LETTER

Re-Evaluating Syntax-Based Negation Scope Resolution*

Asahi YOSHIDA[†a)], *Nonmember*, Yoshihide KATO[††], *Member*, *and* Shigeki MATSUBARA[†,††], *Senior Member*

SUMMARY Negation scope resolution is the process of detecting the negated part of a sentence. Unlike the syntax-based approach employed in previous researches, state-of-the-art methods performed better without the explicit use of syntactic structure. This work revisits the syntax-based approach and re-evaluates the effectiveness of syntactic structure in negation scope resolution. We replace the parser utilized in the prior works with state-of-the-art parsers and modify the syntax-based heuristic rules. The experimental results demonstrate that the simple modifications enhance the performance of the prior syntax-based method to the same level as state-of-the-art end-to-end neural-based methods.
key words: *negation, negation detection, negation scope resolution, syntactic parser*

1. Introduction

Negation is a common linguistic phenomenon that frequently appears in natural language. Consequently, its detection is crucial for various NLP applications, including sentiment analysis, relation extraction and medical data mining. Typically, the negation detection task is broken down into two subtasks: (i) detecting negation cues (words, affixes, or phrases that express negations) and (ii) resolving their scopes (parts of a sentence affected by the negation cue). In example (1) below, the word "not" is the negation cue (marked in bold) and word sequences "He did" and "go to school" form the scope (underlined parts).

(1) He did **not** go to school and stayed home.

This work addresses the second subtask: negation scope resolution. Prior works used syntactic features for resolving the scope of negations [1]–[4]. Read et al. [1] tackled this issue with syntax-based approach and obtained the best performance on the token-level evaluation in *SEM2012 shared task [5]. Recently, many studies treat this task as a sequence labeling problem and use deep-learning techniques [6]–[8]. Without explicitly utilizing syntactic structure, they argued that end-to-end neural approaches can outperform earlier syntax-based ones. However, the prior works proposed in *SEM2012 shared task used the parser of that time**. The performances of parsers have considerably improved since. The effectiveness of the syntax-based approach will increase with the usage of accurate parsers. Furthermore, syntax-based methods have an advantage over deep-learning techniques: high interpretability.

Motivated by the point mentioned above, this work revisits the syntax-based approach for negation scope resolution. We use state-of-the-art parsers to re-evaluate the earlier syntax-based approach. We also modify the syntactic-based heuristic rules used in the prior syntax-based method. Our experimental results demonstrate that the prior method, based on heuristics for syntax structure, can obtain the same level of performance as state-of-the-art methods based on end-to-end neural networks.

2. Related Work

This section describes the syntax-based method proposed by Read et al. [1], based on which we re-evaluate the usefulness of syntax for negation scope resolution. Their approach assumes that the scope of negation corresponds to a constituent. As an example, let us consider the sentence (2).

(2) I know that he is **not** a student.

Figure 1 shows the constituent parse tree of the sentence. In this sentence, the scope of the negation cue "not" corresponds to the constituent S whose left end is "he" and

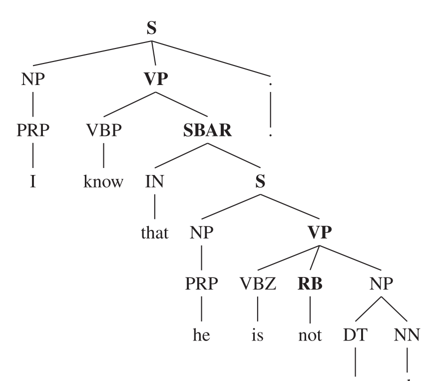

Fig. 1 Constituent parse tree of sentence (2), highlighting candidate scope constituents.

Manuscript received August 6, 2023.
Manuscript publicized October 16, 2023.
[†]The authors are with Graduate School of Informatics, Nagoya University, Nagoya-shi, 464–8601 Japan.
[††]The authors are with Information and Communications, Nagoya University, Nagoya-shi, 464–8601 Japan.
*This paper was presented at the 12th Joint Conference on Lexical and Computational Semantics (*SEM 2023).
a) E-mail: yoshida.asahi.y7@s.mail.nagoya-u.ac.jp
DOI: 10.1587/transinf.2023EDL8050

**The syntactic information provided by the parser is annotated on the datasets utilized in *SEM2012 shared task. Participants in the shared task applied this syntactic information.

whose right end is “student”. This method resolves the scope of the negation cue according to the following steps:

1. Parse the sentence and select the constituents on the path from the cue to the root as candidates (The candidates are marked in bold in Fig. 1).
2. Select one constituent corresponding to the scope either using heuristics or the Support Vector Machine classifier.
3. Adjust the scope by removing certain elements from the constituent selected in the second step.

