Abstract

This paper presents a framework to understand negation in positive terms. Specifically, we extract positive meaning from negation when the negation cue syntactically modifies a noun or adjective. Our approach is grounded on generating potential positive interpretations automatically, and then scoring them. Experimental results show that interpretations scored high can be reliably identified.

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). Acquiring and understanding negation poses unique challenges. For example, children acquire negation after learning to communicate in positive terms (Nordmeyer and Frank, 2013), and adults take longer to process sentences containing negation (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). But negation is ubiquitous (Morante and Sporleder, 2012): In scientific papers, 13.76% of sentences contain a negation (Szavas et al., 2008); in product reviews, 19% (Councill et al., 2010); and in Conan Doyle stories, 22.23% (Morante and Daelemans, 2012).

From a theoretical perspective, it is accepted that negation has scope and focus, and that humans intuitively understand positive meanings from negation (Rooth, 1992; Huddleston and Pulham, 2002). For example, from (1) John didn’t earn a steady paycheck until he was 40 years old, humans understand that (1a) John earned unsteady paychecks before he was 40 years old, and that (1b) John earned steady paychecks when he was 40 years old. This kind of positive interpretations would benefit language understanding in general. For example, a question answering system would benefit from interpretation (1b) when answering question Did John ever earn a steady paycheck?

Within computational linguistics, automated approaches to extract positive meanings from negation target verbal negation (Section 3), i.e., when the negation cue is grammatically associated with a verb, as in (1). Verbal negation accounts only for a portion of all negations, e.g., out of all syntactic dependencies indicating a negation modifier (neg dependency) in OntoNotes (Hovy et al., 2006), 64.4% modify verbs, 19.6% nouns, 10.3% adjectives, and 5.7% other part-of-speech tags. Non-verbal negation also conveys positive meanings, e.g., from (2) No media were allowed inside the venue (No modifies noun media), humans understand that (2a) Somebody (e.g., invited guests) were allowed inside the venue and that (2b) Media were allowed somewhere outside the venue (presumably in a designated press area).

Similarly, from (3) She was not alive when she got to the Lafayette area (not modifies adjective alive), humans understand that (3a) She was dead when she got to the Lafayette area and that (3b) She was alive before she got to the Lafayette area.

This paper presents new corpora and experimental results to extract positive interpretations from negation when the negation cue modifies a noun or adjective. The main contributions are: (1) analysis of negation in OntoNotes beyond verbal negation; (2) procedure to automatically generate potential positive interpretations from non-verbal negation, specifically, when the negation cue modifies a noun or adjective; (3) annotations validating and scoring potential interpretations according to their likelihood;\(^1\) and (4) experimental results showing that the task can be automated.

\(^1\)Available at http://www.cse.unt.edu/~blanco/
2 Terminology and Background

Negation can be expressed by verbs (e.g., avoid the highway), nouns (e.g., lack of knowledge), adjectives (e.g., it is useless), adverbs (e.g., John never drives on the highway), and others (van der Wouden, 1997). The primary negative prefixes in English are in-, il-, im-, ir-, an-, non-, anti- and a- (Garner, 2009, p. 563). We refer to the token, prefix or suffix that indicates negation (emphasized in the examples above) as negation cue. As we shall see, in this paper we target negations whose cue syntactically modifies a noun or adjective.

In philosophy and linguistics, it is generally accepted that negation conveys positive meanings (Horn, 1989). We use the term positive interpretation to refer to the positive meaning intuitively understood by humans when reading sentences that contain negation. Positive interpretations range from implicatures (Blackburn, 2008), to entailments. Potential positive interpretations are positive interpretations whose validity is unknown.

Scope and Focus. Negation is generally understood in terms of scope and focus. Scope is “the part of the meaning that is negated” and focus is “the part of the scope that is most prominently or explicitly negated” (Huddleston and Pullum, 2002). Scope and focus are not exclusive of negation. Among many others, there has been work on detecting the scope of uncertainty cues (Farkas et al., 2010), and other focus-sensitive phenomena include adverbs and conditionals (Rooth, 1985).

