May 18th, 9:00 AM - May 21st, 5:00 PM

Commentary on “Argumentation Mining in Parliamentary Discourse”

Moira Kloster
University of Fraser Valley

Commentary on "Argumentation Mining in Parliamentary Discourse"

MOIRA KLOSTER
Philosophy
University of the Fraser Valley
33844 King Road, Abbotsford, B.C. V2S 7M9
Canada
Moira.Kloster@ufv.ca

1. Introduction

Naderi is engaged in a formidable task. For those of us who examine a single argument through a jeweller's loupe for a good close-up of its detail, the thought of combing many years of parliamentary records for all their political arguments is utterly intimidating. As her paper indicates, the task is not easy, even with the benefit of current computing power, because the computer must be trained to pick out from written transcripts the details needed to identify that an argument is being given, and identify at least enough of its structure to show which line of reasoning is being used to reach what conclusion.

This challenging line of research offers promises that include, for example, the ability to answer questions about how argumentation changes over the years and from place to place. I remember a conversation with John Woods, about eight years ago at OSSA, in which we speculated that argumentation in Canadian public discourse has changed since the 1960s. We both suspected that argument had become less common overall, and that arguments are now shorter and simpler in structure than they used to be. Neither of us could see any easy way to answer this question. What was not possible then is now open through this large-scale "mining" of sources such as parliamentary debates and news media opinion sections.

I am not qualified to assess the details of the computational analysis Naderi employs. It has been an interesting learning experience just to come up to speed on the literature about argument mining and frame identification. My comments, therefore, will focus on the connection of her research to existing questions in argumentation.

2. Framing and supervising argument mining

First, I would like to extend the background provided in the paper. Argument mining is the empirical search for sequences of claims made in arguments and sequences of arguments employed in debate. Lawrence and Reed (2015) define argument mining as "the automatic identification of the argumentative structure contained within a piece of natural language text" (p. 127). One research institute, the Ubiquitous Knowledge Processing (UKP), lists the following three objectives for its argument-mining research:

- Identifying argument components in different text types
- Recognizing relations between argument components
- Automatic assessment of argumentation quality

Bondy, P., & Benacquista, L. (Eds.). *Argumentation, Objectivity, and Bias: Proceedings of the 11th International Conference of the Ontario Society for the Study of Argumentation (OSSA), 18-21 May 2016*. Windsor, ON: OSSA, pp. 1-5.

If it is possible to do all of this, then much of what we currently teach in critical thinking courses could be done for anyone by computer. My students will welcome the day when an argument-processor becomes part of their calculator or their cellphone and they no longer need a textbook, let alone a course. More seriously, as researchers in argumentation we would then have access to benefits such as the ability to diagram and summarize arguments as they actually occur, in sources such as meeting transcripts, online discussion, media sources, and, as here, parliamentary debate. In addition, there are applications to on-line learning, especially for large courses such as MOOCs, where argument-mining of student responses to essay questions or on-line forums would provide an automated way to locate and analyse argument structure which in turn would identify students' level of success in comprehending and producing cogent arguments (2nd Workshop on Argumentation Mining, 2015).

Meeting these goals, however, is evidently a considerable challenge for computing. Nevertheless, researchers in this field are clearly making progress towards being able to extract quite detailed structure from written text. Naderi's paper is one valuable step in the process of making it possible to identify and classify arguments, and the challenges in this step illustrates some of the complexity in reaching the desired goals.

What can be done already is to identify a stance or attitude, for or against a position, and Hasan and Ng (2014) use the "close interplay" between stance and reason to help identify reasons as well (p. 751).

What can also be done is to identify differences in reasons offered for taking this stance, by annotating the reasons according to a classification system. The principal choice of classification system is categorization by "frames". Although, as Naderi notes, there is no clear common strand of agreement on what a frame is, Tsur, Calacci, & Lazer (2015) offer a definition particularly relevant to argumentation: "Framing is a sophisticated form of discourse in which the speaker tries to induce a cognitive bias through consistent linkage between a topic and a specific context (frame)" (p. 1659).

How to identify and then select frames which will adequately indicate the relationship between reason and conclusion is evidently tricky, as indicated in Naderi's paper. The same frame can be offered on either side: the frame only indicates which type of reason is salient, such as whether the key element of an issue is freedom of speech, not whether the arguer will conclude that freedom of speech must outrank other concerns. There are patterns of correlation between frames and positions.

Naderi contrasts her study of parliamentary discourse with earlier studies focussed on user-generated content in online forums, but does not indicate why it is important to make this switch. I take it that part of the answer is that earlier studies of political speeches have, as she says, only identified the topics discussed, without tackling the challenge of connecting frames to positions. Another part of the answer appears to be that, as Chong and Druckman (2007) state, parliamentary debate is a prime example of multiple frames competing for the attention of the audience.

> virtually all public debates involve competition between contending parties to establish the meaning and interpretation of issues. When citizens engage an issue—be it social security, foreign aid, a hate-group rally, affirmative action, or the use of public funds for art—they must grapple with opposing frames that are intended by opinion leaders to influence public preferences. (Chong & Druckman, 2007, p. 100)

Typically, this competition for public support results not simply in two opposing frames but in "multiple frames [presented] with varying frequencies" (Chong & Druckman, 2007, p. 102).

When there are multiple frames, and at least two competing stances—exactly the sort of discourse that argumentation theorists most want to be able to consider—it evidently becomes more challenging to connect frames and stances accurately. In this paper, the first concern is how best to select frames. Part of what emerges from Naderi's study is that predicted frames are not always reliable: anecdotes and recognition of opposing points can generate errors.

