task, and automation is possible.

## 1 Introduction

Negation is a complex phenomenon present in all human languages, allowing for the uniquely human capacities of denial, contradiction, misrepresentation, lying, and irony (Horn and Wansing, 2015). Despite negation always being marked—in the absence of a negation cue, statements are positive—acquiring and understanding sentences that contain negation is more challenging than those that do not. Children acquire negation after learning to communicate (Nordmeyer and Frank, 2013), and adults take longer to process negated statements than positive ones (Clark and Chase, 1972).

In any given language, humans communicate in positive terms most of the time, and use negation to express something unusual or an exception (Horn, 1989). Albeit most sentences are affirmative, negation is ubiquitous (Morante and Sporleder, 2012): In scientific papers, 13.76% of statements contain a negation (Szarvas et al., 2008); in product reviews, 19% (Councill et al., 2010); and in Conan Doyle stories, 22.23% (Morante and Daelemans, 2012). In

OntoNotes (Hovy et al., 2006), 10.15% of statements contain a verb negated with *not*, *n't* or *never*.

From a theoretical point of view, it is accepted that negation conveys positive meaning (Rooth, 1992; Huddleston and Pullum, 2002). For example, when reading (1) *John didn't order the right parts*, humans intuitively understand that (1a) *John ordered something*, or more specifically, (1b) *John ordered the wrong parts*. Interpretation (1a) can be obtained after determining that *n't* does not negate verb *order*, but its THEME, i.e., *the right parts*. Interpretation (1b) can be obtained after determining that *n't* is actually negating *right*, an adjective modifying *parts*.

Determining which words are intended to be negated—identifying the foci of negation, thereby revealing positive interpretations—is challenging. First, as exemplified in (1a, 1b), there is a granularity continuum yielding interpretations that entail each other, e.g., (1b) entails (1a). Second, a single negation often yields several positive interpretations, e.g., from (2) *John doesn't eat meat*, we can extract that (2a) *John eats something other than meat* and (2b) *Some people eat meat, but not John*.

This paper presents a methodology to extract positive interpretations from verbal negation. The main contributions are: (1) deterministic procedure to generate potential interpretations by manipulating syntactic dependencies; (2) analysis showing that dependencies yield finer-grained interpretations and better results than previous work using semantic roles; (3) a corpus of negations and their positive interpretations;<sup>1</sup> and (4) experimental results with gold-standard and predicted linguistic information.

<sup>1</sup>Available at <http://www.cse.unt.edu/~blanco/>

## 2 Terminology, Scope and Focus

Negation is well-understood in grammars, which detail the valid ways to form a negation (Quirk et al., 2000; van der Wouden, 1997). Negation can be expressed by verbs (e.g., *avoid* running), nouns (e.g., the *absence* of evidence), adjectives (e.g., it is *pointless*), adverbs (e.g., I *never* tried Persian food before), prepositions (e.g., you can exchange it *without* a problem), determiners (e.g., the new law has *no* direct implications), pronouns (e.g., *nobody* will keep election promises), and others. In this paper, we focus on verbal negation, i.e., when the negation mark—usually an adverb such as *never* and *not*—is grammatically associated with a verb.

**Positive Interpretations.** In philosophy and linguistics, it is generally accepted that negation conveys positive meaning (Horn, 1989). This positive meaning ranges from implicatures, i.e., what is suggested in an utterance even though neither expressed nor strictly implied (Blackburn, 2008), to entailments. Other terms used in the literature include implied meanings (Mitkov, 2005), implied alternatives (Rooth, 1985) and semantically similars (Agirre et al., 2013). We do not strictly fit into any of this terminology, we reveal positive interpretations as intuitively done by humans when reading text.

### 2.1 Scope and Focus

From a theoretical perspective, it is accepted that negation has scope and focus, and that the focus—not just the scope—yields positive interpretations (Horn, 1989; Rooth, 1992; Taglicht, 1984). Scope is “the part of the meaning that is negated” and focus “the part of the scope that is most prominently or explicitly negated” (Huddleston and Pullum, 2002).