Consider statement (2) again, No media were allowed inside the venue. 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). If any of the truth conditions below were false, statement (2) would be true, thus the scope of the negation marked with No is (2a-2c):

2a. Some people were allowed somewhere.
2b. Media were allowed somewhere.
2c. Some people were allowed inside the venue.

Choosing the element of the scope which is the focus is more challenging than identifying the scope. A natural reading of statement (2) indicates that there were people (e.g., invited guests) allowed inside the venue, and that media were (probably) allowed in a press area outside (but not far from) the venue. The former positive interpretation corresponds to choosing Media as focus, and the latter corresponds to choosing inside as focus. Statement (2) exemplifies two core properties of the work presented here: We choose several foci for a single negation, and as a result, reveal several positive interpretations per negation. Further, we see positive interpretations as probabilistic knowledge, i.e., knowledge that may be likely but not necessarily certain.

In general, the task of identifying foci and revealing positive interpretations is natural to humans, but hard to automate. Consider modified statement (2’) No media were allowed inside the venue to record the presentation. The scope is conditions (2a-2c) and (2d) Somebody was allowed to record the presentation. The positive interpretations from (2’) are different than from (2), e.g., Media were allowed inside the venue, but they weren’t allowed to record the presentation can only be extracted from (2’).

3 Previous Work

Within computational linguistics, most approaches to process negation target scope or focus detection. Generally speaking, there are corpora with scope annotations for all types of negations, but corpora with focus annotations are restricted to verbal negation, i.e., when the negation cue is grammatically associated with a verb.

Scope of Negation. There are two main corpora with scope of negation annotations: BioScope in the medical domain (Szarvas et al., 2008) and CD-SCO (Morante and Daelemans, 2012). The annotations schemas differ substantially; CD-SCO annotates negation cues, their scopes, and the negated events or properties. There have been several supervised proposals to detect the scope of negation using BioScope and CD-SCO (Morante and Daelemans, 2009; Velldal et al., 2012; Basile et al., 2012). Fancellu et al. (2016) present the best results to date with CD-SCO using neural networks. They also perform out-of-domain evaluation with new annotations on Wikipedia, and analyze the main sources of errors.

Outside BioScope and CD-SCO, Reitan et al. (2015) present a scope detector for negation in tweets, and use it for sentiment analysis.

As the examples throughout this paper show (e.g., Section 2), detecting the scope of negation is insufficient to reveal the positive interpretations we target in this work.
Focus of Negation and Positive Interpretations.
Identifying the focus of negation is equivalent to revealing positive interpretations—everything but the focus is actually positive. The definition of focus does not specify annotation guidelines, and most existing efforts are grounded on semantic roles. Blanco and Moldovan (2011) annotate focus on the negations marked with ARGm-NEG role in PropBank (Palmer et al., 2005). They select a single focus per negation, specifically, they select the role that reveals the “most useful [positive] information.” Anand and Martell (2012) refine these annotations and differentiate positive interpretations arising from focus identification, scalar implicature and neg-raising predicates. Blanco and Sarabi (2016) propose a similar approach that scores the likelihood of several potential foci per negation. The main limitations of all these previous works is that selecting as focus a semantic role is only suitable when the negation cue modifies a predicate, and roles often yield coarse-grained interpretations. Sarabi and Blanco (2016) bypass these drawbacks by working with syntactic dependencies to refine coarse-grained interpretations.

All these previous efforts to reveal positive interpretations from negation target exclusively verbal negation, i.e., when the negation cue modifies a verb. While verbal negation is more frequent (64.4% of neg dependencies in OntoNotes, Section 4), in this paper we target two understudied yet important negations: when the negation cue modifies a noun or adjective (19.6% and 10.3% of neg dependencies). Our approach is not grounded on semantic roles but syntactic dependencies. Doing so allows us to tackle negation when the negation cue modifies nouns or adjectives.

