DT-Neg: Tutorial Dialogues Annotated for Negation Scope and Focus in Context

Rajendra Banjade, Vasile Rus
Department of Computer Science, Institute for Intelligent Systems
The University of Memphis,
Memphis, TN, USA
{rbanjade, vrus}@memphis.edu

Abstract

Negation is often found more frequent in dialogue than commonly written texts, such as literary texts. Furthermore, the scope and focus of negation depends on context in dialogues than other forms of texts. Existing negation datasets have focused on non-dialogue texts such as literary texts where the scope and focus of negation is normally present within the same sentence where the negation is located and therefore are not the most appropriate to inform the development of negation handling algorithms for dialogue-based systems. In this paper, we present DT-Neg corpus (DeepTutor Negation corpus) which contains texts extracted from tutorial dialogues where students interacted with an Intelligent Tutoring System (ITS) to solve conceptual physics problems. The DT-Neg corpus contains annotated negations in student responses with scope and focus marked based on the context of the dialogue. Our dataset contains 1,088 instances and is available for research purposes at http://language.memphis.edu/dt-neg.

Keywords: Negation, Cue, Scope, Focus, Dialogue, Corpus, Annotation

1. Introduction

According to SIL International (Summer Institute of Linguistics), negation is a morphosyntactic operation in which a lexical item denies or inverts the meaning of another lexical item or construction. A *negator* (or negation cue), is a lexical item that expresses negation. **Morphological negation** is when a word is negated by an affix (prefix or suffix) as in **un**-*happy* or *sense*-**less** whereas in **syntactic negation** an entire clause is negated explicitly (using a negator) or implicitly, e.g. verbs or nominalizations that negate their complements such as *fail* or *deny*. In explicitly negated statements, negation is marked using cue words, such as *not, no* and *neither nor*. A negation cue word or negator can affect the meaning of a part of the sentence in which it appears or part of another sentence from the discourse context. The part of the sentence affected by the negation cue is called *negation scope*. The part of the scope that is most prominently or explicitly negated is called *negation focus* (Huddleston & Pullum, 2002).

An example of negation is shown in the following sentence where we indicate the negation cue (in <>), the negation scope (in []) and the negation focus (in {}): *The desk stops moving because **[**there is**]** <**no**> **[{**net force**}** acting on it**]***.

Negation is a frequent and complex phenomenon in natural language. Tottie (1991) reports that negation is twice as frequent in spoken sentences (27.6 per 1,000 words) as in written text (12.8 per 1,000 words). Councill and Velikovich (2010) report that 19% of the product review sentences contain negations. The negation frequency and its key role in many applications such as intelligent tutoring, sentiment analysis, and information extraction, emphasize the importance of the negation handling problem. For instance, in order to understand student responses and provide appropriate feedback in Intelligent Tutoring Systems (ITS; Graesser et al. (2004); Dzikovska et al. (2010); Rus et al. (2013)), the ITSs should be able to handle different linguistic phenomena including negation which is often context dependent as illustrated below.

Example 1:
Question: Do these balls (red ball and blue ball) ever have the same speed?
A1: *They do **not** have the same speed.*
A2: ***No***.
A3: *The balls **never** have the same speed.*

The example above shows three different actual answers from high-school students while interacting with the DeepTutor tutoring system (DT in short; Rus et al., 2013). The answer A2 has a negation cue *no* but the explicit scope is retrieved from the context of the dialogue (the most recent system question in this case). Similarly, to resolve negated pronouns we have to look at the context (Councill, McDonald, & Velikovich, 2010). Due to the frequency (in an analysis of 27,785 student utterances we have found that 9.36% of them contained some form of explicit negation) and complexity of negation in tutorial dialogues, there is a need for a systematic effort to address this problem. To this end, we have annotated a data set to further our understanding of negation and to foster the development of negation handling algorithms.

Existing datasets mostly contain literary or scientific texts and product reviews, and the negation scope and focus are typically located in the same sentence where the negation cue is present (Councill et. al, 2010; Konstantinova et. al, 2012; Morante & Blanco, 2012; Vincze et al, 2008; Pyysalo et al., 2007). However, in conversational texts, the scope and focus of negation may be located in the same utterance or in the previous dialogue context (i.e., inter-sentential negation), such as the previous utterance. In particular, our focus was on annotating explicit negations in student responses recorded during actual tutorial dialogue interactions

between high-school students and an intelligent tutoring system (called DeepTutor) while considering the tutorial dialogue context to interpret students' negations. To the best of our knowledge, there is no such annotated dataset that focuses on the identification of negation cues, scope, and focus in dialogues and in particular in tutorial dialogues.

Our dataset contains 1,088 instances having explicit negations and which were extracted from a collection of 27,785 student utterances.

