Identifying the Focus of Negation Using Discourse Structure

Zahra Sarabi1, Eduardo Blanco1
1University of North Texas
1155 Union Circle, Denton, Texas 76203
zahrasarabi@my.unt.edu, eduardo.blanco@unt.edu

Abstract
This paper presents experimental results showing that discourse structure is a useful element in identifying the focus of negation. We define features extracted from RST-like discourse trees. We experiment with the largest publicly available corpus and an off-the-shelf discourse parser. Results show that discourse structure is especially beneficial when predicting the focus of negation in long sentences.

Introduction
Negation “relates an expression e to another expression with a meaning that is in some way opposed to the meaning of e” (Horn and Wansing 2017). It is present in all human languages, and it often conveys positive meanings within a discourse. For example, The seller didn’t ship the right parts implicitly conveys that the seller shipped the wrong parts.

Negation is understood in terms of scope and focus in theoretical grammars (Huddleston and Pullum 2002) and computational linguistics (Morante and Blanco 2012). 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 focus of negation is particularly interesting because it reveals positive meanings: everything but the focus is positive. Following with the example above, The seller did n’t ship the right parts, the focus is right. Knowing the focus reveals the positive meaning The seller shipped parts. (But not the right parts). After identifying the focus, one could rewrite right with wrong to come up with the positive counterpart of the negation: The seller shipped the wrong parts.

Identifying the focus of negation is a challenging task. Previous works obtain F1 measures ranging from 0.58 to 0.66, and while some incorporate inter-sentential information, none of them exploit discourse structure. Identifying the focus of negation, however, intuitively benefits from discourse structure. Going beyond sentences is important since they do not stand on their own in a coherent text. Discourse units (sentences or clauses) are logically connected:

the meaning of a unit relates to the previous and next units. Consider the following text: Analysts continued to speculate that Ford may try to force the issue by calling for a special shareholder’s meeting. But a Ford spokeswoman said Friday the company hasn’t requested such a meeting yet. The contrast discourse relation between these sentences signals that the focus of the negation is requested, revealing that Ford has avoided such a meeting.

In this paper, we incorporate discourse structure into the task of identifying the focus of negation. Specifically, we work with the Rhetorical Structure Theory (RST) framework (Mann and Thompson 1988). The main contributions are: (a) features extracted from RST-like trees, and (b) experiments quantifying the benefits of these features.

Background and Previous Work
Corpus. We work with PB-FOC (Morante and Blanco 2012), the largest corpus with focus of negation annotations. It contains 3,548 verbal negations from PropBank (Palmer, Gildea, and Kingsbury 2005). We chose PB-FOC over other more recent corpora (e.g., (Sarabi and Blanco 2016)) because PB-FOC: (a) contains an order of magnitude more negations (≈3,500 vs. <250), and (b) does not impose any restriction in sentence length. As we shall see, discourse structure is especially useful with long sentences.

PB-FOC was created by (a) selecting verbal negations marked with argm-neg role in PropBank and (b) manually annotating as focus of each verbal negation the most likely semantic role. For example in [The American Bankers Association]ARG0 did [n’t]ARGM-NEG [have]Verb [any comment on the plan]ARG1, the ARG0 semantic role is selected as the focus of negation. Table 2 provides basic counts. Note that most negations have ARG0 and ARG1 roles (2,393 and 3,207 out of 3,548), and that the probability of a role being the focus varies greatly, e.g., ARG0: 6.1%, ARGM-TMP: 42.0%, ARGM-MNR: 73.0%.

Rhetorical Structure Theory (RST). Discourse is a “descriptive linguistic approach to a range of phenomena in the organization of discourse,”(Mann and Thompson 1988) and was originally developed to guide computational text generation (Taboada and Mann 2006). RST captures discourse structure by first dividing text into elemental discourse units,
Table 1: Features used to predict the focus of negation. RST-based features are novel features extracted from the discourse structure obtained with an RST discourse parser.

<table>
<thead>
<tr>
<th>Role</th>
<th># negs.</th>
<th>Focus</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARG_0</td>
<td>2,393</td>
<td>145</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>ARG_1</td>
<td>3,207</td>
<td>1,546</td>
<td>48.2</td>
<td></td>
</tr>
<tr>
<td>ARG_2</td>
<td>519</td>
<td>196</td>
<td>37.8</td>
<td></td>
</tr>
<tr>
<td>ARG_3</td>
<td>53</td>
<td>14</td>
<td>26.4</td>
<td></td>
</tr>
<tr>
<td>ARG_4</td>
<td>28</td>
<td>18</td>
<td>64.3</td>
<td></td>
</tr>
<tr>
<td>ARGM-ADV</td>
<td>516</td>
<td>117</td>
<td>22.7</td>
<td></td>
</tr>
<tr>
<td>ARGM-CAU</td>
<td>98</td>
<td>4</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>ARGM-DIR</td>
<td>38</td>
<td>9</td>
<td>23.7</td>
<td></td>
</tr>
<tr>
<td>ARGM-DIS</td>
<td>291</td>
<td>10</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>ARGM-EXT</td>
<td>19</td>
<td>16</td>
<td>84.2</td>
<td></td>
</tr>
<tr>
<td>ARGM-LOC</td>
<td>131</td>
<td>36</td>
<td>27.5</td>
<td></td>
</tr>
<tr>
<td>ARGM-MNR</td>
<td>267</td>
<td>195</td>
<td>73.0</td>
<td></td>
</tr>
<tr>
<td>ARGM-NEG</td>
<td>3,548</td>
<td>925</td>
<td>26.1</td>
<td></td>
</tr>
<tr>
<td>ARGM-PNC</td>
<td>85</td>
<td>53</td>
<td>62.4</td>
<td></td>
</tr>
<tr>
<td>ARGM-TMP</td>
<td>605</td>
<td>254</td>
<td>42.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Number of verbal negations that have each semantic role, and number of times each role is the focus of the negation. The total number of negations is 3,548.

