Negation Scope Conversion: Towards a Unified Negation-Annotated Dataset

Asahi Yoshida, Yoshihide Kato, Shigeki Matsubara (Nagoya University, Japan)

1. Negation Detection Task & Negation-Annotated Corpora

This work addresses **negation scope resolution**, which is challenging due to its complexity.

- Negation detection task

Examples of negation cues: ***not***, ***no***, ***im**possible*, ***un**known*, ***no longer***

1. Detecting the words that **express negations** **(negation cue detection)**
2. Detecting the **negated parts** of a sentence **(negation scope resolution)**

- Work on negation in natural language processing
  - Automated negation detection [Wu&Sun 2023, Truong+ 2022] etc.
  - Application for downstream tasks [Barnes+ 2021, Mukherjee+ 2021] etc.
  - Analysis of negation processing ability of language models using the negation detection system [Hossain+ 2022, Hossain+ 2020] etc.

- The major corpora of negation scope resolution:

BioScope corpus [Szarvas+ 2008], SFU review corpus [Konstantinova+ 2012], Sherlock dataset [Morante&Daelemans 2012]

Each corpus adopts different annotation schemes.

Corpus	Annotation of negation (bold: negation cue, underline: negation scope)
BioScope	*The park allows fishing in designated areas only and does <u>**not** allow swimming</u>.*
SFU	*The park allows fishing in designated areas only and does **not** <u>allow swimming</u>.*
Sherlock	*<u>The park</u> allows fishing in designated areas only and <u>does</u> **not** <u>allow swimming</u>.*

2. Motivation & Objective

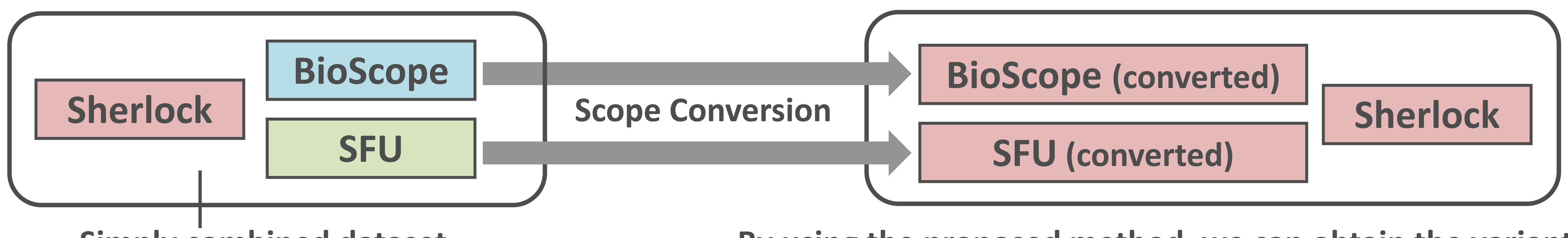
- Motivation
  - Due to the different annotations, negation scope resolution models **perform worse when fine-tuned on the simply combined dataset consisting of the three corpora** [Barnes+ 2021]
  - Solution: **automated conversion of negation scope** **and merging the corpora into a unified dataset**
    - Manual re-annotation consumes much time and effort (negation scope annotation is a complex task, which requires expert linguists)
    - No previous study developed an automated negation scope conversion

- Objective

To propose a method for automatically converting the scopes of BioScope and SFU to those of Sherlock

We select the scope annotation of Sherlock as the target of conversion because it can represent more complex negation scopes compared to those of BioScope and SFU.

By merging Sherlock and the converted version of BioScope and SFU, we have a large training data for negation scope resolution models.



Simply combined dataset consisting of the three corpora

By using the proposed method, we can obtain the variants of BioScope and SFU whose annotation is almost identical to Sherlock.

3. Negation Scope Conversion

A method for automatically converting the scopes of BioScope and SFU (B&S) to those of Sherlock

Our strategy for conversion:
to utilize the correct scope annotations of B&S

Conversion Method

$$\boldsymbol{S}_{\text{left}} \cup \boldsymbol{S}_{\text{cue}} \cup \boldsymbol{S}_{\text{mid}} \cup \boldsymbol{S}_{\text{right}}$$

$$S_{\text{left}} = \begin{cases} \mathrm{L}_c(S_{\text{B\&S}}) & (\mathrm{L}_c(S_{\text{B\&S}} \neq \emptyset)) \\ \mathrm{L}_c(S_{\text{res}}) & (\mathrm{L}_c(S_{\text{B\&S}} = \emptyset)) \end{cases},$$

$$S_{\text{mid}} = \mathrm{M}_c(S_{\text{B\&S}}), \quad S_{\text{right}} = \mathrm{R}_c(S_{\text{B\&S}})$$

