Toward audience-aware argument generation

Milad Alshomary1, 2 and Henning Wachsmuth1
1Department of Computer Science Paderborn University, Paderborn, Germany
2Correspondence: milad.alshomary@uni-paderborn.de

The maturity of the computational argumentation field, demonstrated with the first live debate between a machine and a human,1 triggers a demanding question: how can we build argumentation technologies that bring people together? We believe that an important part of the answer is to include the audience’s beliefs into the process.

Argumentation is omnipresent. We face it in everyday life, from small talks at the dinner table to political debates on large podiums. One main goal of argumentation is to resolve disputes between disagreeing parties. While this goal seems straightforward, many times we find ourselves frustrated because we fail to convince others of our views or even to reach a middle ground when discussing some sort of controversial topic. The source of this disagreement is grounded in one’s prior beliefs. Skilled debaters take into consideration their target audience and frame their arguments accordingly to achieve better agreement. While the field of social psychology is ripe with studies on the role of prior beliefs in argumentation, the field of computational argumentation, which aims to model the argumentation process, is still catching up. If we aim to build argumentation technology that reaches people and bridges the disagreement gap, then this technology must be aware of the prior beliefs of its target audience and acts upon that.

This article will discuss the need to consider the audience’s prior beliefs in computational argumentation. We will then highlight two areas of research that could contribute toward building an audience-aware argumentation technology along with the potential challenges and possible opportunities in each area.

What is computational argumentation?

In the field of computational argumentation, we work on computationally modeling the way humans argue. These computational models enable the machine to understand and synthesize argumentation in natural language texts. This is inherently an interdisciplinary field of study where argumentation, linguistics, social psychology, and computer science meet. The recent advances in natural language processing enabled the delivery of end-user applications in which a machine can engage in debates against humans1 or allow users to search diverse arguments on controversial topics.2 Other applications like writing support systems and decision-making assistants can also benefit directly from this research field. One shared goal of these applications is to achieve agreement between disputed parties—similar to the aim of argumentation.

The role of audience in argument generation

The audience plays an essential role in argumentation. A successful argument that could bridge the disagreement gap is the one carried out with audience awareness. While several studies in the social psychology field have analyzed the effect of audience’s beliefs on persuasiveness, the computational argumentation field still lags behind such analysis. In his book The Righteous Mind,3 Jonathan Haidt proposed the moral foundation theory to understand the source of disagreement between people. As put by the author, different people have different sets of moral foundations they subconsciously adhere to when judging controversial issues. The five moral foundations are fairness, care, loyalty, authority, and sanctity. A good example of this is how arguments achieve more agreement when framed with consideration of the target audience’s moral foundations. For example, to reach an agreement on the topic of military funding with an audience whose morals are founded on loyalty and authority, an argument that frames the military as the protector of the nation would be more effective than one depicting the military as a means for overcoming poverty and inequality. Full disclosure on intentions is essential here, so that effective argumentation is not mistaken for manipulation. To build argumentation technologies that bring people together, we should openly and considerately take the audience’s beliefs into account through the process.

Numerous studies have proved the applicability of Haidt’s theory in understanding people’s behaviors and decisions in daily life and found an apparent effect of these morals on the language and rhetoric used by various news portals. Feinberg and Willer4 demonstrated how arguments achieve more agreement when framed with consideration of the target audience’s moral foundations. For example, to reach an agreement on the topic of military funding with an audience whose morals are founded on loyalty and authority, an argument that frames the military as the protector of the nation would be more effective than one depicting the military as a means for overcoming poverty and inequality. Full disclosure on intentions is essential here, so that effective argumentation is not mistaken for manipulation. To build argumentation technologies that bring people together, we should openly and considerately take the audience’s beliefs into account through the process.

Research in computational argumentation has tended to focus solely on analyzing argumentation strategies and their role in persuasiveness. Aside from the work of Durmus and Cardie5 who showed that prior beliefs affect persuasion, not much work has been done on the study of the target audience. With this motivation, we aim to draw more attention from the computational argumentation research toward integrating the audience’s prior beliefs into the process of argument generation. For this, we kickstart the process by proposing a framework for this task consisting of two blocks, which we will present in the following along with rising concerns and challenges.
Where to start?
To kickstart the process of audience’s belief integration, two complementary research questions can be explored. The first question is how to model audience’s beliefs from a given representation, this representation being texts, preferences, etc. The second question is how to encode a model of beliefs into argumentative texts. In the following sections, we will touch on each of these in detail.

Modeling audience’s beliefs
Modeling audience’s beliefs could be of different granularity levels from too general, including religion or political ideology (liberal/conservative), to a particular set of stances/opinions on controversial issues. For example, in our experiments, we represent beliefs as a simple bag-of-words model, learned from stances on popular issues by aggregating words representing pro and con sides of each issue. A pro-abortion user would likely be pro-choice. Hence, words like “right” and “choice” are candidates to be included in their belief-based bag-of-words. While this model seems straightforward, it is not very practical since such information about users’ stances is not always available. Another promising model of beliefs can be based on the moral foundation theory. Here, the audience can be mapped into a latent space in which equivalent moral language can be used to better reach them. For this, we can learn a lot from the research on the moral foundation theory. We emphasize here on the importance of exploring and learning from the social psychology field. Similar to a recent paper, mining moral foundations from argumentative texts is a good start.