4 Corpus Creation
We create a corpus of negations and their positive interpretations following three steps. First, we select negations whose negation cue syntactically modifies either a noun or adjective. Second, we automatically generate potential positive from those negations by manipulating syntactic dependencies and part-of-speech tags. Third, we gather manual annotations to validate and score potential interpretations. While asking annotators to suggest positive interpretations would potentially yield more natural interpretations, we found experimentally that a generate-and-rank approach yields higher quality annotations.

Negation in OntoNotes. Instead of building our corpus from plain text, we decided to work on top of OntoNotes (Hovy et al., 2006), a publicly available corpus including texts in several genres (news, transcripts, magazines, etc.) OntoNotes includes, among other gold linguistic annotations, part-of-speech tags and parse trees. We transformed the parse trees into syntactic dependencies using Stanford CoreNLP (Manning et al., 2014).

We reduce the problem of finding negations to retrieving syntactic dependencies neg, which stands for negation modifier. Doing so ignores negation cues that are prefixes or suffixes (e.g., unlimited, motionless), but also simplifies the process. There are 9,507 neg syntactic dependencies in OntoNotes; 6,120 of them modify verbs (64.4%), 1,866 nouns (19.6%), 979 adjectives (10.3%), and 543 other part-of-speech tags (5.7%). Since verbal negation has been tackled

---

---
No condemned murderer has been granted clemency in California since nineteen sixty-seven.

Negated statement: No condemned murderer has been granted clemency in California since nineteen sixty-seven.
Positive counterpart: A condemned murderer has been granted clemency in California since nineteen sixty-seven.
Relevant tokens: A condemned murderer has been granted clemency in California since nineteen sixty-seven.

Potential positive interpretations:

Intpn. 1, root: A condemned murderer has been granted {some verb} clemency in California since nineteen sixty-seven, but not granted.
Intpn. 2, nsubjpass: (Someone) has been granted clemency in California since nineteen sixty-seven, but not [a] condemned murderer.
Intpn. 3, dobj: A condemned murderer has been granted clemency {some noun} in California since nineteen sixty-seven, but not [a] condemned murderer.
Intpn. 4, prep: A condemned murderer has been granted clemency {somewhere} since nineteen sixty-seven, but not in California.
Intpn. 5, prep: A condemned murderer has been granted clemency in California {at some point of time}, but not since nineteen sixty-seven.

But she was not alive when she got to the Lafayette area.

Negated statement: But she was not alive when she got to the Lafayette area.
Positive counterpart: But she was alive when she got to the Lafayette area.
Relevant tokens: She was alive when she got to the Lafayette area.

Potential positive interpretations:

Intpn. 1, root: She was {some adjective} when she got to the Lafayette area but not alive.
Intpn. 2, nsubj: (Somebody) was alive when she got to the Lafayette area but not she.
Intpn. 3, advcl: She was alive {at some point of time} but not when she got to the Lafayette area.

Table 1: Examples of negations and the steps to generate potential positive interpretations (the negated token is either a noun (top) or adjective (bottom)). We also indicate the dependency between a token in the potential focus and a token outside the potential focus.

4.1 Selecting Negations

Annotating all negations that modify a noun or adjective is outside the scope of this paper. To alleviate the annotation effort, we discard negations that belong to sentences that do not have at least one verb and one subject (nsubj or nsubjpass dependencies), and sentences that contain more than two negations, conditionals, questions or commas. Additionally, we skip negations if the negated token is the noun one, as scoring their potential positive interpretations is straightforward. Out of the 1,866 and 979 negation modifying nouns and adjectives, 635 and 320 pass the above filters. Out of these, we randomly select 309 and 75 respectively (approximately 50% and 25%).

