Revisiting Syntax-Based Approach in Negation Scope Resolution

Asahi Yoshida[1], **Yoshihide Kato**[2], **Shigeki Matsubara**[1,2]
[1]Graduate School of Informatics, Nagoya University
[2]Information and Communications, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan
yoshida.asahi.y7@s.mail.nagoya-u.ac.jp

Abstract

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.

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 (Read et al., 2012; Carrillo de Albornoz et al., 2012; Abu-Jbara and Radev, 2012; White, 2012). Read et al. (2012) tackled this issue with syntax-based approach and obtained the best performance on the token-level evaluation in *SEM2012 shared task (Morante and Blanco, 2012). Recently, many studies treat this task as a sequence labeling problem and use deep-learning techniques (Fancellu et al., 2016; Khandelwal and Sawant, 2020; Truong et al., 2022). 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[1]. 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. (2012), 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.

[1]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.

*Proceedings of the The 12th Joint Conference on Lexical and Computational Semantics (*SEM 2023)*, pages 18–23
July 13-14, 2023 ©2023 Association for Computational Linguistics

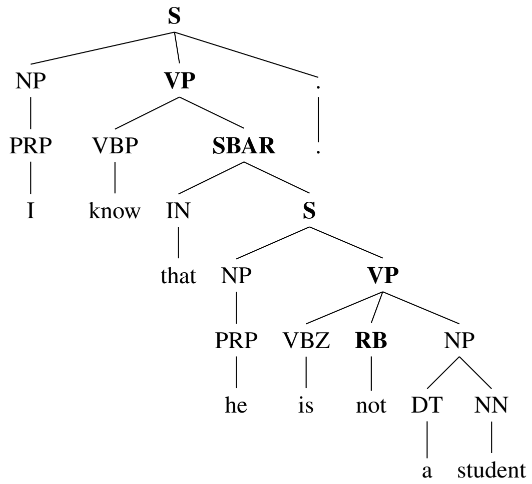

Figure 1: Constituent parse tree of sentence (2), highlighting candidate scope constituents.

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 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 Figure 1).

2. Select one constituent corresponding to the scope 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 Figure 1 are selected as candidates. In the second step, one constituent is selected from the candidates using 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 Figure 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 "`DT//SBAR if SBAR\WHADVP`" will

```
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
```

Figure 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.

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. (2012) adopted some heuristics to remove certain elements from the constituent in the following way:

- Remove the constituent-initial and -final punctuations from the scope.

- Remove certain elements at the beginning or the end of the constituent using *slackening rules*, which consist of five heuristics.

- Apply two post-processing heuristics to handle discontinuous scopes:
  - Remove previous conjuncts from the scope if the cue is in a conjoined phrase.
  - Remove sentential adverbs from the scope.

For sentence (1), the scope "He did, go to school" is correctly resolved using the series of process.

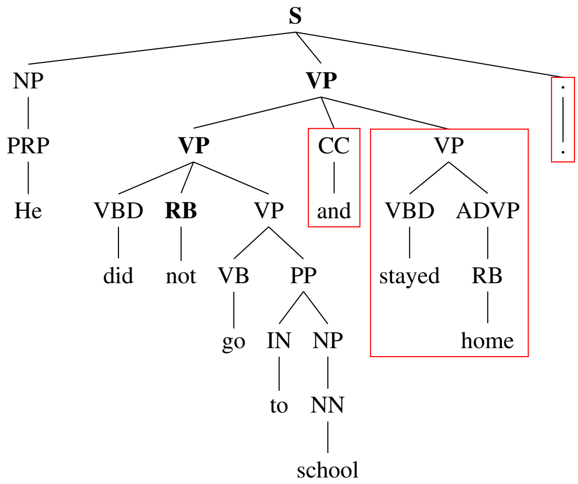

Figure 3: Constituent parse tree of sentence (1), enclosing removed parts in boxes.

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", and its conjunct "stayed home" are removed by the first heuristic rule for discontinuous scope, and the punctuation "." is removed by the above first heuristic rule (removed parts are enclosed in Figure 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. (2012)'s method, respectively.

3.1 Replacement of the Parser

The dataset used in *SEM2012 shared task (Morante and Daelemans, 2012), 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 (2005). As Read et al. (2012) 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 (Kitaev and Klein, 2018; Kitaev et al., 2019) with BERT (Devlin et al., 2019), and Attach Juxtapose Parser (Yang and Deng, 2020) with XLNET (Yang et al., 2019). Table 1 shows the performances of the parsers on Penn Treebank (Marcus et al., 1993).

Parser	**F_1 score (%)**
Reranking Parser (2005)	91.02
Berkeley Neural Parser (2018)	95.77
Attach Juxtapose Parser (2020)	96.34

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

3.2 Modification of *Scope Resolution Heuristics*

Read et al. (2012) used *scope resolution heuristics* shown in Figure 2 to detect the constituent corresponding to the scope of the negation cue. The first rule of Read et al. (2012) (denoted with (#) in Figure 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 Section 2, Read et al. (2012)'s method adjusts the constituent in the third step. This work partially modifies *slackening rules* and post-processing.

In the case of *slackening rules*, we present the following additional rule to the original five rules:

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

This modification was motivated by the annotation guideline of the Conan Doyle dataset (Morante et al., 2011). 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, he should **not** go.

For the post-processing, we modify the second processing: removing sentential adverbs from the

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)

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 Sections 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. (2012)'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.

scope. Read et al. (2012) 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. (2012)'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) There was **no** trace, however, of anything.

