# Learning Interpretable Negation Rules via Weak Supervision at Document Level: A Reinforcement Learning Approach

# Supplementary Materials

## A Learning Parameters

We perform 4000 learning iterations with a higher exploration rate as given by the following parameters[1]: exploration $\varepsilon$ = 0.001, discount factor $\gamma$ = 0 and learning rate $\alpha$ = 0.005. In a second phase, we run 1000 iterations for fine-tuning with exploration $\varepsilon$ = 0.0001, discount factor $\gamma$ = 0 and learning rate $\alpha$ = 0.001. These parameters affect the learning behavior of the agent as follows:

- $\alpha$: The learning rate, set between 0 and 1. Setting it to 0 means that the Q-values are never updated and, hence, nothing is learned. Setting a high value, such as 0.9, means that learning can occur quickly.
- $\gamma$: Discount factor, set between 0 and 1. Determines the importance of future rewards. A factor of 0 will render the agent short-sighted by only considering current rewards, while a factor approaching 1 will cause it to strive for a greater reward over the long term.
- $\varepsilon$: Exploration parameter, set between 0 and 1. Defines the exploration mechanism in $\varepsilon$-greedy action selection. In this strategy, the agent explores the environment by selecting an action at random with probability $\varepsilon$. Alternatively, the agent exploits its current knowledge by choosing the optimal action with probability $1 - \varepsilon$.

## B Comparison to Human Annotations

As part of our robustness checks, we compare the implications of the reinforcement learning policy to annotations of human judges. For this purpose, we use a disjunct dataset that is labeled manually by two external persons (Annotator A and Annotator B). This dataset consists of 500 sentences from movie reviews, with each sentence containing at least one explicit negation phrase from the list of (Jia et al., 2009). Table 1 details the number of equally labeled words and compares to what extend the classifications agree. The comparison is based on all words that are labeled as negated by at least one of the annotators. In addition, we present the inter-rater reliability in terms of Krippendorff's alpha coefficient (Krippendorff, 2013). Here, a reliability value of 1 indicates a perfect overlap between the classifications of words in negated and not negated, whereas a negative value denotes a systematic disagreement.

According to the table, we observe a relatively low agreement between the two human annotators. Only 50.20 % of the words exhibit the same labeling as negated. Unsurprisingly, this also results in a relatively low inter-rater reliability of −0.33. This confirms the high subjectivity of negation scopes as found in previous works (Councill et al., 2010). We also see a large disagreement between the human annotations and the classifications based on the negation policy of our method. For instance, in the case of Annotator A, only 18.81 % of all words are labeled equally. This results in a low inter-rater reliability of −0.68. We observe a similar pattern for Annotator B, where only 25.19 % of all words exhibit the same labeling (inter-rater reliability of −0.60).

## C Exemplary Negation Scope Resolutions

We now illustrate the benefits of our method using two exemplary sentences from movie reviews (see Figure 1). For this purpose, we compare the implications of the reinforcement learning policy to manual annotations of two external language experts (Annotator A and Annotator B).

The first example is given by the sentence *"it's not an original task"*. Here, Annotator A negates all words except *original*, whereas Annotator B negates all words in the sentence. Obviously, these annotations result in sentences with completely different meanings. While Annotator A neglects the fact that there is a *tale*, Annotator B neglects the fact that there is an *original tale*. In our opinion, both interpretations are not completely accurate. In contrast, our method only negates the word *original*, indicating that the reviewer observed a tale that is, however, not original.

[1] Further details regarding the learning parameters are provided in the supplementary materials.

|  | Consensus classification |  |  | Reliability |  |  |
|---|---|---|---|---|---|---|
| **Annotator** | **Annotator A** | **Annotator B** | **Our method** | **Annotator A** | **Annotator B** | **Our method** |
| **Annotator A** | 1.0000 |  |  | 1.0000 |  |  |
| **Annotator B** | 0.5020 | 1.0000 |  | −0.3313 | 1.0000 |  |
| **Our method** | 0.1881 | 0.2519 | 1.0000 | −0.6832 | −0.5974 | 1.0000 |

Table 1: This table compares human annotations for 500 random sentences from movie reviews to the optimal negation policy derived from out method. Reliability is measured in terms of Krippendorff's alpha coefficient (Krippendorff, 2013).

We observe a similar pattern for our second example, which is given by the sentence *"the story is not especially original"*. In this case, Annotator A neglects the fact that the story is original, whereas Annotator essentially neglects the fact that there is a story. Also here, the negation policy based on our method produces more accurate annotations. Specifically, our method only negates the words *especially* and *original*, indicating that there is a story that is, however, not especially original.

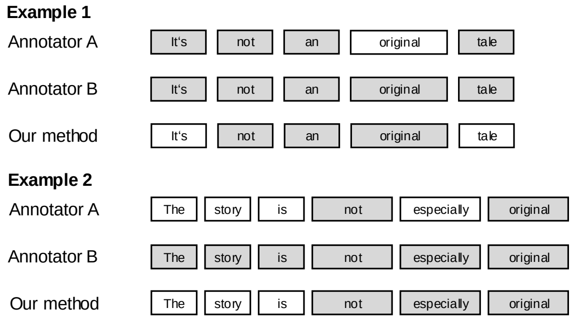

Figure 1: Labels for exemplary sentences from movie reviews. Boxes with gray background color denote words labeled as negated, whereas white boxes correspond to words labeled as not negated.

## References

Isaac G. Councill, Ryan McDonald, and Leonid Velikovich. 2010. What's great and what's not: Learning to classify the scope of negation for improved sentiment analysis. In *Workshop on Negation and Speculation in Natural Language Processing*, pages 51–59. ACL.

Lifeng Jia, Clement Yu, and Weiyi Meng. 2009. The effect of negation on sentiment analysis and retrieval effectiveness. In *CIKM*, pages 1827–1830.

Klaus Krippendorff. 2013. *Content Analysis: An Introduction to Its Methodology*, 3 edition. SAGE Publications, Thousand Oaks, CA.