4.2 Generating Potential Positive Interpretaions

We generate potential positive interpretations automatically using a deterministic procedure that manipulates part-of-speech tags and syntactic dependencies. The first step is to remove the negation cue to obtain the positive counterpart. Then, we use dependencies to select the tokens relevant to the negation—the main motivation is to select the eventuality to which the negation belongs. Finally, we generate potential interpretations using a battery of deterministic rules. Table 1 shows the output of each step with two sample negations.
Selecting Relevant Tokens. Negation may occur in sentences with multiple clauses. We simplify the original sentence and identify the eventuality to which the negation belongs by using the rules below. We defined these rules after analyzing several examples and the Stanford dependencies manual (de Marneffe and Manning, 2008).

When the negated token is a noun, we have two scenarios. Scenario (1) occurs when the negated token is the root or it has a cop dependency (copula) with the verb to be. In this case, we select all the dependents of the negated token. Scenario (2) occurs when neither of the two rules above apply. In that case, we select all the dependents of the closest verb, where the closest verb is the first verb found when traversing the dependency tree from the negated token to the root. For example, we simplify Article three says that no law shall prohibit any religious belief to No law shall prohibit any religious belief.

When the negated token is an adjective, we select all the dependents of the negated token. For example, from His estimate of 3.3% for third-quarter GNP is higher than the consensus because he believes current inventories aren’t as low as official figures indicate, we select Current inventories aren’t as low as official figures indicate.

Manipulating Dependencies to Generate Potential Positive Interpretations. After removing the negation cue and selecting relevant tokens, we use syntactic dependencies to select potential foci. Once potential foci are identified, generating positive interpretations is straightforward: each focus yields one interpretation, where everything but the focus is positive. The main idea is to select as potential foci subtrees rooted at selected tokens.

When the negated token is a noun, we select as potential foci the subtrees rooted in all the direct dependents of the negated token in Scenario (1), and the subtrees rooted in all the direct dependents of the closest verb in Scenario (2). When the negated token is an adjective, we select as potential foci the subtrees rooted in all the direct dependents of the negated token. This strategy to select potential foci has a few exceptions to avoid foci that yield meaningless interpretations. Specifically, we discard potential foci:

- whose root has dependency aux, auxpass, cop, poss, dep, prt or punct;
- that consist of
  - the determiner the, a, an, it and there;
  - the adverbs so, too, though, even, still, as, quite, either, however, anymore, moreover, therefore, furthermore, hence, thus, further, apparently, clearly, specifically, actually, fortunately, and unfortunately;
  - a single token with part-of-speech tag TO, CC, UH, POS or IN.

These exceptions were defined after manual examination of several examples. For example, consider sentence It’s not just women and girls who are affected. We avoid generating interpretations It’s just women and girls who {X} affected (focus would be are, with dependency aux), and It’s just women {X} girls who are affected (focus would be and, with part-of-speech tag CC).

After potential foci are selected, we generate potential positive interpretations by rewriting each focus with “someone / some people / something / etc.”, and appending “but not text of focus” at the end. Table 1 details the steps to generate potential positive interpretations.

4.3 Scoring Potential Interpretations

Once potential interpretations are generated automatically, we manually annotate them. The annotation interface shows the sentence containing the negation, the previous and next sentences as context, and one potential positive interpretation at a time. Annotators are asked Given the text snippet below [previous sentence, sentence containing the negation and next sentence], do you think the statement [positive interpretation] is true?, and must answer with a score ranging from 0 to 5, where 0 means certainly no and 5 means certainly yes. During pilot annotations, we found that certainty must be taken into account as forcing annotators to answer yes or no proved too restrictive. Note that some negations do not have any positive interpretation scored high, e.g., all interpretations from Utter no words receive a low score.

5 Corpus Analysis

Table 2 shows basic counts and statistics of the annotated potential positive interpretations. Dependency indicates the syntactic dependency between a token within the potential focus and a token outside the potential focus. The total number of potential positive interpretations is 777 when...
Table 3: Annotation examples. We show the original sentence containing a negation (the negated token is either a noun or an adjective), its context, and all potential interpretations with their scores.

