The Moral Debater:

A Study on the Computational Generation of Morally Framed Arguments

Authors:

Milad Alshomary, Timon Gurcke and Henning Wachsmuth

Paderborn University, Paderborn, Germany

Roxanne El Baff

German Aerospace Center, Oberpfaffenhofen, Germany

The Problem

The issue that the paper addresses are that an audience's prior beliefs and morals are rarely accounted for in argumentation technology, as knowing what a user believes in can help them come to a mutual understanding with someone of differing beliefs. Understanding these prior beliefs and the role they play in the life of an individual can help people craft more effective arguments when targeting a particular audience. The paper studies the feasibility of generating morally framed arguments automatically, as well as the effect of said arguments on different audience groups, specifically liberals and conservatives.

Prior Works

Research has been conducted to study prior beliefs in argumentation, but with limited reach. There is an entire branch of psychology, known as social psychology, that employs the notion of morals to understand people's judgments on controversial topics. Nearly all studies and experiments concluded that when an argument matches the morals of the target audience, they become significantly more effective than if the argument is misaligned with the morals of the target audience.

Works have also been done in computational linguistics where the persuasive effectiveness of arguments depending on the target audience was closely analyzed. These works showed that audience-based features were able to reliably and consistently predict and augment effectiveness on an audience.

These works were analyzed in different ways, ranging from the interests and personality traits of a person to their stances on popular hot-button issues. Authors of the latter used the stances to influence generation of the argumentative texts, but failed to assess the effectiveness of the texts on the audience specifically, thereby leaving the importance of encoding beliefs unclear.

Aside from these bodies of work and separate social psychology experiments, little work has been done on generating arguments that are tailored toward a specific audience, and even less on the importance of the audience's morals as a factor in achieving agreement.
Also previously known and researched was the Moral Foundation Theory and the general classification of liberals and conservatives into its five projections. According to the theory, humans subconsciously adhere to five basic moral foundations when judging controversial issues: fairness (importance of justice, rights, and equality), care (being kind, and avoiding harm), loyalty (self-sacrifice, solidarity, belongingness), authority (respect to traditions and hierarchy), and purity (sacredness of religion and human). Based on this theory, the disagreement between liberals and conservatives can be explained by the moral gap between the two parties. While liberals rely mainly on care and fairness (individualizing morals) in their assessment of controversial issues, conservatives consider all moral foundations more evenly, somewhat skewed towards the binding morals though, that is, loyalty, authority, and purity

Unique Contributions

This work studies the feasibility of computationally generating morally framed arguments and the effect these arguments have on various audiences. This work used the Moral Foundation Theory, which breaks down morals into five foundations: care, fairness, loyalty, authority, and purity. These foundations were used as inputs to the program on top of using the capabilities of the “Project Debater” to generate high-quality arguments that compete with human arguments. Building on this existing technology helped to make the project replicable and would allow for the full focus to be on the evaluation of the impact of morally framed arguments.
The extended system takes a controversial topic, a stance, and a set of morals as input. It then retrieves a set of argumentative texts, filters the ones conveying the input morals,
and finally phrases an argument holding the given stance on the given topic focusing on the given morals.

Evaluation of the Experiment

There was an evaluation done by the authors at each step of the process when conducting the experiments. For the first step, during the query generation and sentence retrieval, the Project Debater index was used, compiling 400 million news articles which were all split into sentences and indexed with several meta-annotations. These retrieved sentences were first automatically checked to ensure they were annotated with both sentiment or causality markers, then were manually reviewed by a couple of the paper’s authors to ensure their quality. The next step was the moral tagging and filtering as well as component extraction, where the trained classifier was used to annotate each of the argumentative sentences, and discarded any sentences that either had no moral or contained a moral not given as input. Then, during the input aggregation and narrative generation, the narrative generation identified the stance of the claims towards the topic and only kept the arguments that matched the input stance. Redundant elements were also filtered out at this stage.

Along with these internal and manual quality checks to ensure only relevant and high-quality data was being used, the researchers also conducted an external study to ensure the results would remain in line with what they had concluded. The results did line up with their conclusion. Another key point that was brought up was the size of the experiment. There were a total of six individuals, 3 conservatives and 3 liberals, that the experiment was done on. The small sample size could be influencing the results in some way.

Stats

Names with the number of citations:

Milad Alshomary: 10
Timon Gurcke: 4
Henning Wachsmuth: 23
Roxanne El Baff: 1

Many of these authors' works are in argument generation and analysis, with a narrow focus on biases and explanations. A couple of them have several other papers with each other as well on the subject.