The Thirty-Second International Florida
Artificial Intelligence Research Society Conference (FLAIRS-32)

Identifying the Focus of Negation Using Discourse Structure

Zahra Sarabi,[1] **Eduardo Blanco**[1]
[1]University 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 ($\approx 3,500$ 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]*$_{\text{ARG}_0}$ *did [n't]*$_{\textit{ARGM-NEG}}$ *[have]*$_{\text{Verb}}$ *[any comment on the plan]*$_{\text{ARG}_1}$, the ARG_0 semantic role is selected as the focus of negation. Table 2 provides basic counts. Note that most negations have ARG_0 and ARG_1 roles (2,393 and 3,207 out of 3,548), and that the probability of a role being the focus varies greatly, e.g., ARG_0: 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,

	#	Description
Prev. Work	1–16	binary flags indicating semantic role presence (ARG_0, ARG_1, etc.)
	17–80	first word and POS tag, number of tokens, and the position within all roles for each role
	81–83	syntactic node of ARG_1 and flags indicating if it contains selected POS tags or keywords
	84–85	labels of the first and last semantic roles
	86–87	the main verb and the sequence of words in the VP containing the verb
	88–89	the POS tag of the main verb and the sequence of POS tags in the VP containing the verb
	90–91	flags indicating whether the VP containing the verb has a CC and an RB
RST-based	92–108	discourse relation between (a) each role and the verb and (b) the rest of the discourse tree
	109–125	whether each role and the verb are satellite or nucleus in the relations from features 92–108
	126–142	discourse relations in the path from each role and the verb to the root of the discourse tree
	143–159	whether each discourse unit in features 126–142 is satellite or nucleus
	160–175	flags indicating whether each role belongs to the same elemental discourse unit than the verb
	177	flag indicating whether the sentence containing the verb is one elemental discourse unit

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.

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

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 aforecited 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 Okurowski 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.[1] Features 1–80 characterize all semantic roles of the negated verb (5 features $\times$ 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

[1] We discarded the features corresponding to the thematic roles of numbered semantic roles because they did not bring performance improvements during the tuning process.

Text: [It's not [a crop]$_{\text{ARG}_1}$]$_{1A}$ [that ca*n't* be [doubled]$_{verb}$,]$_{1B}$ [says Mr. Krishnamurthy.]$_{1C}$ [But no one has made a serious effort]$_{1D}$ [to transplant the crop.]$_{1E}$ [In Sidhpur, it is almost time]$_{1F}$ [to sow this year's crop.]$_{1G}$

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*

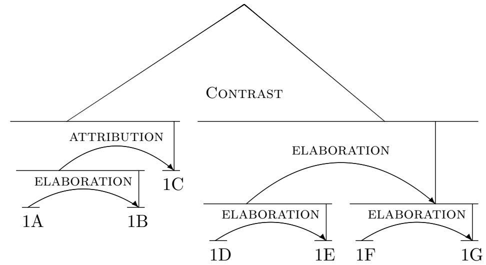

Figure 1: Sample text, automatically obtained discourse tree (right) and values for RST-based features (left).

with four features extracted from each semantic role and the verb (features 92–159, $4 \times 17 = 68$), flags indicating whether each role belongs to the same elemental discourse unit than the verb (features 160–175), and a flag indicating whether one elemental discourse unit contains the whole sentence to which the negation belongs. More specifically, features 92–159 are extracted as follows. First, we retrieve the discourse unit that contains each semantic role and verb. This discourse unit may be a leaf in the discourse tree (i.e., an elemental discourse unit) or an internal node. Then, for each discourse unit retrieved, we extract: the outgoing discourse relation (features 92–108), whether the outgoing relation is nucleus or satellite (features 109–125), the sequence of discourse relations in the path to the root of the discourse tree (features 126–142), and the sequence of argument types (i.e., nucleus or satellite) in the path to the root (features 143–159). Figure 1 exemplifies the RST-based features.

Results and Analysis

Table 3 presents results obtained with models trained with the features from (Blanco and Moldovan 2011) and incorporating the RST-based features proposed in this paper. Overall, the benefit of discourse-based features is minimal (F1: 0.68 vs. 0.69). The benefit is, however, substantial for long sentences, which we define as sentences consisting of 20 or more tokens. Indeed, for long sentences, the overall difference in F1 (0.65 vs. 0.69) is significant with a 0.10 p-value (Z-score). This is not surprising, as shorter sentences tend to belong to a single elemental discourse unit and thus most RST-features have the same values.

When working with long sentences, RST-based features are always beneficial regardless of which role is the focus (there are two exceptions: ARGM-ADV and ARGM-LOC). ARG$_1$ and ARGM-NEG, the most frequent roles to be the focus (214 and 126 out of 497 instances respectively), benefit: F1 improves from 0.77 to 0.80 with ARG$_1$, and from 0.53 to 0.55 with ARGM-NEG. Other roles benefit more in both absolute and relative terms, but because of their lower frequencies have less overall impact. More specifically, the F1 for ARGM-PNC improves from 0.73 to 0.83 (+13.7%), for ARG$_0$ from 0.19 to 0.28 (+47.4%), and for ARGM-EXT from 0.50 to 0.91 (+82.0%).