<table>
<thead>
<tr>
<th>Dependency</th>
<th>#</th>
<th>%</th>
<th>Score Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>nsubj</td>
<td>231</td>
<td>30.0</td>
<td>4.45</td>
<td>0.88</td>
</tr>
<tr>
<td>root</td>
<td>169</td>
<td>21.7</td>
<td>3.96</td>
<td>1.41</td>
</tr>
<tr>
<td>pobj</td>
<td>70</td>
<td>9.0</td>
<td>3.99</td>
<td>1.31</td>
</tr>
<tr>
<td>amod</td>
<td>50</td>
<td>6.4</td>
<td>4.72</td>
<td>0.8</td>
</tr>
<tr>
<td>ccomp</td>
<td>45</td>
<td>5.7</td>
<td>3.62</td>
<td>1.34</td>
</tr>
<tr>
<td>xcomp</td>
<td>24</td>
<td>11.0</td>
<td>4.91</td>
<td>0.29</td>
</tr>
<tr>
<td>total</td>
<td>777</td>
<td>100.0</td>
<td>3.86</td>
<td>1.38</td>
</tr>
</tbody>
</table>

Table 2: Corpus analysis. For each dependency, we show the total number and percentage of interpretations, mean score and standard deviation.

Table 3 presents two complete annotation examples when the negated token is a noun and adjective. We show all potential interpretations generated and the manually assigned scores.

We generate three potential positive interpretations from the first example, The victim in this case is not a young child, and two of them received high scores (4 and 5). When reading the statement in context, it is clear that the judge implied that Something (a younger human) would be a younger child (Interpretation 2), and that The victim of this case is an older (not young) child (Interpretation 3). Interpretation (1), The victim in this case is a young something, receives a low score (2 out of 5). One could argue that this interpretations should receive a higher score because The victim in this case is a young adult, we simply provide real annotations drawn from our corpus.

The procedure to generate potential positive interpretations also generates three interpretations from the second example, But she was alive when she got to the Lafayette area, and two of them receive the highest score (5 out of 5). Interpretation (1) encodes the intuitive meaning that She was dead when she got to the Lafayette area, and Interpretation (3) captures that She was alive before she got to the Lafayette area. Interpretation (2) receives a low score (2 out of 5), as there is no evidence suggesting which individuals were alive when they got to the Lafayette area.

5.1 Annotation Examples

The negated token is a noun, and 200 when it is an adjective. On average, we generate 2.5 potential interpretation per negation when the negated token is a noun, and 2.7 when it is an adjective (we selected 309 and 75 negations respectively).

When the negated token is a noun, scores are overall lower (3.86 vs. 4.40), and scores are higher when the dependency is either nsubj (nominal subject, 4.45) or amod (adjectival modifier, 4.72). Regardless of dependency, mean scores are always over 4 when the negated token is an adjective.

Annotation Quality. The procedure to generate potential interpretations (Section 4.2) was tuned iteratively until we achieved considerable annotation agreement in pilot annotations. To ensure quality, we calculated Pearson correlations with 20% of annotations, and stopped the refinement process when we achieved 0.76 Pearson correlation. Note that Pearson is better suited than agreement measures designed for categorical labels, as not all disagreements are equally bad, e.g., 4 vs. 1 is worse than 4 vs. 5.
### Table 4: Features used to assign scores to automatically generated potential interpretations.