2. Related Work

Most of the earlier work in negation handling has been done in the biomedical domain. Vincze et al. (2008) annotated negation and hedge cues and their scopes in the BioScope corpus. GENIA Event corpus (Kim et al., 2008) contains annotation of biological events with negation and two types of uncertainty. The BioInfer corpus (Pyysalo et. al, 2007) contains biological relations annotated for negation.

In 2011, Morante, Schrauwen, and Daelemans published a comprehensive guideline for the annotation of negation cues and their scope. In fact, one of the shared tasks in the *SEM 2012 conference was dedicated to negation scope and focus detection (Morante & Blanco, 2012). They released a dataset annotated for scope and focus detection. Scope was annotated in texts from Conan Doyle stories (CD-SCO corpus). Focus was annotated on top of PropBank, which uses the WSJ section of the Penn TreeBank (PB-FOC corpus; Blanco & Moldovan (2011)).

Negation was also studied in the context of sentiment analysis. Councill et al. (2010) focused on explicit negation and created a product review corpus annotated with negation cue and scope. Konstantinova et al. (2012) annotated the SFU Review Corpus which includes annotation of negative/speculative information and their linguistic scope in the review domain.

In summary, though there are several publicly available corpora created to foster the research in the area of negation handling across different domains, to the best of our knowledge, these datasets only focus on non-dialogue texts and typically have the negation scope and focus within the same sentence where the negation cues are located (or at least the data are so annotated). Also, only one dataset (PB-FOC corpus) contains focus annotations but it is not freely available. As mentioned before, our dataset focuses on conversational texts where negation scope and focus detection often relies on context.

3. Data Collection

We created the DT-Neg dataset by extracting student answers containing explicit negation cues from logged tutorial interactions between high-school students and the DeepTutor tutoring system. During the interactions, students solved conceptual physics problems and the interactions were in the form of pure natural language texts (i.e., with no mathematical expressions and special symbols). Each problem contained multiple questions. In 27,785 student responses, we found 2,603 (9.36%) of them contained at least one explicit negation cue word, such as *no* and *not*. We have not considered affixal negations, such as *un* in *un-identified*.

We tokenized the collected dialogue utterances using Stanford CoreNLP Toolkit (Manning et al., 2014). As we focused on explicit negation, we identified student answers containing negation cue words based on a list of cue words which we compiled from different research reports (Morante, Schrauwen, & Daelemans, 2011; Vincze et al., 2008) as well as our own data. If a student response contained multiple negations, they were treated as separate instances in our corpus. We then annotated each such candidate negation instance for negation cue (we simply verified whether the automatically annotated cue word was correct or not), negation scope, and negation focus. The annotation process is described next.

4. Annotation

During the annotation process, annotators were asked to validate the automatically detected negation cue words and identify the corresponding negation scope and focus. It should be noted that we only targeted student responses for negation handling and not all the dialogue utterances, because as the system/tutor utterances are system generated and therefore their interpretation is assumed to be known.

The annotation was conducted by a group (5 people in total) comprised of graduate students and researchers who were first trained before being asked to annotate the data. The annotators had access to an annotation manual for reference. The guideline presented in this paper has been adapted from the guideline prepared by Morante, Schrauwen, and Daelemans (2011). Annotators were instructed to use contextual information to best disambiguate the scope and focus. For this, annotators were shown the student response containing the negation as well as the previous system turn (tutor question). The Example 2 and Example 3 illustrate annotations where the negation information is in the same sentence (Example 2) whereas the negation scope and focus in Example 3 are located in the dialogue context, i.e. the previous dialogues utterance generated by the tutor. The negation cue, scope, and focus are enclosed in <>, [], and {}, respectively.

Example 2:
Question: Do these balls (red ball and blue ball) ever have the same speed?
A: [*They do*] <*not*> [*have the* {*same*} *speed*].

Example 3:
Question: Do [these balls (red ball and blue ball)] ever [have the {same} speed]?
A: <*No*>.

In order to measure the annotation agreement, a randomly selected subset of 500 instances was divided equally into five subsets and each one was annotated by two different annotators. If at least one of the annotators

said the example was invalid (such as *no* in a misspelled word "*pia no*" is not a valid cue), we completely discarded that example. The percentage of agreement of each group was calculated and then averaged. The agreement on scope location, i.e. the same sentence or in the previous dialogue context was very good at 94.33%. When they agreed on the location of scope (focus), we measured the agreement for the actual scope and focus. The average token (sentence) level agreement was 89.43% (66.60%) and 94.20% (66.95%) for scope and focus, respectively. The main disagreement was on how to use the contextual information. The disagreements were discussed among the annotators and fixed. The role of the discussion was to both reach an agreement and improve consistency of future annotations. The rest of the data was divided among the annotators. In total, we have annotated 1,088 valid instances.