- $S_{\text{B\&S}}$: the scopes of B&S
- S_{res} : the results of a scope resolution method (★)
- c : a negation cue
- S_{cue} : internal structure of the negation cue (◆)
- $\mathrm{L}_c(S)$: the left part of the scope S
- $\mathrm{M}_c(S)$: the middle part of the scope S (if present)
- $\mathrm{R}_c(S)$: the right part of the scope S

【Example of the case where $\mathrm{L}_c(S_{\text{B\&S}}) \neq \emptyset$】
*<u>This book</u> **wasn't** <u>published before the end of 2000</u>.*

【Example of the case where $\mathrm{L}_c(S_{\text{B\&S}}) = \emptyset$】
*This computer **isn't** <u>worth your time or money</u>.*

<u>The conversion method is based on these observations:</u>

- $\mathrm{R}_c(S_{\text{B\&S}})$ and $\mathrm{R}_c(S_{\text{SH}})$, as well as $\mathrm{M}_c(S_{\text{B\&S}})$ and $\mathrm{M}_c(S_{\text{SH}})$ can be regarded as almost identical.
  (S_{SH} : the scope of Sherlock)
- $\mathrm{L}_c(S_{\text{B\&S}})$ and $\mathrm{L}_c(S_{\text{SH}})$ can be regarded as almost identical if $\mathrm{L}_c(S_{\text{B\&S}}) \neq \emptyset$

Example of Conversion — In this example, $\mathrm{L}_c(S_{\text{B\&S}}) = \emptyset$.

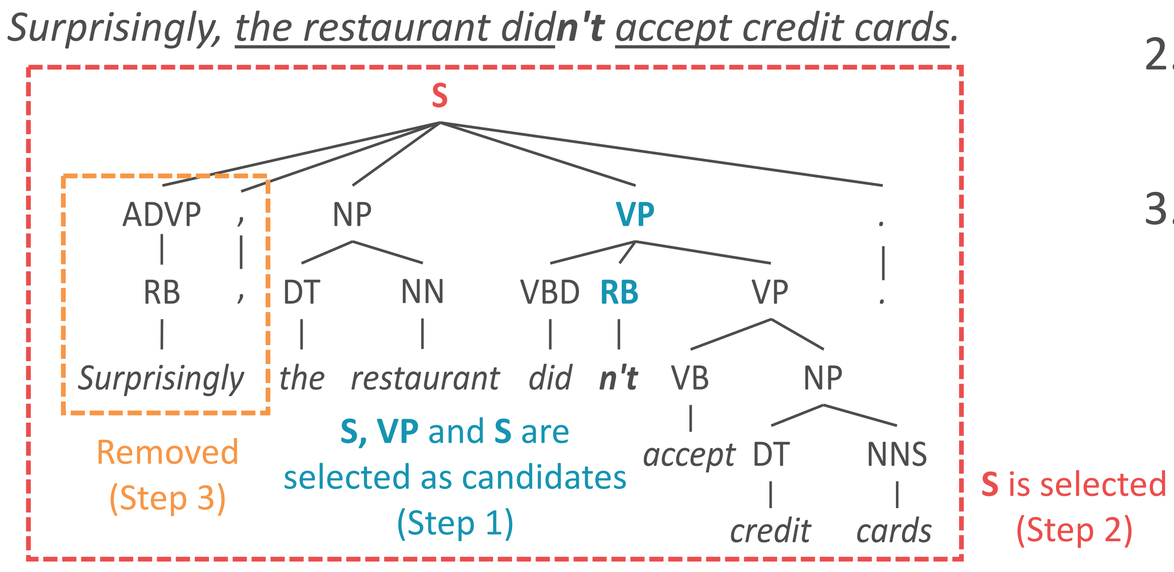
*Surprisingly, the restaurant **didn't** <u>accept credit cards</u>.*

processing for the contracted cue (◆)

*Surprisingly, the restaurant <u>did</u>**n't** <u>accept credit cards</u>.*

scope resolution for the left part of the cue (★) — use the scope of B&S for the right part of the cue

*Surprisingly, <u>the restaurant did</u>**n't** <u>accept credit cards</u>.*



◆ Affixial and Contracted Cues

If B&S's cue c has an affix in V_{aff} or V_{cont} , $\boldsymbol{S}_{\text{cue}}$ is a singleton set consisting of the result by removing the affix from c .

$$\mathrm{V}_{\text{aff}} = \{\text{dis}, \text{im}, \text{in}, \text{ir}, \text{un}, \text{less}\}, \quad \mathrm{V}_{\text{cont}} = \{\text{n't}, \text{not}\}$$

★ Negation Scope Resolution

We obtain $\boldsymbol{S}_{\text{res}}$ using the scope resolution method of Yoshida et al. (2023) with **several modifications**, **which adapt to the domains of BioScope and SFU**.