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>neg_cue</td>
<td>word form of the negation cue</td>
</tr>
<tr>
<td></td>
<td>neg_token</td>
<td>word form and part-of-speech tag of negated token</td>
</tr>
<tr>
<td>Path</td>
<td>syn_path_dep</td>
<td>path of dependencies from focus to negated token (or verb)</td>
</tr>
<tr>
<td></td>
<td>syn_path_pos</td>
<td>path of POS tags from focus to negated token (or verb)</td>
</tr>
<tr>
<td></td>
<td>last_syn_path_dep</td>
<td>last syntactic dependency in syn_path_dep</td>
</tr>
<tr>
<td></td>
<td>last_syn_path_pos</td>
<td>last part-of-speech tag in syn_path_pos</td>
</tr>
<tr>
<td>Focus</td>
<td>focus_length</td>
<td>number of tokens in potential focus</td>
</tr>
<tr>
<td></td>
<td>focus_first_word</td>
<td>word form and part-of-speech tag of first word in focus</td>
</tr>
<tr>
<td></td>
<td>focus_last_word</td>
<td>word form and part-of-speech tag of last word in focus</td>
</tr>
<tr>
<td></td>
<td>focus_direction</td>
<td>flag indicating whether focus occurs before or after neg_token</td>
</tr>
<tr>
<td></td>
<td>focus_head_word</td>
<td>word form of the head of focus</td>
</tr>
<tr>
<td></td>
<td>focus_head_pos</td>
<td>part-of-speech tag of the head of focus</td>
</tr>
<tr>
<td></td>
<td>focus_head_rel</td>
<td>syntactic dependency of the head of focus</td>
</tr>
</tbody>
</table>

### Table 5: Pearson correlations obtained with the test split and several combination of features. We detail results by the part-of-speech tag of the negated token (noun or adjective).

<table>
<thead>
<tr>
<th>Feature set</th>
<th>Gold Nouns</th>
<th>Adjectives</th>
<th>All</th>
<th>Predicted Nouns</th>
<th>Adjectives</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>neg_cue</td>
<td>-0.10</td>
<td>-0.24</td>
<td>0.06</td>
<td>0.02</td>
<td>-0.26</td>
<td>-0.02</td>
</tr>
<tr>
<td>basic</td>
<td>0.12</td>
<td>0.10</td>
<td>0.11</td>
<td>0.05</td>
<td>-0.39</td>
<td>0.01</td>
</tr>
<tr>
<td>basic + path</td>
<td>0.36</td>
<td>0.39</td>
<td>0.40</td>
<td>0.17</td>
<td>0.58</td>
<td>0.24</td>
</tr>
<tr>
<td>basic + path + focus</td>
<td>0.36</td>
<td>0.32</td>
<td>0.42</td>
<td>0.33</td>
<td>0.39</td>
<td>0.34</td>
</tr>
</tbody>
</table>

### 6 Learning to Score Potential Interpretations

We solve the task of scoring potential positive interpretations using standard supervised machine learning. We divide negations and their corresponding interpretations into training (80%) and test (20%), and use SVM with RBF kernel as implemented in scikit-learn (Pedregosa et al., 2011). We tune parameters $C$ and $\gamma$ using 10-fold cross-validation using the training set.

#### 6.1 Feature Selection

Table 4 lists the full feature set. We extract features from the negated token (nouns or adjectives), part-of-speech tags and dependency tree.

**Basic** features are straightforward. They include the negation cue, and the word form and part-of-speech tag of the negated token.

**Path** features are derived from the syntactic path between the subgraph selected as focus and the negated token or closest verb. If the negated token is a noun, we extract the path between the subgraph and the negated token in Scenario (1), and between the subgraph and the closest verb in Scenario (2) (Section 4.2). If the negated token is an adjective, we extract the path between the subgraph and the negated token. We include two paths (dependencies and part-of-speech tags), and the last dependency and part-of-speech tag.

**Focus** features characterize the dependency subgraph chosen as focus to generate the potential interpretation. 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 negated token. Additionally, we extract the word form, part-of-speech tag and dependency of the head of the focus, which we define as the token whose syntactic head is outside the focus.

### 7 Experimental Results

We perform two kinds of experiments. First, we score all potential positive interpretations automatically generated (Section 7.1). Second, we identify interpretations scored with the highest score, 5 out 5 (Section 7.2). We always build separate models for nouns and adjectives, and train with gold linguistic information (POS tags and dependencies). We report results on the test set using both gold and predicted linguistic information. For gold, we use the annotations provided with the CoNLL-2011 release, and for auto, we use the output of SyntaxNet (Andor et al., 2016).