5. DT-Neg Dataset Statistics

Some general characteristics about the DT-Neg corpus are offered in Table 1. Different forms of the same cue, such as *n't* or *not* or *NOT* were considered identical while counting unique cues. We observe that 42% of the negation instances required contextual information to properly identify their scopes (and foci).

Parameter	Value
Total number of instances	1088
Instances with scope/focus in context	458
Number of unique cues	10

Table 1: Summary of DT-Neg corpus.

Negation Cue	Frequency
not	250
no	533
n't	104
NO	8
no longer	15
nothing	5
impossible	1
none	43
without	22
Nothing	1
Neither	8
never	5
No	72
NOT	4
Not	4
absence	2
Never	2
neither	8
NOTHING	1

Table 2: Negation cues and their frequencies.

The negation cues and their frequencies are presented in Table 2. Please note that the choice of negation cue may be found different from corpus to corpus as there is no comprehensive list of negation cues. Also, it may vary depending on the domain. For example, *negative* is considered as negation cue in medical domain whereas it is not usually considered as negation cue in other domains.

6. Conclusion

We have presented in this paper our work on creating a new corpus (called DT-Neg) which focuses on negation in tutorial dialogues. However, as the MITRE Corporation explained in their report on why negation is not solved and emphasized on reconsidering negation annotation and evaluation (Wu et. al, 2013), there may be inconsistencies in annotations proposed by various teams. For example, some negation corpora include cue within the scope whereas others don't.

We also developed a system for negation scope and focus detection using this dataset (Banjade, Niraula, & Rus, 2016) and we would like to integrate the system for the purpose of answer grading. In the future, we intend to annotate more samples incorporating the feedback we receive. The dataset and the supplementary documents including annotation details are available at http://language.memphis.edu/dt-neg.

7. Acknowledgments

This research was supported by the Institute for Education Sciences (IES) under award R305A100875 to Dr. Vasile Rus. All opinions and findings presented here are solely the authors'.

8. References

Banjade, R., Niraula, N.B., & Rus, V. (2016). Towards Detecting Intra- and Inter-Sentential Negation Scope and Focus in Dialogue. In *Proceedings of FLAIRS Conference*.

Councill, I. G., McDonald, R., & Velikovich, L. (2010). What's great and what's not: learning to classify the scope of negation for improved sentiment analysis. In *Proceedings of the workshop on negation and speculation in natural language processing* (pp. 51-59). Association for Computational Linguistics.

Dzikovska, M. O., Moore, J. D., Steinhauser, N., Campbell, G., Farrow, E., & Callaway, C. B. (2010, July). Beetle II: a system for tutoring and computational linguistics experimentation. In *Proceedings of the ACL 2010 System Demonstrations* (pp. 13-18). Association for Computational Linguistics.

Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 180-192.

Huddleston, R., & Pullum, G. K. (2002). The Cambridge Grammar of English *Language. Cambridge: Cambridge University Press*, 1-23.

Kim, J. D., Ohta, T., & Tsujii, J. I. (2008). Corpus annotation for mining biomedical events from literature. *BMC Bioinformatics*, *9*(1), 10.

Konstantinova, N., de Sousa, S. C., Díaz, N. P. C., López, M. J. M., Taboada, M., & Mitkov, R. (2012, May). A review corpus annotated for negation, speculation and their scope. In *LREC* (pp. 3190-3195).

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. (2014, June). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 55-60).

Morante, R., & 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-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation* (pp. 265-274). Association for Computational Linguistics.

Morante, R., Schrauwen, S., & Daelemans, W. (2011). *Annotation of negation cues and their scope: Guidelines v1*. 0. Technical report, University of Antwerp. CLIPS: Computational Linguistics & Psycholinguistics technical report series.

Pyysalo, S., Ginter, F., Heimonen, J., Björne, J., Boberg, J., Järvinen, J., & Salakoski, T. (2007). BioInfer: a corpus for information extraction in the biomedical domain. *BMC Bioinformatics*, *8* (1), 50.

Rus, V., D'Mello, S., Hu, X., & Graesser, A. (2013). Recent advances in conversational intelligent tutoring systems. *AI magazine*, *34*(3), 42-54.

Tottie, G. (1991). *Negation in English speech and writing: A study in variation*. San Diego/New York: Academic Press.

Vincze, V., Szarvas, G., Farkas, R., Móra, G., & Csirik, J. (2008). The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes. *BMC Bioinformatics*, *9* (Suppl 11), S9.

Wu, S., Miller, T., Masanz, J., Coarr, M., Halgrim, S., Carrell, D., ... & Bedford, M. A. Negation is not solved: reconsidering Negation Annotation and Evaluation. 2013.

http://www.mitre.org/sites/default/files/publications/13-3434.pdf (accessed on Sept 18, 2015).