and then adding discourse relations (CONCESSION, ELABORATION, etc.) to build a discourse tree. Arguments of discourse relations belong to two types: nuclei (the most important part) or satellites (subordinate, secondary part to nuclei). Relations such as CONCESSION have one nucleus and a satellite, other relations such as CONTRAST are multinuclear. Figure 1 draws a discourse tree, we refer the reader to the aforementioned works for more details. We obtain RST-like trees using the discourse parser by (Surdeanu, Hicks, and Valenzuela-Escarcega 2015), which is trained using the corpus by (Carlson, Marcu, and Okurouki 2001).

**Previous Work.** Scope has received considerably more attention than focus in computational linguistics (Özgür and Radev 2009; Li et al. 2010; Reitan et al. 2015). The state of the art obtains 0.89 F1 (scope tokens) and 0.78 F1 (exact match) (Fancellu, Lopez, and Webber 2016).

Existing focus detectors are less sophisticated than scope detectors. (Rosenberg and Bergler 2012) present a set of heuristics grounded on syntax, and (Blanco and Moldovan 2011) use features extracted from the verb-argument structure of the negation. (Zou, Zhou, and Zhu 2014) are the only ones working with inter-sentential features using a graph model. The work presented here is the first to incorporate discourse structure for identifying focus of negation and outperforms previous work.

Finally, we note that discourse structure has been proven useful for other tasks including question answering (Jansen, Surdeanu, and Clark 2014) and machine translation (Guzmán et al. 2014).

**Experiments and Results**

We experiment with an SVM trained with features proposed in previous work and features derived from discourse structure. Each verbal negation becomes an instance, and the focus prediction task is reduced to predicting which semantic role is the focus. We use the implementation provided in scikit-learn (Pedregosa et al. 2011), and tune hyperparameters C and γ using 10-fold cross-validation with the train and development splits in PB-FOC. We report results obtained with the test split in PB-FOC after training a model with the train and development splits and the best hyperparameters found during the tuning process. For each negation, we obtain the RST-like discourse tree automatically using a window size of 11 sentences (5 before and 5 after).

**Feature Set**

The feature set is described in Table 1. Features proposed by previous work mainly characterize the verb-argument structure to which the negation belongs. Features 1–80 characterize all semantic roles of the negated verb (5 features × 16 semantic roles). Features 81–83 further characterize ARG_1, the most likely role to be the focus (Table 2). Features 84–85 indicate the first and last semantic role, as focus is often the first or last role. Finally, features 86–91 characterize the verb based on word forms and part-of-speech tags.

The bottom block of Table 1 describes the features extracted from the RST-like discourse tree. We experiment

---

1We discarded the features corresponding to the thematic roles of numbered semantic roles because they did not bring performance improvements during the tuning process.
It’s not a crop [ARG_1] that can’t be [doubled] [verb] [ARG_2] [by Mr. Krishnamurthy].

But no one has made a serious effort [ARG_1] to transplant the crop. [ARG_2] In Sidhpur, it is almost time [ARG_1] to sow this year’s crop. [ARG_2]

RST-based features:
- Features 92–108: elaboration (ARG_1 and verb)
- Features 109–125: nucleus (ARG_1) and satellite (verb)
- Features 126–142: elaboration + attribution + contrast (ARG_1 and verb)
- Features 143–159: nucleus + nucleus + none (ARG_1) and satellite + nucleus + none (verb)
- Features 160–175: false (ARG_1)
- Feature 177: false

Figure 1: Sample text, automatically obtained discourse tree (right) and values for RST-based features (left).

**Comparison with previous work.** Our models outperform previous approaches to identify the focus of negation. Overall, we obtained 0.69 F1, while Rosenberg and Bergler 2012 obtained 0.58 F1 and Zou, Zhou, and Zhu 2014 obtained 0.66 F1. Beyond results, we note that our approach is not much more complex than previous work. The key difference is that we incorporate discourse structure automatically obtained with an off-the-shelf discourse parser.

**Conclusions**

This paper shows that discourse structure helps when identifying the focus of negation. We have defined features extracted from an RST-like discourse tree. Using an off-the-shelf discourse parser, these features improve focus identification with long sentences (>20 tokens, F1: 0.65 vs. 0.69), and are not detrimental with short sentences (<20 tokens). The model presented here obtains the best results to date on PB-FOC, the largest publicly available corpus with annotations on focus of negation. Our future plans include investigating other discourse frameworks such as the Penn Discourse Treebank (Prasad et al. 2008), and exploring sources of knowledge from the discourse structure beyond the actual discourse relation labels and argument types.

**Acknowledgments**

This material is based upon work supported by the National Science Foundation under Grants Nos. 1734730, 1832267 and 1845757. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

**References**


Table 3: Results obtained in the test set without and with features extracted from discourse structure. Long sentences are those with 20 or more tokens. We only provide detailed results for roles that are the focus at least five times in the test set, but the weighted average includes all of them. We indicate statistical significance between the Weighted Avg. F1 measures obtained with Prev. Work and Prev. Work + RST with * (Z-score, p < 0.10).


Palmer, M.; Gildea, D.; and Kingsbury, P. 2005. The Propo-


