HiER 2013

Griesbaum, Heuwing, Ruppenhofer, Werner (Hrsg.)

HiER 2013

Proceedings des 8. Hildesheimer Evaluierungs- und Retrievalworkshop

Hildesheim, 25.–26. April 2013

J. Griesbaum, B. Heuwing, J. Ruppenhofer, K. Werner (Hrsg.):
HIER 2013 - Proceedings des 8. Hildesheimer Evaluierungs-
und Retrievalworkshop, Hildesheim 2013

IGGSA-STEPS: Shared Task on Source and Target Extraction from Political Speeches

Josef Ruppenhofer[1], *Julia Maria Struß*[1], *Jonathan Sonntag*[2], *Stefan Gindl*[3]

[1]University of Hildesheim
Department of Information Science and Natural Language Processing
Marienburger Platz 22, 31141 Hildesheim
{josef.ruppenhofer | julia.struss} @ uni-hildesheim.de

[2]University of Potsdam
Computational Linguistics
Karl-Liebknecht-Straße 23-24, 14476 Potsdam
jsonntag @ uni-potsdam.de

[3]MODUL University Vienna
Department of New Media Technology
Am Kahlenberg 1, 1190 Wien
stefan.gindl @ modul.ac.at

Abstract

In this paper, we report on the definition of a shared task considering source (whose opinion?) and target (about what?) extraction in protocols of the Swiss parliament that will be conducted by the Interest Group on German Sentiment Analysis (IGGSA)[1].

1 http://iggsa.sentimental.li/

1 Introduction

In fine-grained opinion mining/sentiment analysis, the source or holder of an opinion as well as the target or object of opinion often need to be extracted from the text rather than being available from metadata. This is the case especially for text types that contain multiple sources and targets. While it may appear that sources and targets should be intuitively and reliably recognizable to humans, and we should thus find high inter-rater agreement among annotators, actual attempts at defining and annotating the concepts show that this is not so (see e. g. NTCIR-6: OAT (Seki et al. 2007)).

We report on the definition of a shared task related to source and target extraction in protocols of the Swiss parliament that the IGGSA is in the process of preparing and that is to be held in the run-up to the KONVENS conference in 2014. We discuss various ways of defining the notions source and target and report on the development of the annotation scheme.

2 Related Work

Sentiment analysis systems must be able to reliably tie opinions or subjective states to their sources and targets. This is a non-trivial task as some sentiment-bearing expressions are not linked to the sources, and some not even to the targets, of opinion. In the best case, the source and target correspond to semantic roles of sentiment-bearing predicates that can be expressed as syntactic arguments (cf. Ruppenhofer et al. 2008). However, a direct tie-in with semantic role labeling is usually not the chosen way of handling the source and target extraction. In what follows, we discuss the reasons for this and some of the adopted alternative problem statements.

In the case of one important sub-class of sentiment-bearing expressions, called expressive subjective elements by Wiebe et al. (2005), a grammatical link exists only between the opinion expression and the target, but not to the source. For instance, in the case of *idiotic* we know that what the adjective modifies or is predicated of is the target of the sentiment conveyed. Note, however, that the sources may differ between the two following examples: in

(1), the source is the writer of the text, whereas in (2) it is the quoted speaker Irvine.

(1) His rude, crude response and IDIOTIC exit from his duties is hardly deserving of the praise he has attracted. [www]

(2) "That was IDIOTIC," Irvine told talkSPORT. [www]

Rather than connect expressions of opinion only to their immediate sources, it is desirable to keep track of the chain of transmission. In the MPQA-corpus (Wiebe et al. 2005), levels of nesting are recorded that would show for a sentence like (2) that not only is Irvine the source of the opinion but also the writer of the text.

The main issue regarding targets is whether the analysis should address only what one may call "local" targets – expressions that are semantic valents and syntactic dependents of a particular sentiment-bearing predicate – or whether other pragmatically relevant targets should also be taken into account. To illustrate the difference, consider the following pair of examples:

(3) a. I am not a Dortmund fan - I am a Schalke fan - but I am GLAD + [Dortmund beat Bayern TARGET].

b. I am not a Dortmund fan - I am a Schalke fan - but I am GLAD Dortmund beat - [Bayern TARGET].

(3a) displays the stable, "literal" sentiment that is conveyed by the sentence: that the speaker is glad about the reported event. (3b), by contrast, displays an inferred sentiment: that the speaker specifically dislikes Bayern's team.

Very much a pragmatic understanding of targets is adopted by Stoyanov and Cardie (2008). They suggest a definition of opinion topic and present an algorithm for opinion topic identification that casts the task as a problem in topic co-reference resolution. They distinguish between *Topic* (subject of the opinion as intended by the opinion holder), *Topic span* (minimal span of text mentioning the topic) and *Target span* (syntactic surface form comprising the contents of the opinion). Example (4) gives an example of Stoyanov and Cardie's, showing through the question marks that there are multiple entities that might be perceived to refer to a relevant topic. Notice that *think* is treated as the relevant subjective expression.

(4) [OPINION HOLDER Al] thinks that [TARGET SPAN [TOPIC SPAN? the government] should [TOPIC SPAN? tax gas] more in order to [TOPIC SPAN? curb [TOPIC SPAN? CO2 emissions]]].

Thus, what Stoyanov and Cardie are after are certain inferred targets which may be more important on the text-level than the overt target and which may not have syntactic relations to the subjective expression.