In the first step, the sentence is parsed and all the constituents that dominate the negation cue are considered as scope candidates. For example, in sentence (2), six constituents highlighted in Fig. 1 are selected as candidates. In the second step, one constituent is selected from the candidates using either heuristics or a classifier. We describe the heuristic method, which we use in this work. This method selects one constituent from the candidates using *scope resolution heuristics* shown in Fig. 2. The 14 rules that form the heuristics are applied in order from top to bottom; the rules are listed in a specific-to-general order. Each rule is represented as a path pattern and some rules have additional constraints (if part). For example, the fifth rule

```
RB//VP/SBAR if SBAR\WH*  (#)
RB//VP/S
RB//S
DT/NP if NP/PP
DT//SBAR if SBAR\WHADVP
DT//S
JJ//ADJP/VP/S if S\VP\VB* [@lemma="be"]
JJ/NP/NP if NP\PP
JJ//NP
UH
IN/PP
NN/NP//S/SBAR if SBAR\WHNP
NN/NP//S
CC/SINV
```

Fig. 2 Scope resolution heuristics. Each row displays one rule, which is presented in the order that they should be applied. Each rule is represented as a path pattern. A/B denotes that B is the parent of A, A//B implies B is an ancestor of A, and A\B means B is a child of A. (#) is the rule we modify in this work.

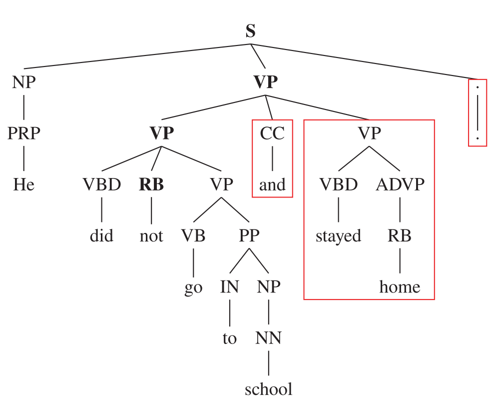

Fig. 3 Constituent parse tree of sentence (1), enclosing removed parts in boxes.

“DT//SBAR if SBAR\WHADVP” will be activated and the constituent SBAR is selected when the negation cue is a determiner (DT), provided that it has an ancestor SBAR if the SBAR has a child WHADVP. If no rule is activated, it uses a *default scope*, which expands the scope to the left and the right of the negation cue until either a sentence boundary or a punctuation is found.

The alignment of the constituent and the scope is not always straightforward. Sentence (1) is one of such illustration. In this sentence, the scope of the negation cue “not” does not cross the coordination boundary: the coordinating conjunction “and”, its following conjunct “stayed home” and the punctuation “.” are not included in the scope. To deal with such a case, Read et al. [1] adopted some heuristic rules to remove certain elements from the constituent. These rules remove elements that are not within the scope, such as constituent-initial interjections and sentential adverbs.

For sentence (1), the scope “He did, go to school” is correctly resolved using the series of process. The constituent S is selected as the scope of the cue according to the first and second steps. In the third step, the coordinating conjunction “and”, its conjunct “stayed home” and the punctuation “.” are removed by the heuristic rules (removed parts are enclosed in Fig. 3).

3. Revisiting the Syntax-Based Method

In this section, we revise the method described in the previous section to re-evaluate the syntax-based approach in negation scope resolution. Section 3.1 describes the parsers we use in this work. Sections 3.2 and 3.3 discuss the modifications we made for the second and the third steps of Read et al.’s method [1], respectively.

3.1 Replacement of the Parser

The dataset used in *SEM2012 shared task [9], also known as the Conan Doyle dataset, is one of the primary datasets used for negation scope resolution. This dataset also contains syntactic information, which was assigned using the reranking parser of Charniak and Johnson [10]. As Read et al. mentioned, syntactic information contains parse errors. They suspected that parse errors cause scope resolution errors in their method. To mitigate this issue, we parse the sentences in the dataset using state-of-the-art, high-accuracy parsers. We use two parsers: Berkeley Neural Parser [11], [12] with BERT [13], and Attach Juxtapose Parser [14] with XLNET [15]. Table 1 shows the performances of the parsers on Penn Treebank [16].

Table 1 Performances of the parsers in Penn Treebank Section 23.

Parser	F_1 score (%)
Reranking Parser [10]	91.02
Berkeley Neural Parser [11], [12]	95.77
Attach Juxtapose Parser [14]	96.34

Table 2 Scope resolution performances for gold cues using the three different parsers. The upper figure in each row demonstrates the result with modified rules discussed in Sects. 3.2 and 3.3; the lower figure shows the result without modifications. Note that in the case of the rule to remove sentential adverbs from the scope in the third step, we were not able to reproduce the Read et al.'s method because the sentential adverb list is not publicly available. Thus, both the upper and the lower figures describe the results of our modified rule.