Consider the following statement in the context of the recent refugee crisis: (2) *Mr. Haile was not looking for heaven in Europe*. By definition, scope refers to “all elements whose individual falsity would make the negated statement strictly true”, and focus is “the element of the scope that is intended to be interpreted as false to make the overall negative true” (Huddleston and Pullum, 2002). The falsity of any of the truth conditions below makes statement (2) true, thus the scope of the negation is (2a–2d):

2a. Somebody was looking for something somewhere. [verb *looking*]

2b. Mr. Haile was looking for something somewhere. [AGENT of looking, *Mr. Haile*]

2c. Somebody was looking for heaven somewhere. [THEME of looking, *heaven*]

2d. Somebody was looking for something in Europe. [LOCATION of looking, *in Europe*]

Determining the focus is almost always more challenging than the scope. The challenge relies on determining which of the truth conditions (2a–2d) is intended to be interpreted as false to make the negated statement true: all of them qualify, but some are more likely. A natural reading of statement (2) suggests that *Mr. Haile was looking for something (a regular life, a job, etc.) in Europe, but not heaven*. Determining that the focus is *heaven*, i.e., that everything in statement (2) is positive except the THEME of *looking*, is the key to reveal the intended positive interpretation. Note that scope on its own does not identify positive interpretations, and other foci yield unlikely positive interpretations, e.g., *Mr. Haile was looking for heaven somewhere, but not in Europe*.

It is worth noting that while scope is defined from a logical standpoint, in most negations there are several possible foci and corresponding positive interpretations. For example, given (3) *Most jobs now don't last for decades*, the following are valid positive interpretations: (3a) *Few jobs now last for decades*, (3b) *Most jobs in the past lasted for decades*, and (3c) *Most jobs now last for a few years*.

**Granularity of Focus.** The definition of focus does not provide guidelines about identifying the element of the scope that is the focus. The larger the focus, the more generic the corresponding positive interpretation; and the smaller the focus, the more specific the corresponding positive interpretation. Let us consider statement (3) again. A possible focus is *Most jobs*, yielding the positive interpretation *Something now lasts for decades, but not most jobs*. Another possible focus is *Most*, yielding the interpretation *Few (not most) jobs now last for decades*. We argue that the latter is preferable, as it yields a more specific interpretation and it entails the former: if some jobs last for decades, then something lasts for decades, but not the other way around.

We use the term *coarse-grained focus* to refer to foci that include *all* tokens belonging to an argument of a verb (e.g., *Most Jobs* above), and *fine-grained focus* to refer to foci that do not (e.g., *Most* above).

### 3 Previous Work

Within computational linguistics, approaches to process negation are shallow, or target scope and focus detection. Popular semantic representations such as semantic roles (Palmer et al., 2005; Baker et al., 1998) or AMR (Banarescu et al., 2013) do not reveal the positive interpretations we target in this paper. Shallow approaches are usually application-specific. In sentiment and opinion analysis, negation has been reduced to marking as negated all words between a negation cue and the first punctuation mark (Pang et al., 2002), or within a five-word window of a negation cue (Hu and Liu, 2004). The examples throughout this paper show that these techniques are insufficient to reveal implicit positive interpretations.

#### 3.1 Scope Annotations and Detection

Scope of negation detection has received a lot of attention, mostly using two corpora: BioScope in the medical domain (Szarvas et al., 2008) and CD-SCO (Morante and Daelemans, 2012). BioScope annotates negation cues and linguistic scopes exclusively in biomedical texts. CD-SCO annotates negation cues, scopes, and negated events or properties in selected Conan Doyle stories.

There have been several supervised proposals to detect the scope of negation using BioScope and CD-SCO (Özgür and Radev, 2009; Øvrelid et al., 2010). Automatic approaches are mature (Abujbara and Radev, 2012): F-scores are 0.96 for negation cue detection, and 0.89 for negation cue and scope detection (Velldal et al., 2012; Li et al., 2010). ?) present the best results to date using CD-SCO, and analyze the main sources of errors. Outside BioScope and CD-SCO, Reitan et al. (2015) present a negation scope detector for tweets, and show that it improves sentiment analysis. As shown in Section 2, scope detection is insufficient to reveal positive interpretations from negation.

#### 3.2 Focus Annotation and Detection

While focus of negation has been studied for decades in philosophy and linguistics (Section 2), corpora and automated tools are scarce. Blanco and Moldovan (2011) annotate focus of negation in the 3,993 negations marked with ARG-M-NEG semantic role in PropBank (Palmer et al., 2005). Their an-

notations, PB-FOC, were used in the \*SEM-2012 Shared Task (Morante and Blanco, 2012). Their guidelines require annotators to choose as focus the semantic role that “is most prominently negated” or the verb. If several roles may be the focus, they prioritize “the one that yields the most meaningful implicit [positive] information”, but do not specify what *most meaningful* means. Their approach has 2 limitations. First, because they select one focus per negation, they only extract one positive interpretation per negation. Second, because they select as focus a semantic role, they only consider coarse-grained foci. Consider again statement (3) from Section 2.1. By design, their approach is limited to extract a single interpretation even though interpretations (3a–3c) are valid. Similarly, their approach is limited to select as focus *Most jobs*—all tokens belonging to a semantic role—although *Most* yields a “more meaningful” interpretation: *Something now lasts for decades* (generic, worse) vs. *Few jobs now last for decades* (specific, better).