#### 7.1 Scoring all Potential Interpretations

We score all potential interpretations using SVM for regression, and calculate Pearson correlation for evaluation purposes. Table 5 shows results obtained with several combinations of features.

When extracting features form gold linguistic...
Table 6: Precision, Recall and F-measure obtained with the test split for instances with the highest score (5 out of 5). Predicting these interpretations correctly allows our models to identify which of the automatically generated potential interpretations are valid given the negation.

<table>
<thead>
<tr>
<th></th>
<th>Nouns</th>
<th>Adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
</tr>
<tr>
<td>majority baseline</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>neg_mark</td>
<td>0.66</td>
<td>0.33</td>
</tr>
<tr>
<td>basic</td>
<td>0.66</td>
<td>0.37</td>
</tr>
<tr>
<td>basic + path</td>
<td>0.78</td>
<td>0.64</td>
</tr>
<tr>
<td>basic + path + focus</td>
<td>0.71</td>
<td>0.63</td>
</tr>
</tbody>
</table>

information (part-of-speech tags and syntactic dependencies), using only the negation cue as feature or basic features (negation cue and negated token) is rather useless (Pearson correlations range from -0.24 to -0.10). Using basic + path features yields 0.36 and 0.59 Pearson correlations for nouns and adjectives respectively, and including focus features is detrimental (0.36 and 0.52).

When extracting features from predicted linguistic information, we observe a similar trend in Pearson correlations. Results using predicted linguistic information are not directly comparable with those using gold linguistic information. Our methodology to generate potential interpretations relies heavily on syntactic dependencies, and using predicted interpretations implies that some interpretations present in our corpus cannot be automatically generated because of mistakes made by the parser. Out of the 196 potential interpretations in the original test set (20% of 977 annotated interpretations), we evaluate with the 94 interpretation generated with predicted dependencies (48%).

7.2 Identifying Valid Potential Interpretations

While scoring all potential positive interpretation generated is interesting, determining which of those interpretations are certain (scored 5 out of 5) is arguably more useful in a real system. Indeed, an inference tool would ideally extract certain interpretations from negation, and identify other potential interpretations as a byproduct of our generate-and-rank approach.

To estimate performance under this scenario, we approach the task as a standard binary classification task. Interpretations scored 5 receive the positive label, and other interpretations receive the negative label. Table 6 presents results in the test set (Precision, Recall and F-score) for the positive label using the majority baseline and several combinations of features, and gold and predicted linguistic information.

The majority baseline fails to detect any interpretation scored with 5 when the negated token is a noun (in this case, the majority of interpretations are not scored 5), and obtains a modest 0.64 F-score when the negated token is an adjective. Using gold linguistic information, neg_mark as the only feature or basic features improve performance when the negated token is a noun (F-score: 0.44 and 0.47), but does not improve results when the negated token is an adjective. Adding path features brings performance up (nouns: 0.70, adjectives: 0.73), and adding focus features yields similar results (nouns: 0.67, adjectives: 0.76).

Using predicted linguistic information, we observe the same general trends, but adding focus features brings a substantial improvement over basic + path for adjectives: 0.54 vs. 0.62.

8 Conclusions

This paper presents a framework to extract positive meaning from negation when the negation cue modifies a noun or adjective. First, we generate potential positive interpretations deterministically. Second, we rank them according to their likelihood. On average, we generate 2.5 potential interpretations when the negated token is a noun, and 2.7 when the negated token is an adjective.

Experimental results show that scoring all potential positive interpretations is challenging, we obtain overall Pearson correlation of 0.42 using features extracted from gold linguistic information (0.36 for nouns and 0.52 for adjectives). But when identifying interpretations annotated with the highest score (5 out of 5), F-scores are relatively high: 0.67 for nouns and 0.76 for adjectives. The latter evaluation is more suitable, as the ultimate goal is to identify valid positive interpretations and discard other potential interpretations generated with our generate-and-rank approach.
References