In terms of prior shared tasks in opinion mining, the most relevant work to ours was done in the context of the Japanese NTCIR[2] Project. In the NTCIR-6 Opinion Analysis Pilot Task (cf. Seki et al. 2007), offered for Chinese, Japanese and English, sources and targets had to be found relative to whole opinionated sentences. If multiple opinions were expressed in a given sentence, the task allowed for multiple opinion holders, possibly occurring anywhere in the document. In the evaluation, where necessary, co-reference information was used to (manually) check whether a system response was part of the correct referent's chain of mentions. The sentences in the document were judged as either relevant or non-relevant to the topic (=target). Polarity was determined for each opinionated sentence and in case of multiple opinions, the polarity of the main opinion expressed was chosen.

All sentences were annotated by three assessors, allowing for strict and lenient (by majority vote) evaluation. The successor tasks, NTCIR-7 and -8: Multilingual Opinion Analysis, (Seki et al. 2008, 2010) were basically similar in their setup.

While the STEPS-task will focus on German, the most important difference to the shared tasks organized by NTCIR is that it defines the source and target extraction task at the level of individual subjective expressions. A key aspect that is yet to be resolved in the development of the STEPS-TASK is what kinds of units may be considered as the relevant subjective expressions, as discussed further below in section 4.

2 NII [National Institute of Informatics] Test Collection for IR Systems

3 Data

The data set comes from the Swiss parliament[3]. The choice of this particular data set is motivated as follows. On a pragmatic level, it was important that the source data along with our annotations could be redistributed freely. In terms of the task itself, we were interested in having a text type in which multiple sources and targets would be relevant. Further, several IGGSA-members are interested in, and collaborating with researchers in, political science in the context of the digital humanities.

The data set does have some special characteristics, though, when compared to other German-language data sets. First, it reflects the Swiss variety of German. Second, the Swiss parliament is a body that operates multilingually. In order to keep the influence of these characteristics to a minimum, only speeches in German discussing foreign affairs are selected for the tasks' gold standard.

4 Development of the annotation scheme

In order to meet different research interests of the IGGSA-members, a new annotation scheme is being developed based on the annotations of a first explorative annotation step. In this step 50 sentences are annotated by four independent annotators with respect to opinions, targets and sources with the only requirement, that sources and targets are to be annotated at the level of individual subjective expressions and all nested targets and holders have to be considered. The annotators are asked to comment their annotation decisions, which will then be used for the development of an initial annotation scheme that will be validated in a second annotation step with different annotators on the same data, before the actual annotation of the gold standard sentences will be conducted.

3 Schweizer Bundesversammlung: http://www.parlament.ch/ab/frameset/d/index.htm

While we want the task to be focused on the extraction of sources and targets, it is also crucial to have good agreement on subjective expressions since source and targets exist only relative to them. Consider example (5), slightly modified from the MPQA.

(5) The United States has [CONSISTENTLY ATTACHED] [GREAT IMPORTANCE] to this issue.

The MPQA annotations treat *consistently attached* as one subjective expression (a direct subjective) and *great importance* as another (an expressive subjective element). If one accepts this analysis, what should the sources and targets be for these two expressions? An alternative analysis that fuses the two subjective expressions into one, recognizing that *attach* is a support verb for *importance*, would avoid the problem of what to treat as the target of *attach*. Other problems may result when the (relative) lack of constraints on what may count as a subjective expression leads to a situation where the subjective expression overlaps the target.

5 Task description

We plan on offering one main task and two subtasks in the context of STEPS. In the main task the participants are asked to identify subjective expressions *and* their respective opinion holders and opinion targets. In the subtasks, these two aspects are divorced. For subtask 1 the participants are given the subjective expressions and are only asked to identify opinion holders and/or opinion targets. Subtask 2 is considered with the identification of subjective expressions and their polarities and strength. Participating in the main task precludes participation in subtask 1 and vice versa.

6 Conclusion

In this paper we discussed various ways of defining opinion targets and holders with more semantic/syntactic ways on the one hand and more pragmatic

ways on the other and reported on the task definition as wells as first ideas for the development of a new annotation scheme to be employed in the development of the gold standard for the shared task that will be held in the run-up of the KONVENS conference in 2014.

References

Ruppenhofer, J.; Somasundaran, S.; Wiebe, J. (2008): Finding the Sources and Targets of Subjective Expressions. In: Proceedings of the LREC, Marrakech, Morocco, 26 May - 1 June 2008, pp. 2781–2788.

Seki, Y.; Evans, D.; Ku, L.; Chen, H.; Kando, N.; Lin, C. (2007): Overview of Opinion Analysis Pilot Task at NTCIR-6. In: Proceedings of the 6th NTCIR. Tokyo, Japan, May 15-18, 2007, pp. 265–278.

Seki, Y.; Evans, D.; Ku, L..; Le Sun; Chen, H.; Kando, Noriko (2008): Overview of Multilingual Opinion Analysis Task at NTCIR-7. In: Proceedings of the 7th NTCIR, Tokyo, Japan, December 16-19, 2008, pp. 185–203.

Seki, Y.; Ku, L.; Sun, L.; Chen, H.; Kando, N. (2010): Overview of Multilingual Opinion Analysis Task at NTCIR-8. A Step Toward Cross Lingual Opinion Analysis. In: Proceedings of the 8th NTCIR. Kyoto, Japan, June 15-18, 2010, pp. 209–220.

Stoyanov, V.; Cardie, C. (2008): Topic identification for fine-grained opinion analysis. In: Proceedings of the 22nd COLING. Manchester, UK, 18-22 August 2008, pp. 817–824.

Wiebe, J.; Wilson, T.; Cardie, C. (2005): Annotating Expressions of Opinions and Emotions in Language. In *Language Resources and Evaluation* 39 (2), pp. 165–210.