Blanco and Sarabi (2016) present a complimentary approach to extract and score several positive interpretations from a single verbal negation. Their methodology is grounded on semantic roles and does not consider fine-grained foci. In this paper, we improve upon their work: we extract both coarse- and fine-grained interpretations, and also extract several interpretations from one negation.

Anand and Martell (2012) reannotate PB-FOC and argue that positive interpretations arising from scalar implicatures and neg-raising predicates should be separated from those arising from focus detection. They argue that 27.4% of negations with a focus annotated in PB-FOC do not have one. In this paper, we are not concerned about annotating foci per se, but about extracting positive interpretations from negation, as intuitively done by humans.

Automatic systems to detect the focus of negation yield modest results. Blanco and Moldovan (2011) obtain an accuracy of 65.5 using supervised learning and features derived from gold-standard linguistic information. With predicted linguistic information, ?) report an F-measure of 58.4 using 4 linguistically sound heuristics, and ?) an F-measure of 65.62 using contextual discourse information. Blanco and Sarabi (2016) obtain Pearson correlation of 0.642 ranking coarse-grained interpretations. Unlike the

work presented here, none of these systems extract fine-grained interpretations from a single negation.

## 4 Corpus Creation

Our goal is to create a corpus of negations and their positive interpretations. We put a strong emphasis on automation and simplicity. First, we deterministically generate potential positive interpretations from verbal negations by manipulating syntactic dependencies (Section 4.1). Second, we ask annotators to score potential positive interpretations (Section 4.2). Positive interpretations and their scores are later used to learn models to rank potential interpretations automatically (Section 6). Generating potential interpretations deterministically prior to scoring them proved very beneficial. After pilot experiments, it became clear that asking annotators to propose positive interpretations complicates the annotation effort (lower agreements) as well as learning.

We decided to work on top of OntoNotes (Hovy et al., 2006)<sup>2</sup> instead of plain text or other corpora for several reasons. First, OntoNotes includes gold linguistic annotations such as part-of-speech tags, parse trees and semantic roles. Second, unlike BioScope, CD-SCO and PB-FOC (Section 3.2), OntoNotes includes sentences from several genres, e.g., newswire, broadcast news and conversations, magazines, the web. We transformed the parse trees in OntoNotes into syntactic dependencies using Stanford CoreNLP (Manning et al., 2014).

### 4.1 Manipulating Syntactic Dependencies to Generate Potential Positive Interpretations

OntoNotes contains 63,918 sentences. Annotating all positive interpretations from all negations is outside the scope of this paper. Instead, we target selected representative negations.

**Selecting Negations.** We first select all verbal negations by retrieving all tokens whose syntactic head is a verb and dependency type *neg*.<sup>3</sup> Then, we discard negations from sentences that contain two negations, conditionals, commas or questions. Finally, we discard negations if the negated verb is *to be* or it does not have a subject (dependency *nsubj* or *nsubjpass*).

<sup>2</sup>We use the CoNLL-2011 Shared Task distribution (Pradhan et al., 2011), <http://conll.cemantix.org/2011/>

<sup>3</sup>The Stanford manual describes and exemplifies all syntactic dependencies (de Marneffe and Manning, 2008).

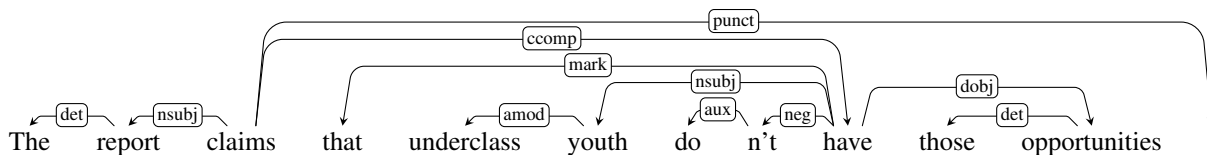
**Converting Negated Statements into their positive counterparts.** We apply 3 steps inspired after the grammatical rules to form negation detailed by Huddleston and Pullum (2002, Ch. 9):

1. Remove the negation mark by deleting the token with syntactic dependency *neg*.
2. Remove auxiliaries, expand contractions, and fix third-person singular and past tense. For example (before: after), *doesn't go: goes, didn't go: went, won't go: will go*. We loop through the tokens whose head is the negated verb with dependency *aux*, and use a list of irregular verbs and grammar rules to convert to third-person singular and past tense.
3. Rewrite negatively-oriented polarity-sensitive items. For example (before: after), *anyone: someone, any longer: still, yet: already, at all: somewhat*. We use the correspondences between negatively-oriented and positively-oriented polarity-sensitive items by (Huddleston and Pullum, 2002, pp. 831).