Parser	Scope-level			Token-level		
	Pre. (%)	**Rec.** (%)	**F$_1$** (%)	**Pre.** (%)	**Rec.** (%)	**F$_1$** (%)
Reranking Parser	97.21 (97.14)	69.88 (68.27)	81.31 (80.19)	86.87 (85.48)	93.07 (93.63)	89.86 (89.37)
Berkeley Neural Parser	98.91 (98.88)	72.69 (70.68)	83.80 (82.43)	89.78 (87.88)	92.96 (93.57)	91.34 (90.64)
Attach Juxtapose Parser	98.94 (98.90)	74.70 (72.29)	**85.13** (83.53)	90.62 (88.70)	94.68 (95.24)	**92.61** (91.85)

3.2 Modification of *Scope Resolution Heuristics*

Read et al. used *scope resolution heuristics* shown in Fig. 2 to detect the constituent corresponding to the scope of the negation cue. The first rule of Read et al. (denoted with (#) in Fig. 2) is considered to extract relative clauses, but this rule does not work properly. In relative clauses in Penn Treebank, SBAR directly dominates not VP but S (and the S has a child VP). To accurately capture this structure, we modify the rule as follows:

(3) `RB//VP/S/SBAR if SBAR\WHNP`

This modification is based on the preliminary experiment conducted on the training data.

3.3 Modification of Scope Adjustment

As indicated in Sect. 2, Read et al.'s method adjusts the constituent in the third step. This work partially modifies their heuristics.

First, we present the following additional rule to the original ones:

- Remove initial PP (prepositional phrase) if delimited by a comma.

This modification was motivated by the annotation guideline of the Conan Doyle dataset [17]. According to this guideline, discourse markers are excluded from the scope. Comma-delimited prepositional phrases often function as discourse markers, such as "In my opinion" in example (4). In this case, we should remove them from the scope.

(4) In my opinion, <u>he should</u> **not** <u>go</u>.

Also, we modify one of the Read et al.'s rules: removing sentential adverbs from the scope. Read et al. compiled a list of sentential adverbs from the training data and used it for this processing. Instead, in this work, we simply remove "comma-delimited ADVP (adverbial phrase) or INTJ (interjection)" from the scope along with the commas. This is a generalization of Read et al.'s processing. As an example of a comma-delimited ADVP that functions as a discourse-level adverbial and should be excluded from the scope, see sentence (5) below.

(5) <u>There was</u> **no** <u>trace</u>, however, <u>of anything</u>.

Again, this modification of scope adjustment rules is based on the training data.

4. Experiment

4.1 Experimental Settings

To re-evaluate the syntax-based approach to negation scope resolution, we conducted an experiment† using the evaluation data of the Conan Doyle dataset, which was employed in *SEM2012 shared task. We created new constituent parse trees for the sentences in the dataset using Berkeley Neural Parser and Attach Juxtapose Parser. We conducted negation scope resolution using the modified version of Read et al. [1] discussed in Sects. 3.2 and 3.3. Other experimental setups are similar to those of *SEM2012 shared task, with the scope-level F_1 score and the token-level F_1 score as the evaluation metrics. We used the official script distributed in the shared task†† for evaluation.

4.2 Experimental Results

Table 2 shows the experimental results with three different parsers to provide the constituent parse trees. The results demonstrate that the use of accurate parsers leads to an increase in performance in negation scope resolution for both scope-level and token-level metrics. We also verified that the rule modifications introduced in this work contributed to the performance improvement.

Several previous works, including state-of-the-art methods, incorporate punctuation tokens for evaluation, which were omitted in *SEM2012 shared task. To compare our results with these methods, we also assessed F_1 score including punctuation tokens. Table 3 shows the results. The performance of the syntax-based method tested in this work obtained 91.74% in F_1 score including punctuations, which is only 1.11 points behind values reported by the state-of-the-

†The code is available at https://github.com/asahi-y/revisiting-nsr

††https://www.clips.ua.ac.be/sem2012-st-neg/data.html

Table 3 Comparison to previous methods. The results of this work are the ones obtained by using syntactic information generated by Attach Juxtapose Parser, and by applying modified rules. Note that the results are for negation scope resolution using gold cues.

Method	**Token-level F$_1$** (%) Including punctuations	Excluding punctuations
This work	91.74	92.61
Fancellu et al. (2016) [6]	88.72	-
Li and Lu (2018) [19]	-	89.4
Khandelwal and Sawant (2020) [7]	92.36	-
Truong et al. (2022) [8]	91.24	-
Wu and Sun (2023) [18]	92.85	-

art method (92.85%), obtained by Wu and Sun [18]. This result shows that the prior method based on heuristics for syntax, with the use of a high-performance parser, can obtain performance close to the results obtained by the best-performing deep learning methods.