The focus of the debate in the literature is on the effectiveness of a frame in shaping or changing public opinion—what does happen when arguments are used, as opposed to what should happened when arguments are used. It also stays alert to the difficulty of agreeing on what should count as a frame. I would like to highlight another connection, only touched on in the literature, but of particular relevance to argumentation: the connection between framing and inaccurate reasoning.

Framing is already notorious as a contributor to fallacious reasoning. The impact of cognitive bias on logical acumen is considerable. Research in cognitive psychology has demonstrated how significantly a shift in frame can affect a reasoner's ability to detect whether two arguments or logical problems are identical in logical structure. Borah (2011), reviewing a decade's worth of literature on the theory of framing, notes that

> Kahneman and Tversky (1979, 1984) were the first to demonstrate how different presentations of essentially the same information can have an impact on people's choices. They found that individuals were inclined to take risks when ''losses'' are highlighted. But when the same information is presented in terms of ''gains,'' individuals shy away from risks. (p. 248)

While Kahneman and Tversky were among the earliest and most notable of the researchers in this area, other lines of research have noted the same effect. The "Wason 4-card test", a standard test of logical reasoning ability, has been tried in many forms which indicate that framing significantly affects test subjects' ability to recognize logical equivalence between two versions of the same problem. Subjects' accuracy varies depending on whether the problem is purely abstract or is presented in a familiar social context such as getting permission. Accuracy even varies when the problem itself is not varied in wording but a key background detail is changed. For example, when test subjects were asked which two islanders must be checked to see if all four of them are following the rule, "If you go out at night, you must wear a volcanic rock tied around your ankle" (a version of the Wason 4-card test), the accuracy rate was greater if subjects were also told that they were the mother of the four islanders than if they were told they were anthropologists investigating the island's customs. [I can't find the precise reference for this, but I believe it is in P. Cheng and K. Holyoak's research, 1985-1986. Another study testing the "social context" hypothesis, with a good review of this literature, is Deshon, Smith, Chan and Schmitt (1998).]

Fortunately, as Entman (1993) notes, the impact of a frame is not identical on everyone: neither good reasoning nor logical fallacies are hardwired in the human brain (p. 54). Once argument-mining can drill down to the level of individual responses to the same argument it may help us get a broader picture of which people most successfully resist errors under which circumstances. Some studies have moved in this direction. Chong and Druckman (2007) note that the impact of a frame varies depending on the audience's knowledge, motivation, frequency of exposure, and subjective impression of the strength or weakness of a particular frame.

One additional application of Naderi's research is to identify whether particular frames change in frequency over time. I confess to being a little disappointed that the present study does not mention whether there were significant differences between the two different parliamentary debates on same-sex marriage under the Liberals in 2005 and the Conservatives in 2006. Perhaps there weren't, but it would be interesting to use large-scale mining to tell which frames are used most often and are presumably considered most central to an issue, and be able to test whether there are changes over time in where opposing sides find common ground.

3. Conclusion

According to Zhu and Goldberg (2009), Naderi is right to expect that a semi-supervised learning model will be more cost-effective and less labour intensive than the supervised approach used in the present study. If so, this will presumably make it possible to enlarge considerably the number of arguments that can be accurately identified. However, she also notes that given the errors which can still occur with the study's present framing strategy, "more complex approaches will be required to successfully recognize the frames". Evidently, considerable challenges still lie ahead.

I wish Naderi and other researchers in this field the best of luck in training computers to put the rest of us out of business as argument analysts.

References

2nd Workshop on Argumentation Mining (2015). Workshop Description. Retrieved from https://www.cs.cornell.edu/home/cardie/naacl-2nd-arg-mining/

Borah, P. (2011). Conceptual issues in framing theory: A systematic examination of a decade's literature. *Journal of Communication* 61 (2), 246-263.

Chong, D., & Druckman, J. N. (2007). A theory of framing and opinion formation in competitive elite environments. *Journal of Communication* 57 (1): 99-118.

Deshon, R., Smith, M., Chan, D., & Schmitt, N. (1998). Can racial differences in cognitive test performance be reduced by presenting problems in a social context? *Journal of Applied Psychology* 83 (3), 438-451.

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication* 43 (4), 51–58.

Hasan, K. S., & Ng, V. (2014). Why are you taking this stance? Identifying and classifying reasons in ideological debates. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, 25-29 October 2014*, (pp. 751–762). Stroudsburg, PA: Association for Computational Linguistics.

Lawrence, J., & Reed C. (2015). Combining argument mining techniques. *Proceedings of the 2nd Workshop on Argumentation Mining, 4 June 2015*, (pp. 127–136). Stroudsburg, PA: Association for Computational Linguistics.

Tsur, O., Calacci D., & Lazer D. (2015). A frame of mind: Using statistical models for detection of framing and agenda setting campaigns. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Vol. 1: Long papers), 26-31 July 2015*, (pp. 1629–1638). Red Hook, NY: Association for Computational Linguistics.

Ubiquitous Knowledge Processing (UKP) (2016). Argumentation mining. Retrieved from https://www.ukp.tu-darmstadt.de/ukp-home/research-areas/argumentation-mining/

Zhu, X., & Goldberg, A. B. (2009). Introduction to semi-supervised learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*. Retrieved from http://www.morganclaypool.com/toc/aim/3/1