**Selecting Relevant tokens.** Verbal negation often occurs in multi-clause sentences. In order to identify the relevant (syntactically negated) eventuality, we simplify the original statement by including only the negated verb and all tokens that are dependents of the verb, i.e., tokens reachable from the negated verb traversing dependencies. For example, from *Individuals familiar with the Justice Department's policy said that Justice officials hadn't any knowledge of the IRS's actions in the last week*, after getting the positive counterpart and selecting relevant tokens, we obtain *Justice officials had some knowledge of the IRS's actions in the last week*.

**Generating Interpretations.** Given the simplified positive counterpart, generating all combinations of tokens as potential foci would result in  $2^t$  potential positive interpretations for  $t$  tokens. To avoid a brute-force approach that generates many nonsensical potential interpretations, we define a procedure grounded on syntactic dependencies.

The main idea is to run a modified breadth-first traversal of the dependency tree to select subtrees that are potential foci. We start the traversal from the negated verb and stop it at depth 3, selecting as potential foci the subtrees rooted at all tokens except those whose syntactic dependency is *aux*, *auxpass* or *punct* (auxiliary, passive auxiliary and punctua-



Negated statement: The report claims that underclass youth don't have those opportunities.			
Positive counterpart	Step 1	The report claims that underclass youth do have those opportunities.	
	Step 2	The report claims that underclass youth have those opportunities.	
	Step 3	The report claims that underclass youth have those opportunities. (idem)	
Relevant tokens	Underclass youth have those opportunities.		
Potential positive interpretations	<i>none</i>	coarse	Underclass youth [some verb] those opportunities, <i>but not have</i> .
	<i>nsubj</i>	coarse	[Some people] have those opportunities, but not <i>Underclass youth</i> .
	<i>amod</i>	fine	[Some adjective] youth have those opportunities, but not <i>Underclass youth</i> .
	<i>nsubj</i>	fine	Underclass [some people] have those opportunities, but not <i>Underclass youth</i> .
	<i>dobj</i>	coarse	Underclass youth have [something], but not <i>those opportunities</i> .
	<i>det</i>	fine	Underclass youth have [some] opportunities, but not <i>those opportunities</i> .
	<i>dobj</i>	fine	Underclass youth have those [something], but not <i>those opportunities</i> .

**Table 1:** Negated statement and syntactic dependencies (top), and automatically generated positive counterpart and potential positive interpretations (bottom). For potential interpretations, we include the dependency from the focus to the rest of the interpretation.

tion). Additionally, we discard potential foci that consist only of (1) the determiners *the*, *a* and *an*, or (2) a single token with part-of-speech tag TO, CC, UH, POS, XX, IN, WP or dependency relation *prt*. These rules were defined after manually observing several examples and concluding that the corresponding positive interpretation was useless. For example, from the negated statement *And our credit standards haven't changed one iota*, we avoid generating the useless potential interpretation *Our credit standards X changed one iota, but not have changed*. (focus would be *have*, with dependency *aux*). Similarly, from *It is not supported by the text or history of the Constitution*, we avoid generating potential interpretation *It is supported by X text or history of the Constitution, but not by the text or history of the Constitution* (focus would be *the*); and from *You don't want to get yourself too upset about these things*, potential interpretation *You want X get yourself too upset about these things, but not to get* (focus would be *to*, with part-of-speech tag TO).

Once potential foci are selected, we generate positive interpretations by rewriting each focus with “someone/some people/something/etc.” and appending “but not text\_of\_focus” at the end. Additionally, if the first token of the focus is a preposition, we include it to improve readability, e.g., *didn't leave [by noon]: left by sometime, but not by noon*.

Note that potential interpretations obtained from

foci that are direct syntactic dependents of the negated verb are coarse-grained interpretations, and the rest are fine-grained interpretations. Table 1 exemplifies the procedure step by step.

## 4.2 Scoring Potential Positive Interpretations

After generating potential positive interpretations automatically, we asked annotators to score them. Annotators had access to the original negated sentence, the previous and next sentence as context, and one potential positive interpretation at a time. The interface asked *Given the three sentences [previous sentence, negated sentence and next sentence] above, do you think the statement [positive interpretation] below is true?* Annotators were forced to answer with a score from 0 to 5, where 0 means *absolutely disagree* and 5 means *absolutely agree*. We did not provide descriptions for intermediate scores or use categorical labels. These simple guidelines were sufficient to reliably score plausible positive interpretations automatically generated (Section 5).