To empirically model the relation between beliefs and argumentative texts, annotated data is an essential ingredient. Obtaining such data is a challenging task, and debate portals can be used for this purpose. For example, we can use Debate.org, a platform that enables users to engage in debates and build public profiles where they can state their political ideology, religion, and vote their stance (pro/con) on popular controversial issues. Still, moral annotations for argumentative texts are done on a small scale. Bigger annotated corpora are required to build reliable models.

Here, privacy concerns stemming from the need to collect information about the audience must be addressed. General frameworks for protecting user’s privacy exist and must be maintained. Most importantly, the audience should be informed and made aware of how their data are being used to build models of beliefs. Generating conditioned argumentative texts
To generate argumentative texts on a specific topic matching certain target beliefs, we need a way to model argumentative language and a mechanism to condition this language on the given target beliefs. To model and generate argumentative language, several studies in the computational argumentation field proposed various approaches, covering the generation of simple argumentative units to full arguments. However, producing coherent argumentative text remains an open research field with its own challenges. As a start, focusing on generating single-unit argumentative claims could be an option. In our recent paper, we made use of transfer-learning approaches in text generation, in which we fine-tuned a general transformer-based language model (GPT-2) on a large corpus of arguments crawled from debate portals. By triggering this fine-tuned language model with a prompt representing a controversial topic, we can generate an argumentative text on it.

To condition the generation of argumentative text on the target audience’s beliefs, research on conditional text generation is to be considered. Several approaches have been proposed to guide generation models to produce texts that satisfy certain properties. Among them, for example, is the algorithm of Dathathri et al., which doesn’t require extra training and fits the bag-of-words representation of beliefs. This algorithm steers the generation model to produce text that is similar to a certain vocabulary distribution. In our experiments, the generation model is the fine-tuned argumentative language model and the vocabulary distribution represents the target audience’s beliefs.

Aiming to tune arguments toward a target audience can be thought of as a manipulation attempt. However, changing others’ minds is considered manipulation only when it is done deceptively. To avoid manipulation, future work should ensure a level of transparency by highlighting the belief aspects that the argument is adapted toward.

Final words
We would like to conclude with an envisioned scenario of the proposed technology. That is, given two disagreeing parties on a controversial topic, the technology builds belief models of both parties, finds common ground aspects that the two models share, and then synthesizes arguments adjusted toward these shared aspects. We expect these arguments would then challenge both opposing sides and reduce the disagreement gap. To reach this goal, we call on more research efforts to be allocated toward studying the role of the audience in argument generation. We highlighted two research areas to be explored.

The first is modeling the users’ beliefs. Here, learning more about the source of disagreement from the field of social psychology, like the moral foundation theory, could help. The second area is belief-controlled argument generation. For this, the text generation field and specifically controlled text generation techniques provide a good start to explore the potential of encoding beliefs into argumentative texts.

REFERENCES
1. Slonim, N., Bilu, Y., Alzate, C., Bar-Haim, R., Bogin, B., Bonin, F., Choshen, L., Cohen-Karlik, E., Dankin, L., Edelstein, L., et al. (2021). An autonomous debating system. Nature 597, 378–384.
2. Wachsmuth, H., Pothast, M., Al-Khattib, K., Ajjour, Y., Puschmann, J., Ou, J., Dorsch, J., Morari, V., Bevendorff, J., and Stein, B. (2017). Building an argument search engine for the web. In Proceedings of the 4th Workshop on Argument Mining, 49–59.
3. Hauk, J. (2012). The righteous mind: Why good people are divided by politics and religion (Vintage).
4. Feinberg, M., and Willer, R. (2015). From gulf to bridge: when do moral arguments facilitate political influence? Pers. Soc. Psychol. Bull. 41, 1665–1681.
5. Durmus, E. and Cardie, C. (2018). Exploring the role of prior beliefs for argument persuasion. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), 1035–1045.
6. Alshomary, M., Chen, W.-F., Gurcke, T., and Wachsmuth, H. (2021). Belief-based generation of argumentative claims. In Proceedings of the 43rd annual European Conference on Information Retrieval Research.
7. Kobbé, J., Rebein, I., Hulpus, I., and Stucken Schmidt, H. (2020). Exploring morality in argumentation. In Proceedings of the 7th Workshop on Argument Mining (Association for Computational Linguistics). pp. 30–40. https://www.aclweb.org/anthology/2020.argmining-1.4.
8. Guo, B., Wang, H., Ding, Y., Wu, W., Hao, S., Sun, Y., and Yu, Z. (2021). Conditional text generation for harmonious human-machine interaction. ACM Trans. Intell. Syst. Technol. 12, 1–50.
9. Dathathri, S., Madotto, A., Lan, J., Hung, J., Frank, E., Molino, P., Yosinski, J., and Liu, R. (2020). Plug and play language models: A simple approach to controlled text generation. In International Conference on Learning Representations https://openreview.net/forum?id=H1edEyBKDS.

About the authors

Milad Alshomary is a PhD student and research assistant in the computational social science department at Paderborn University in Paderborn, Germany. He graduated from Damascus University and finished his master’s in Bauhaus University. In the realm of computational argumentation, Milad works on computationally modeling argument generation and on how to synthesize arguments in natural language texts.

Henning Wachsmuth is a junior professor at Paderborn University in Paderborn, Germany. He finished his PhD in Paderborn in 2015 before he worked as a postdoc at Bauhaus-Universität Weimar. Since 2018, Henning has led the Computational Social Science Group at the department of computer science in Paderborn. Henning’s group studies how intentions and views of people are reflected in their language and how machines can understand and imitate this with natural language-processing methods. One main research focus is on computational argumentation. The underlying goal is to support self-determined opinion formation in times of fake news and alternative facts through technology.