5. Conclusion

This work re-evaluated the syntax-based approach in negation scope resolution. We replaced the parser used in the prior works with the state-of-the-art parsers. We also slightly modified the syntax-based heuristic rules designed in the prior work. The experimental results demonstrate that the prior syntax-based approach can obtain high performance comparable to those of state-of-the-art methods. This work gives a strong baseline for the negation scope resolution task and opens up the possibility of accurate and interpretable negation scope resolution.

In future work, we will introduce a tree-based neural model into the constituent selection process to enhance the performance of the scope prediction. It would also be interesting to apply the syntax-based approach to the scope resolution of other linguistic phenomena, for example, speculation or quantifier.

Acknowledgments

This research was partially supported by the Grant-in-Aid for Scientific Research (C) (No. 22K12148) of JSPS.

References

[1] J. Read, E. Velldal, L. Øvrelid, and S. Oepen, "UiO1: Constituent-based discriminative ranking for negation resolution," The 1st Joint Conference on Lexical and Computational Semantics, Montréal, Canada, pp.310–318, June 2012.

[2] J. Carrillo de Albornoz, L. Plaza, A. Díaz, and M. Ballesteros, "UCM-I: A rule-based syntactic approach for resolving the scope of negation," The 1st Joint Conference on Lexical and Computational Semantics, Montréal, Canada, pp.282–287, June 2012.

[3] A. Abu-Jbara and D. Radev, "UMichigan: A conditional random field model for resolving the scope of negation," The 1st Joint Conference on Lexical and Computational Semantics, Montréal, Canada, pp.328–334, June 2012.

[4] J.P. White, "UWashington: Negation resolution using machine learning methods," The 1st Joint Conference on Lexical and Computational Semantics, Montréal, Canada, pp.335–339, June 2012.

[5] R. Morante and E. Blanco, "*SEM 2012 shared task: Resolving the scope and focus of negation," The 1st Joint Conference on Lexical and Computational Semantics, Montréal, Canada, pp.265–274, June 2012.

[6] F. Fancellu, A. Lopez, and B. Webber, "Neural networks for negation scope detection," Proc. 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, pp.495–504, Aug. 2016.

[7] A. Khandelwal and S. Sawant, "NegBERT: A transfer learning approach for negation detection and scope resolution," Proc. 12th Language Resources and Evaluation Conference, Marseille, France, pp.5739–5748, May 2020.

[8] T. Truong, T. Baldwin, T. Cohn, and K. Verspoor, "Improving negation detection with negation-focused pre-training," Proc. 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Seattle, United States, pp.4188–4193, July 2022.

[9] R. Morante and W. Daelemans, "ConanDoyle-neg: Annotation of negation cues and their scope in conan doyle stories," Proc. 8th International Conference on Language Resources and Evaluation, Istanbul, Turkey, pp.1563–1568, May 2012.

[10] E. Charniak and M. Johnson, "Coarse-to-fine *n*-best parsing and MaxEnt discriminative reranking," Proc. 43rd Annual Meeting of the Association for Computational Linguistics, Ann Arbor, Michigan, pp.173–180, June 2005.

[11] N. Kitaev and D. Klein, "Constituency parsing with a self-attentive encoder," Proc. 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia, pp.2676–2686, July 2018.

[12] N. Kitaev, S. Cao, and D. Klein, "Multilingual constituency parsing with self-attention and pre-training," Proc. 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, pp.3499–3505, July 2019.

[13] J. Devlin, M.W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," Proc. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis, Minnesota, pp.4171–4186, June 2019.

[14] K. Yang and J. Deng, "Strongly incremental constituency parsing with graph neural networks," Neural Information Processing Systems 33, Online, pp.21687–21698, Dec. 2020.

[15] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R.R. Salakhutdinov, and Q.V. Le, "Xlnet: Generalized autoregressive pretraining for language understanding," Advances in Neural Information Processing Systems 32, Vancouver, Canada, Dec. 2019.

[16] M.P. Marcus, B. Santorini, and M.A. Marcinkiewicz, "Building a large annotated corpus of English: The Penn Treebank," Computational Linguistics, pp.313–330, 1993.

[17] R. Morante, S. Schrauwen, and W. Daelemans, "Annotation of negation cues and their scope: Guidelines v1.0.," Tech. Rep., University of Antwerp, 2011.

[18] Y. Wu and A. Sun, "Negation scope refinement via boundary shift loss," Findings of the Association for Computational Linguistics: ACL 2023, Toronto, Canada, pp.6090–6099, July 2023.

[19] H. Li and W. Lu, "Learning with structured representations for negation scope extraction," Proc. 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia, pp.533–539, July 2018.