## 5 Corpus Analysis

The procedure described in Section 4.1 generates 9729 potential positive interpretations (5865 coarse-grained and 3864 fine-grained) from 1671 verbal negations. Out of all these potential positive interpretations, we annotate 1700 (1008 coarse- and 692 fine-grained). Overall, the mean score is 3.20, and the standard deviation is 1.66. Table 2 shows basic

	Negated statement, context if relevant to determining scores, and all positive interpretations	Score
1	Context, previous statement: You're not giving me enough benefits. Negated Statement: You're not paying me for my overtime work. Context, next statement: Well I think the Walton family does take it personally.	
	- Int. 1.1 [coarse, root]: You're [some verb] me for my overtime work, but not <i>paying</i> .	4
	- Int. 1.2 [coarse, nsubj]: [Some people]'re paying me for my overtime work, but not <i>you</i> .	0
	- Int. 1.3 [coarse, dobj]: You're paying [somebody] for my overtime work, but not <i>me</i> .	1
	- Int. 1.4 [coarse, prep]: You're paying me for [something], but not <i>for my overtime work</i> .	5
	- Int. 1.5 [fine, poss]: You're paying me for [somebody's] overtime work, but not <i>for my overtime work</i> .	0
	- Int. 1.6 [fine, nn]: You're paying me for my [some adjective] work, but not <i>for my overtime work</i> .	5
	- Int. 1.7 [fine, pobj]: You're paying me for my overtime [something] but not <i>for my overtime work</i> .	0
2	Negated Statement: Those concerns aren't expressed in public.	
	- Int. 2.1 [coarse, root]: Those concerns are [some verb] in public, but not <i>expressed</i> .	5
	- Int. 2.2 [coarse, nsubjpass]: [Some things] are expressed in public, but not <i>Those concerns</i> .	5
	- Int. 2.3 [fine, nsubjpass]: Those [some noun] are expressed in public, but not <i>Those concerns</i> .	2
	- Int. 2.4 [fine, det]: [Some] concerns are expressed in public but, not <i>Those concerns</i> .	4
	- Int. 2.5 [coarse, prep]: Those concerns are expressed in [somewhere], but not <i>in public</i> .	5

**Table 3:** Negated statements, all potential positive interpretations automatically generated and their manually assigned scores.

Dependency	#	%	Mean	SD
<i>nsubj</i>	358	21.13%	3.36	1.47
<i>dobj</i>	237	14.05%	3.73	1.57
<i>pobj</i>	178	10.51%	3.48	1.59
<i>ccomp</i>	125	7.29%	3.29	1.77
<i>advmod</i>	108	6.39%	3.33	1.59
<i>xcomp</i>	90	5.28%	3.92	1.50
<i>amod</i>	67	3.96%	4.08	1.29
<i>conj</i>	40	2.38%	2.80	1.60
<i>advcl</i>	40	2.32%	2.84	1.80
<i>nsubjpass</i>	35	2.17%	3.63	1.51
other	209	13.34%	2.9	1.7
verb	213	12.5%	2.01	1.46
All	1,700	100.00%	3.20	1.66

**Table 2:** Basic corpus analysis. For each dependency, we show the number of potential interpretations generated (#) and percentage (%), mean score and standard deviation.

statistics for potential foci, where *dependency* indicates the dependency from the potential focus to a token outside the potential focus. Most foci are *nsubj*, *dobj* and *pobj*, and the mean scores and standard deviation are similar for most dependencies.

**Annotation Quality.** In order to ensure annotation quality, we calculated Pearson correlation. Kappa and other measures designed for categorical labels are ill-suited for our annotations, since not all disagreements between numeric scores are the same, e.g., 4 vs. 5 should be counted as higher agreement, than 1 vs. 5. Overall Pearson correlation was 0.75.

## 5.1 Annotation Examples

Table 3 presents 2 statements that contain verbal negation, the list of positive interpretations automatically generated and the annotated scores.