<u>Yoshida et al. (2023)'s scope resolution method:</u>

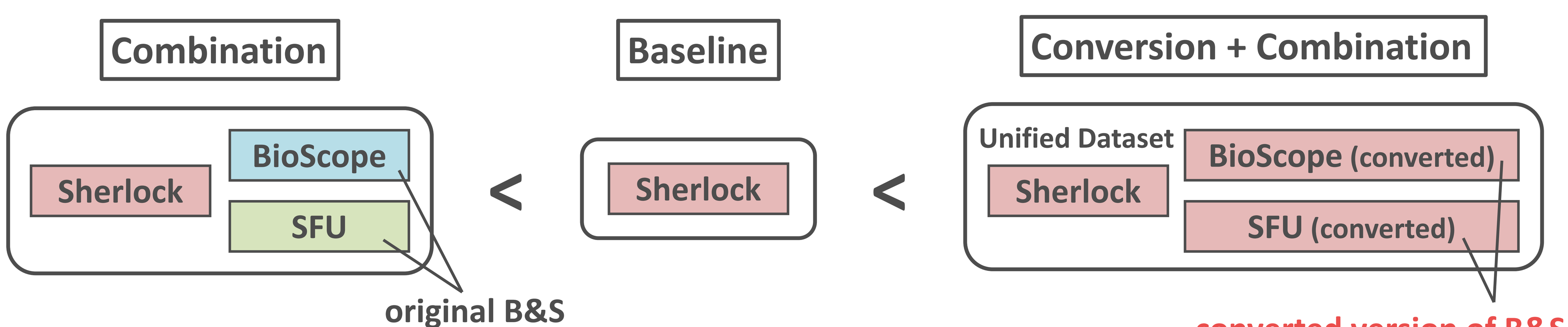
1. Parse the sentence and select the constituents that dominate the cue as scope candidates
2. Select one constituent from the candidates (using the path pattern rules on the syntax tree)
3. Adjust the scope by removing certain elements (using the syntax-based heuristics)

Modification of step 2: **add path pattern rules**
Modification of step 3: **delete some adjustment rules**

4. Negation Scope Resolution with a Unified Dataset

Objective of the experiment is to evaluate the effectiveness of our method in terms of scaling up the dataset through scope conversion.

- Experimental settings
  - We fine-tuned pre-trained language models on the negation-annotated datasets and performed negation scope resolution ※ We used the code and hyper-parameters of Truong et al. (2022)
  - Fine-tuning was conducted with **the three different configurations**:



  - Evaluation metrics:
    token-level and **scope-level** metrics, both of which compute **precision**, **recall**, **$\mathbf{F_1}$ measure**

- Experimental results

Method	Token-level (%) Pre.	Rec.	F_1	Scope-level (%) Pre.	Rec.	F_1
Baseline (BERT)	94.44	89.23	91.76	99.11	71.77	83.25
Combination (BERT)	<u>94.74</u>	87.23	90.83	98.57	66.61	79.48
Conversion + Combination (BERT)	93.84	<u>92.43</u>	**93.13**	98.91	<u>74.21</u>	84.79
Baseline (RoBERTa)	92.08	90.44	91.24	**99.45**	58.60	73.74
Combination (RoBERTa)	93.58	87.29	90.32	<u>99.19</u>	58.44	73.53
Conversion + Combination (RoBERTa)	91.47	92.10	91.76	99.08	60.53	75.14
[Khandelwal&Sawant 2020] (BERT)	–	–	92.36	–	–	–
[Truong+ 2022] (RoBERTa) - Baseline	–	–	91.51	–	–	–
[Truong+ 2022] (RoBERTa) - CueNB	–	–	91.24	–	–	–
[Wu&Sun 2023] (BERT)	**95.12**	90.57	92.77	–	–	<u>85.35</u>
[Wu&Sun 2023] (RoBERTa)	94.54	91.24	<u>92.85</u>	–	–	**87.10**
[Yoshida+ 2023] (heuristics)	89.32	**94.30**	91.74	98.94	**74.70**	85.13

Simple combination of the three corpora (Combination) led to lower performance of the models.
When using the unified dataset (Conversion + Combination), the performance of the models improved.

This supports the effectiveness of our method in terms of scaling up the dataset through scope conversion.