Argumentation Mining in Political Debates

Shohreh Haddadan, Elena Cabrio, Serena Villata, Leendert van der Torre

University of Luxembourg, Universite' Côte d'Azur

Arguments may occur in various contexts from debate forums on the Web to discussions among the jury members in the Court. The automatic process of extracting arguments from natural language text is known as *Argument(ation) Mining* (AM) (Torroni & Lippi, 2016).

Political debates are one of the application scenarios of major interest for this research area. In these debates, arguments are constructed to justify the plans the candidates' parties advocate or the stance they take towards certain topics such as abortion, tax cuts, immigration. In the election debates, the candidates attempt to make their claims appealing, while discrediting the opponent by attacking the strength of his/her claims. Thus, the increasingly available data of online political forums, transcripts of televised discussions or parliamentary debates provide researchers with a huge amount of textual data, from which fallacies, persuasive arguments, incoherence among the arguments proposed by a candidate, can be automatically identified.

This study focuses on extracting arguments from 41 transcripts of the presidential and vice-presidential debates organized by the Commission on Presidential Debates (CPD) in the United States from 1960 to 2016. The dataset contains approximately 6800 text snippets, divided based on the speech turns for each candidate.

The first task when implementing an Argumentation Mining pipeline to extract arguments from raw text is to gather an annotated dataset. The annotation of an argument structure is not a straightforward task, hence it is challenging for annotators to annotate arguments, and their components (i.e., premises and claims) in the text. To support the annotators in their task, we defined a set of guidelines containing the definitions of the components composing an argument, i.e., claims and premises (Stab & Gurevych, 2014). Our annotation scheme does not include the annotation of major claims due to the diversity of topics in each of the transcripts. A claim is the portion of the argument, which needs to be justified by a set of reasons called premises (Cabrio & Villata, 2013). Various examples from the dataset are provided in the guidelines for the annotators.

Political debates are rifled with different types of claims including suggestions about a certain policy which is advocated by the debater or her party, as shown in Example 1. Claims might express stances taken towards a certain subject such as in Example 2, or they can state an opinion or a belief such as in Example 3. The debaters also often make assertions concerning their capabilities such as in Example 4[1].

1. [**I feel that another effort should be made by a new Administration in January of 1961, to renew negotiations with the Soviet Union**[2]]Claim.

---

[1] Claims are written in bold and premises are written in italics. Component boundaries can be distinguished by [square brackets]

[2] Kennedy, 13 October 1960 in debate against Nixon.

2. [**The present tax structure is a disgrace to this country**[3]]Claim.
3. [**George Bush, who I think is one of the finest Vice Presidents this country has ever had**]Claim.
4. [**I believe I'm going to win**[4]]Claim, because [**the American people know I know how to lead**]Claim. [**I've shown the American people I know how to lead**]Claim.

The second component defined in our annotation scheme are premises. Premises in political debates are also proposed in different forms consisting of facts, statistics, quotations, reports or examples, findings.

Example 5 contains two premises in which the debater provides statistics as premise to support his claim. Example 6 is an instance of a premise in which the debater justifies his plan by stating that experts have approved it. In the first three premises of Example 7, some facts as premises, and a report as the last premise to support the asserted claim have been exerted.

5. [**The present tax structure is a disgrace to this country**]Claim. [**It's just a welfare program for the rich**]Claim2. [*As a matter of fact, uh - 25 percent of the total tax deductions, go for only 1 percent of the richest people in this country*]Premise1, and [*over 50 percent of the tax uh credits go for the 14 percent of the richest people in this country*]Premise2.

6. [*I have submitted an economic plan that I have worked out in concert with a number of fine economists in this country, all of whom approve it*]Premise1, and [**believe that over a five year projection, this plan can permit the extra spending for needed refurbishing of our defensive posture**]Claim1, that [**it can provide for a balanced budget by 1983 if not earlier**]Claim2.

7. [**Race remains a significant challenge in our country**]Claim1. Unfortunately, [**race still determines too much**]Claim2, [*often determines where people live*]Premise1, [*determines what kind of education in their public schools they can get*]Premise2, and, yes, [*it determines how they're treated in the criminal justice system*]Premise3. [*We've just seen those two tragic examples in both Tulsa and Charlotte*]Premise4.

Table 1 illustrates the percentage of annotation of the single components, based on sentence-level segmentation on the annotated portion of the dataset.

Table 1: Percentage of annotated components based on sentence-level segmentation

	Sentence-Level %
Claim	44.80
Premise	44.36
Not annotated	13.49

The annotation process is a cyclic process, which consists in reading the guidelines, performing the annotation, evaluating the annotation carried on by the annotators and

---

[3] Carter, 23 September 1976 in debate against Ford.
[4] Bush, 30 September 2004 in debate against Kerry.

revising the annotations inconsistent with respect to the guidelines. The reliability of the annotation is measured through the calculation of the inter-annotator agreement, e.g., by using Kripendorff's α for sentence-level segmentation. An early computation of the agreement has been computed on the portion of the dataset that has been annotated and depicts an observed agreement 0.83 and Kripendorff's α of 0.62 which is slightly less than the accepted threshold of 0.67 for reliable annotation. However, we expect that this agreement elevates as the annotation process advances with more trial and guideline revision cycles.

At the end of the annotation phase, the annotated dataset is going to be used as an input to the AM pipeline to automatically extract the claims made in the arguments and the premises on which the claims are reasoned upon. We also plan to address the further AM task of predicting the attack or support relations between different arguments.

The aim of this study is to help researches in the field of social sciences, public address, rhetoric and history to analyze the large amount of discussions that is now available mostly by means of digitization of textual resources. Applications ranges from studying the impact of persuasive arguments, tracking the coherency of candidates' claims in debates, detecting fallacies in arguments, exploiting the audience's sentiments to construct the arguments, and altering the reasoning methods through time.

References

Cabrio, E., & Villata, S. (2013). A natural language bipolar argumentation approach to support users in online debate interactions. *Argument & Computation 4.3*, 209-230.

Stab, C., & Gurevych, I. (2014). Annotating argument components and relations in persuasive essays. *In Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, (pp. 1501-1510).

Torroni, M., & Lippi, P. (2016). Argumentation Mining: State of the Art and Emerging Trends. *ACM Trans. Internet Technol., 16*, 10:1-10:25. doi:10.1145/2850417