Example (1) is a simple negated clause, yet we generate 7 potential positive interpretations and 3 of them receive high scores (4 or 5). Given *You're not paying me for my overtime work* and the previous statement, it is reasonable to believe that the author is in an employee-employer relationship, and the employer is not fair to the employee. Interpretations 1.1, 1.4 and 1.6 are implicit positive interpretations intuitively understood by humans when reading the original negated statement. Namely, Interpretation 1.1: *You (the employer) are nickel-and-diming me for my overtime work* (focus is *paying*), Interpretation 1.4: *You (the employer) are paying me for something* (focus is *my overtime work*), and Interpretation 1.6: *You (the employer) are paying me for my regular work* (focus is *overtime*). These interpretations show the benefits of fine-grained interpretations: Interpretation 1.6 is a refinement of Interpretation 1.4, and the former is more desirable than the latter as it reveals more specific positive knowledge. The remaining interpretations are legible, but do not make sense given the negated statement, e.g., interpretation 1.2: *Somebody (but not the employer) pays me for my overtime* (focus is *You*).

Example (2) is also a simple negated clause, and 4 out of 5 interpretations receive high scores, capturing valid positive meaning. Specifically, Interpreta-

Type	Name	Description
Basic	neg_mark	word form of negation mark
	verb	word form and part-of-speech tag of verb
	coarse_or_fine	flag indicating whether interpretation is coarse- or fine-grained
Path	syn_path_dep	syntactic path from focus to verb (concatenation of dependencies)
	syn_path_pos	syntactic path from focus to verb (concatenation of part-of-speech tags)
	syn_path_last_dep	last syntactic dependency in syn_path_dep (direct dependent of verb)
	syn_path_last_pos	last part-of-speech tag in syn_path_pos (direct dependent of verb)
Focus	focus_length	number of words in subgraph chosen as focus
	focus_first_word	word form and part-of-speech tag of first word in focus
	focus_last_word	word form and part-of-speech tag of last word in focus
	focus_direction	flag indicating whether focus occurs before or after verb
	focus_head_word	word form of head of focus
	focus_head_pos	part-of-speech tag of head of focus
	focus_head_rel	syntactic dependency of head of focus

**Table 4:** Features used to score potential positive interpretations automatically generated.

tion 2.1: *Those concerns are avoided in public* (focus is *expressed*), Interpretation 2.2: *Something is expressed in public* (focus is *Those concerns*), Interpretation 2.4: *Some concerns (but not problematic or secret concerns) are expressed in public* (focus is *Those*), and Interpretation 2.5: *Those concerns are expressed in private* (focus is *in public*).

## 5.2 Syntactic Dependencies vs. Semantic Roles

The procedure presented in Section 4.1 is not the first to generate potential positive interpretations from negation (Section 3.2). Our approach has 2 advantages with respect to those grounded on semantic roles (Blanco and Sarabi, 2016): (1) it generates both coarse- and fine-grained interpretations, and (2) learning to score interpretations is easier because state-of-the-art tools extract dependencies more reliably than semantic roles.

To support claim (1), we compare the interpretations generated with our procedure and previous work using semantic roles. 96.12% of interpretations generated using roles are also generated using syntactic dependencies. Also, using dependencies allow us to generate 67.9% of additional (fine-grained) interpretations not obtainable with roles.

To support claim (2), we compare interpretations generated with gold and predicted linguistic information (roles or dependencies). The overlap with semantic roles is 70.1%, and with syntactic dependencies, 92.8%. Syntactic dependencies are thus better in a realistic scenario because they allow us to automatically generate (and score) most interpretations.

## 6 Supervised Learning to Score Potential Positive Interpretations

We follow a standard supervised machine learning approach. The 1,700 potential positive interpretations along with their scores become instances, and we divide them into training (80%) and test splits (20%) making sure that all interpretations generated from a sentence are assigned to either the training or test splits. Note that splitting instances randomly would not be sound: training with some interpretations generated from a negation, and testing with the rest of interpretations generated from the same negation would be an unfair evaluation.

We train a Support Vector Machine for regression with RBF kernel using scikit-learn (Pedregosa et al., 2011), which in turn uses LIBSVM (Chang and Lin, 2011). SVM parameters ( $C$  and  $\gamma$ ) were tuned using 10-fold cross-validation with the training set, and results are calculated using the test set.

### 6.1 Feature Selection

Table 4 presents the full feature set. Features are relatively simple and characterize the verbal negation from which a potential interpretation was generated, as well as the interpretation per se, i.e., the dependency subgraph chosen as potential focus.

*Basic* features account for the negation mark, the negation verb (word form and part-of-speech tag) and a binary flag indicating whether we are scoring a coarse- or fine-grained interpretation.

*Path* features are derived from the syntactic path

Features	Gold	Predicted
neg_mark	-0.109	-0.077
basic	0.033	0.026
basic + path	0.474	0.482
basic + path + focus	0.530	0.560

**Table 5:** Pearson correlations obtained with the test split. Results are provided using gold-standard and predicted linguistic information (part-of-speech tags and syntactic dependencies).

between the subgraph selected as focus and the verb. We include the actual path (concatenation of dependencies and up/down symbols), and the modified path using part-of-speech tags. Additionally, we also include the last dependency and part-of-speech tag, i.e., the ones closest to the verb in the path.