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

4 Experiment

To re-evaluate the syntax-based approach to negation scope resolution, we conducted an experiment[2]. This section describes the detail of the experiment. We explain the dataset, settings and results in Sections 4.1, 4.2 and 4.3, respectively.

4.1 Dataset

To evaluate the performance of our work, we used the Conan Doyle dataset, which was employed in *SEM2012 shared task. The dataset is divided into training data, development data and evaluation data. The training data contains 848 sentences including negation, the development data 144 and the evaluation data 235. Note that there can be more than one negation cue in a sentence. Each data contains 984, 173 and 264 negation cues, respectively.

4.2 Experimental Settings

We conducted an experiment using the evaluation data of Conan Doyle dataset. We created new constituent parse trees for the sentences in the dataset using Berkeley Neural Parser and Attach Juxtapose Parser. We did not perform cue detection, that is, we report performance using gold cues. 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. Among the evaluation metrics, the following points should be noted:

- Punctuation tokens are excluded from the evaluation.
- If a sentence contains two or more negation cues, scope predictions for each negation cue are evaluated separately.
- For the scope-level evaluation, a predicted scope is counted as TP if all tokens corresponding to the scope of a negation cue are predicted correctly. Partial matches are counted as FN.

We used the official script distributed in the shared task[3] for evaluation.

4.3 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

[2]The code is available at `https://github.com/asahi-y/revisiting-nsr`.

[3]`https://www.clips.ua.ac.be/sem2012-st-neg/data.html`

Method	Token-level F_1 (%)	
	Including punctuations	Excluding punctuations
This work	91.74	92.61
Fancellu et al. (2016)	88.72	-
Li and Lu (2018)	-	89.4
Khandelwal and Sawant (2020)	92.36	-
Truong et al. (2022)	91.24	-

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.

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 0.62% behind values reported by the state-of-the-art method (92.36%), obtained by Khandelwal and Sawant (2020). 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.

Acknowledgements

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

References

Amjad Abu-Jbara and Dragomir Radev. 2012. UMichigan: A conditional random field model for resolving the scope of negation. In **SEM 2012: 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 (SemEval 2012)*, pages 328–334, Montréal, Canada. Association for Computational Linguistics.

Jorge Carrillo de Albornoz, Laura Plaza, Alberto Díaz, and Miguel Ballesteros. 2012. UCM-I: A rule-based syntactic approach for resolving the scope of negation. In **SEM 2012: 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 (SemEval 2012)*, pages 282–287, Montréal, Canada. Association for Computational Linguistics.

Eugene Charniak and Mark Johnson. 2005. Coarse-to-fine n-best parsing and MaxEnt discriminative reranking. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*, pages 173–180, Ann Arbor, Michigan. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Federico Fancellu, Adam Lopez, and Bonnie Webber. 2016. Neural networks for negation scope detection. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 495–504, Berlin, Germany. Association for Computational Linguistics.

Aditya Khandelwal and Suraj Sawant. 2020. NegBERT: A transfer learning approach for negation detection and scope resolution. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 5739–5748, Marseille, France. European Language Resources Association.

Nikita Kitaev, Steven Cao, and Dan Klein. 2019. Multilingual constituency parsing with self-attention and pre-training. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3499–3505, Florence, Italy. Association for Computational Linguistics.

Nikita Kitaev and Dan Klein. 2018. Constituency parsing with a self-attentive encoder. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2676–2686, Melbourne, Australia. Association for Computational Linguistics.

Hao Li and Wei Lu. 2018. Learning with structured representations for negation scope extraction. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 533–539, Melbourne, Australia. Association for Computational Linguistics.

Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2):313–330.

Roser Morante and Eduardo Blanco. 2012. *SEM 2012 shared task: Resolving the scope and focus of negation. In **SEM 2012: 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 (SemEval 2012)*, pages 265–274, Montréal, Canada. Association for Computational Linguistics.

Roser Morante and Walter Daelemans. 2012. ConanDoyle-neg: Annotation of negation cues and their scope in conan doyle stories. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 1563–1568, Istanbul, Turkey. European Language Resources Association (ELRA).

Roser Morante, Sara Schrauwen, and Walter Daelemans. 2011. Annotation of negation cues and their scope : Guidelines v1.0. Technical report, University of Antwerp.

Jonathon Read, Erik Velldal, Lilja Øvrelid, and Stephan Oepen. 2012. UiO1: Constituent-based discriminative ranking for negation resolution. In **SEM 2012: 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 (SemEval 2012)*, pages 310–318, Montréal, Canada. Association for Computational Linguistics.

Thinh Truong, Timothy Baldwin, Trevor Cohn, and Karin Verspoor. 2022. Improving negation detection with negation-focused pre-training. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4188–4193, Seattle, United States. Association for Computational Linguistics.

James Paul White. 2012. UWashington: Negation resolution using machine learning methods. In **SEM 2012: 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 (SemEval 2012)*, pages 335–339, Montréal, Canada. Association for Computational Linguistics.

Kaiyu Yang and Jia Deng. 2020. Strongly incremental constituency parsing with graph neural networks. In *Neural Information Processing Systems (NeurIPS 2020)*.

Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R Salakhutdinov, and Quoc V Le. 2019. Xlnet: Generalized autoregressive pretraining for language understanding. In *Advances in Neural Information Processing Systems (NeurIPS 2019)*.