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 ($\leq$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

Blanco, E., and Moldovan, D. 2011. Semantic representation of negation using focus detection. In *Proceedings of ACL*, 581–589.

Carlson, L.; Marcu, D.; and Okurowski, M. E. 2001. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of SIGDIAL '01*, SIGDIAL '01, 1–10.

		#	Prev. Work			Prev. Work + RST		
			P	R	F1	P	R	F1
ARG_0	All sentences	29	0.53	0.28	0.36	0.53	0.28	0.36
	Long sentences	19	0.23	0.16	0.19	0.40	0.21	0.28
ARG_1	All sentences	308	0.76	0.85	0.80	0.74	0.85	0.79
	Long sentences	214	0.73	0.81	0.77	0.76	0.85	0.80
ARG_2	All sentences	35	0.54	0.77	0.64	0.56	0.71	0.63
	Long sentences	31	0.60	0.68	0.64	0.61	0.71	0.66
ARGM-ADV	All sentences	21	0.58	0.52	0.55	0.59	0.48	0.53
	Long sentences	19	0.64	0.47	0.55	0.56	0.47	0.51
ARGM-EXT	All sentences	6	0.75	0.50	0.60	0.50	0.17	0.25
	Long sentences	6	1.00	0.33	0.50	1.00	0.83	0.91
ARGM-LOC	All sentences	8	0.22	0.25	0.24	0.30	0.38	0.33
	Long sentences	7	0.33	0.43	0.38	0.25	0.29	0.27
ARGM-MNR	All sentences	41	0.78	0.88	0.83	0.78	0.88	0.83
	Long sentences	30	0.77	0.77	0.77	0.82	0.77	0.79
ARGM-NEG	All sentences	182	0.70	0.46	0.56	0.68	0.48	0.56
	Long sentences	126	0.62	0.47	0.53	0.66	0.47	0.55
ARGM-PNC	All sentences	12	0.86	1.00	0.92	0.86	1.00	0.92
	Long sentences	5	0.67	0.80	0.73	00.71	1.00	0.83
Weighted Avg.	All sentences	704	0.69	0.68	0.68	0.70	0.70	0.69
	Long sentences	497	0.65	0.66	0.65	0.69	0.70	*0.69

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, $p < 0.10$).

Fancellu, F.; Lopez, A.; and Webber, B. 2016. Neural networks for negation scope detection. In *Proceedings of ACL*, 495–504.

Guzmán, F.; Joty, S.; Màrquez, L.; and Nakov, P. 2014. Using discourse structure improves machine translation evaluation. In *Proceedings ACL*, 687–698.

Horn, L. R., and Wansing, H. 2017. Negation. In *The Stanford Encyclopedia of Philosophy*.

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

Jansen, P.; Surdeanu, M.; and Clark, P. 2014. Discourse complements lexical semantics for non-factoid answer reranking. In *Proceedings of ACL*, 977–986.

Li, J.; Zhou, G.; Wang, H.; and Zhu, Q. 2010. Learning the Scope of Negation via Shallow Semantic Parsing. In *Proceedings of ACL (Coling 2010)*, 671–679.

Mann, W. C., and Thompson, S. A. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text-Interdisciplinary Journal for the Study of Discourse* 8(3):243–281.

Morante, R., and Blanco, E. 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)*, 265–274.

Özgür, A., and Radev, D. R. 2009. Detecting speculations and their scopes in scientific text. In *Proceedings EMNLP*, EMNLP '09, 1398–1407.

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

Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; and Duchesnay, E. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12:2825–2830.

Prasad, R.; Dinesh, N.; Lee, A.; Miltsakaki, E.; Robaldo, L.; Joshi, A.; and Webber, B. 2008. The Penn Discourse TreeBank 2.0. In *Proceedings of LREC'08*.

Reitan, J.; Faret, J.; Gambäck, B.; and Bungum, L. 2015. Negation scope detection for twitter sentiment analysis. In *Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 99–108.

Rosenberg, S., and Bergler, S. 2012. Uconcordia: Clac negation focus detection at *sem 2012. In *(SemEval 2012)*, 294–300.

Sarabi, Z., and Blanco, E. 2016. Understanding negation in positive terms using syntactic dependencies. In *Proceedings of EMNLP*, 1108–1118.

Surdeanu, M.; Hicks, T.; and Valenzuela-Escarcega, M. A. 2015. Two practical rhetorical structure theory parsers. In *Proceedings of NAACL*, 1–5.

Taboada, M., and Mann, W. C. 2006. Rhetorical structure theory: looking back and moving ahead. *Discourse Studies* 8(3):423–459.

Zou, B.; Zhou, G.; and Zhu, Q. 2014. Negation focus identification with contextual discourse information. In *Proceedings of ACL*, 522–530.