*Focus* features characterize the dependency subgraph chosen as focus to generate the potential interpretation. Specifically, we include the number of tokens, word form and part-of-speech tags of the first and last tokens, and whether the focus occurs before or after the verb. We also include features derived from the head of the focus, which we define as the token whose syntactic head is outside the focus. We include the word form and part-of-speech of the focus head, as well as its the dependency.

## 7 Experiments and Results

We report results obtained with several combinations of features in Table 5. We detail results obtained with features extracted from gold-standard and predicted linguistic annotations (part-of-speech tags and syntactic dependencies) as annotated in the *gold* and *auto* files from the CoNLL-2011 Shared Task release of OntoNotes (Pradhan et al., 2011). All models are trained with gold-standard linguistic annotations, and tested with either gold-standard or predicted linguistic annotations.

**Testing with gold-standard POS tags and syntactic dependencies.** Training with the word form of the negation mark is virtually useless, it yields a Pearson correlation of  $-0.109$ . *Basic* features (negation mark, verb and flag indicating coarse- or fine-grained interpretation) are also ineffective to score potential interpretations (Pearson: 0.033). Including features derived from the syntactic *path* yields higher correlation, 0.474, even though these features only capture the syntactic relationship be-

tween the focus from which the interpretation was generated and the verb. Finally, adding *focus* features yields the best results (Pearson: 0.53, +11.8%).

**Testing with predicted POS tags and syntactic dependencies.** We selected 20% of positive interpretations in our corpus as test instances, totalling 379 interpretations (Section 6). When executing the procedure to generate potential interpretations (Section 4.1) with predicted linguistic information, however, we are unable to generate all of them due to incorrect and missing syntactic dependencies. Specifically, 352 of the 379 interpretations are generated (92.8%). While we do not generate 7.2% of instances, this percentage is substantially lower than previous work grounded on semantic roles (Section 5.2).

Pearson correlations with predicted linguistic information are calculated using the 352 instances that were also generated with gold dependencies (and thus assigned a score during the manual annotations). Correlations are slightly higher and follow a similar trend than the correlations obtained with gold-standard linguistic information. These results should be taken with a grain of salt: the test instances are not exactly the same, and the 352 test instances in this scenario are presumably easier to score than the remainder 27, as dependencies were predicted correctly.

## 8 Conclusions

Humans intuitively extract positive meaning from negation when reading text. This paper presents an automated procedure to generate potential positive interpretations from verbal negation, and score them according to their likelihood. Our procedure is grounded on syntactic dependencies, allowing us to extract fine-grained interpretations beyond semantic roles (67.9% additional interpretations). Additionally, because dependencies are extracted automatically more reliably than semantic roles, we generate 92.8% of all potential interpretations when using predicted linguistic information, as opposed to 70.1% with semantic roles.

On average, we generate 6.4 potential interpretations per verbal negation (coarse-grained: 3.8, fine-grained: 2.6). Manual annotations show that potential interpretations are deemed likely. The mean



score is 3.20 (out of 5.0), thus we extract a substantial amount of positive meaning.

The work presented in this paper is not tied to any existing semantic representation. While we rely heavily on syntactic dependencies, positive interpretations are generated in plain text, and they could be processed, along with the original negated statement, with any NLP pipeline.

## References

- Amjad Abu-Jbara and Dragomir Radev. 2012. Umichigan: A conditional random field model for resolving the scope of negation. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*, pages 328–334. Association for Computational Linguistics.
- Eneko Agirre, Daniel Cer, Mona Diab, Aitor Gonzalez-Agirre, and Weiwei Guo. 2013. \*sem 2013 shared task: Semantic textual similarity. In *Second Joint Conference on Lexical and Computational Semantics (\*SEM), Volume 1: Proceedings of the Main Conference and the Shared Task: Semantic Textual Similarity*, pages 32–43, Atlanta, Georgia, USA, June. Association for Computational Linguistics.
- Pranav Anand and Craig Martell. 2012. Annotating the focus of negation in terms of questions under discussion. In *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics*, ExProM '12, pages 65–69, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet Project. In *Proceedings of the 17th international conference on Computational Linguistics*, Montreal, Canada.
- Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider. 2013. Abstract meaning representation for sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, pages 178–186, Sofia, Bulgaria, August. Association for Computational Linguistics.
- Simon Blackburn. 2008. *The Oxford Dictionary of Philosophy*. Oxford University Press.
- Eduardo Blanco and Dan Moldovan. 2011. Semantic representation of negation using focus detection. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 581–589, Portland, Oregon, USA, June. Association for Computational Linguistics.
- Eduardo Blanco and Zahra Sarabi. 2016. Automatic generation and scoring of positive interpretations from negated statements. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1431–1441, San Diego, California, June. Association for Computational Linguistics.
- Chih-Chung Chang and Chih-Jen Lin. 2011. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27:1–27:27, May.
- H. H. Clark and W. G. Chase. 1972. On the process of comparing sentences against pictures. *Cognitive Psychology*, 3(3):472–517, July.
- Isaac Council, Ryan McDonald, and Leonid Velikovich. 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 Processing*, pages 51–59, Uppsala, Sweden, July. University of Antwerp.
- Marie-Catherine de Marneffe and Christopher D Manning. 2008. Stanford typed dependencies manual. Technical report, Technical report, Stanford University.
- Laurence R. Horn and Heinrich Wansing. 2015. Negation. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Summer 2015 edition.
- Laurence R. Horn. 1989. *A natural history of negation*. Chicago University Press, Chicago.
- Eduard Hovy, Mitchell Marcus, Martha Palmer, Lance Ramshaw, and Ralph Weischedel. 2006. OntoNotes: the 90% Solution. In *NAACL '06: Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers on XX*, pages 57–60, Morristown, NJ, USA. Association for Computational Linguistics.
- Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177, New York, NY, USA. ACM.
- Rodney D. Huddleston and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press, April.
- 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 2010)*, pages 671–679, Beijing, China, August. Coling 2010 Organizing Committee.
- Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky.

2014. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, pages 55–60.
- Ruslan Mitkov. 2005. *The Oxford handbook of computational linguistics*. Oxford University Press.
- 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)*, pages 265–274, Montréal, Canada, June.
- Roser Morante and Walter Daelemans. 2012. Conandoyle-neg: Annotation of negation in conandoyle stories. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation, Istanbul*.
- Roser Morante and Caroline Sporleder. 2012. Modality and negation: An introduction to the special issue. *Comput. Linguist.*, 38(2):223–260, June.
- Ann E Nordmeyer and Michael C Frank. 2013. Measuring the comprehension of negation in 2-to 4-year-old children. *Proceedings of the 35th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society*.
- Lilja Øvrelid, Erik Velldal, and Stephan Oepen. 2010. Syntactic Scope Resolution in Uncertainty Analysis. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, pages 1379–1387, Beijing, China, August. Coling 2010 Organizing Committee.
- Arzucan Özgür and Dragomir R. Radev. 2009. Detecting Speculations and their Scopes in Scientific Text. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 1398–1407, Singapore, August. Association for Computational Linguistics.
- Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106.
- Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing*, pages 79–86. Association for Computational Linguistics, July.
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Sameer Pradhan, Lance Ramshaw, Mitchell Marcus, Martha Palmer, Ralph Weischedel, and Nianwen Xue. 2011. Conll-2011 shared task: Modeling unrestricted coreference in ontonotes. In *Proceedings of the Fifteenth Conference on Computational Natural Language Learning: Shared Task*, pages 1–27, Portland, Oregon, USA, June. Association for Computational Linguistics.
- Randolph Quirk, Sidney Greenbaum, and Geoffrey Leech. 2000. *A comprehensive grammar of the English language*. Longman, London.
- Johan Reitan, Jørgen Faret, Björn Gambäck, and Lars Bungum. 2015. Negation scope detection for twitter sentiment analysis. In *Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 99–108, Lisboa, Portugal, September. Association for Computational Linguistics.
- Mats Rooth. 1985. Association with focus.
- Mats Rooth. 1992. A theory of focus interpretation. *Natural language semantics*, 1(1):75–116.
- György Szarvas, Veronika Vincze, Richárd Farkas, and János Csirik. 2008. The BioScope corpus: annotation for negation, uncertainty and their scope in biomedical texts. In *Proceedings of BioNLP 2008*, pages 38–45, Columbus, Ohio, USA. ACL.
- Josef Taglicht. 1984. *Message and emphasis: On focus and scope in English*, volume 15. Addison-Wesley Longman Limited.
- Ton van der Wouden. 1997. *Negative contexts: collocation, polarity, and multiple negation*. Routledge, London.
- Erik Velldal, Lilja Øvrelid, Jonathon Read, and Stephan Oepen. 2012. Speculation and negation: Rules, rankers, and the role of syntax. *Comput. Linguist.*, 38(2):369–410, June.