

Analyzing and Detecting Persuasion in an Online Discussion Forum

Christopher Hidey, Elena Musi, Alyssa Hwang, Smaranda Muresan,
Kathleen McKeown

Columbia University

July 26, 2019



Goal: Understanding aspects of persuasion that occur at the intersection of discourse and conceptual form

Research questions:

- Can we empirically validate an annotation scheme for conceptual aspects of argumentation?
- What relationship exists between the semantic type of argumentative components and persuasion?

Change My View

- Subreddit from reddit.com where users post an opinion on a topic and other users attempt to change their view on that topic.

Change My View

- Subreddit from reddit.com where users post an opinion on a topic and other users attempt to change their view on that topic.
- The subject of the posts are always of the form “CMV: opinion” and if their view is changed the original poster will indicate this with a delta character: Δ .

Change My View

- Subreddit from reddit.com where users post an opinion on a topic and other users attempt to change their view on that topic.
- The subject of the posts are always of the form “CMV: opinion” and if their view is changed the original poster will indicate this with a delta character: Δ .
- Tan et al. (2016) mined this forum and obtained 18,363 discussions from 12,351 unique original posters.

Data

A

CMV: Patriotism is the belief that being born on one side of a line makes you better

...



I would define patriotism quite simply as supporting one's country, but not *necessarily* disparaging others

B

...

Someone who assists another country that is in worse shape instead of assisting their own can still be a patriot, but also recognize significant need in other nations and decide to assist them as well



A

This is true, but, I think, supporting the common good is also more important than supporting your country



B

Yes, but the two are often one the same, especially when you live in a country as large as the U.S. most acts which serve the common good generally support your country.

- Persuasion
 - predicting persuasiveness of comments in CMV (Tan et al, 2016)
 - ranking arguments in CMV according to “karma” scores (Wei et al., 2016)
 - identifying influencers in social media using agreement, dialog patterns, credibility (Rosenthal and McKeown, 2017)

- Persuasion
 - predicting persuasiveness of comments in CMV (Tan et al, 2016)
 - ranking arguments in CMV according to “karma” scores (Wei et al., 2016)
 - identifying influencers in social media using agreement, dialog patterns, credibility (Rosenthal and McKeown, 2017)
- Argumentation
 - identifying argumentative components (claims and premises) and relations (support/attack) (Stab and Gurevych, 2016; Peldszus and Stede, 2016)
 - predicting persuasiveness of essays using argumentative components and relations (Ghosh et al., 2016)

- Argumentation (continued)
 - Semantics
 - emotion and logic (Habernal and Gurevych, 2017; Walker et al., 2012; Das et al., 2016; Lukin et al., 2017)
 - situation entity types (Becker et al., 2016)
 - subjectivity and objectivity (Park et al., 2015)
 - Pragmatics
 - inference anchoring theory (Budzynska et al., 2014; Budzynska and Reed, 2011)

Goal: Understanding aspects of persuasion that occur at the intersection of discourse and conceptual form

Research questions:

- **Can we empirically validate an annotation scheme for conceptual aspects of argumentation?**
- What relationship exists between the semantic type of argumentative components and persuasion?

Two-stage annotation:

- 1 Label propositions as claims and premises
- 2 Label semantic types
 - claims - one of interpretation, evaluation (emotional or rational), and agreement/disagreement
 - premises - one or more of pathos, ethos, or logos

Claim - stance expressed by a speaker

Based on Freeman's classification of contingent statements

Claim - stance expressed by a speaker

- **interpretation:** prediction or explanation of state of affairs (*I think he will win the election.*)

Based on Freeman's classification of contingent statements

Claim - stance expressed by a speaker

- **interpretation:** prediction or explanation of state of affairs (*I think he will win the election.*)
- **evaluation:** positive or negative judgment. Sub-classified as:

Based on Freeman's classification of contingent statements

Analysis

Claim - stance expressed by a speaker

- **interpretation:** prediction or explanation of state of affairs (*I think he will win the election.*)
- **evaluation:** positive or negative judgment. Sub-classified as:
 - **evaluation-rational:** opinion based on evidence or credible sources (*Her political program is very solid.* or *She is a very smart student.*)

Based on Freeman's classification of contingent statements

Claim - stance expressed by a speaker

- **interpretation**: prediction or explanation of state of affairs (*I think he will win the election.*)
- **evaluation**: positive or negative judgment. Sub-classified as:
 - **evaluation-rational**: opinion based on evidence or credible sources (*Her political program is very solid.* or *She is a very smart student.*)
 - **evaluation-emotional**: opinion based on emotional reasons and/or subjective beliefs (*Going to the gym is an unpleasant activity.* or *I do not like doing yoga.*)

Based on Freeman's classification of contingent statements

Claim - stance expressed by a speaker

- **interpretation:** prediction or explanation of state of affairs (*I think he will win the election.*)
- **evaluation:** positive or negative judgment. Sub-classified as:
 - **evaluation-rational:** opinion based on evidence or credible sources (*Her political program is very solid.* or *She is a very smart student.*)
 - **evaluation-emotional:** opinion based on emotional reasons and/or subjective beliefs (*Going to the gym is an unpleasant activity.* or *I do not like doing yoga.*)
- **agreement or disagreement:** beliefs shared/not shared with another speaker (*I agree that going to the gym is boring.* or *I do not like your ideas.*)

Based on Freeman's classification of contingent statements

Premises - justifications in support of a claim

Based on Aristotle's modes of persuasion

Premises - justifications in support of a claim

- **logos**: appeals to the use of reason, including providing examples and facts (*He will probably win the election. He is the favorite according to the polls.*)

Based on Aristotle's modes of persuasion

Premises - justifications in support of a claim

- **logos**: appeals to the use of reason, including providing examples and facts (*He will probably win the election. He is the favorite according to the polls.*)
- **pathos**: appeals to emotions or personal experience (*You should put comfy furniture into your place. The feeling of being home is unforgettable*)

Based on Aristotle's modes of persuasion

Premises - justifications in support of a claim

- **logos**: appeals to the use of reason, including providing examples and facts (*He will probably win the election. He is the favorite according to the polls.*)
- **pathos**: appeals to emotions or personal experience (*You should put comfy furniture into your place. The feeling of being home is unforgettable*)
- **ethos**: appeals to the credibility established by personal experience/expertise (*I assure you the consequences of fracking are terrible. I have been living next to a pipeline since I was a child.*)

Based on Aristotle's modes of persuasion

Analysis

A

CMV: Patriotism is the belief that being born on one side of a line makes you better

...



[I would define patriotism quite simply as supporting one's country, but not *necessarily* disparaging others] CLAIM_{DISAGREEMENT}

B

...

[Someone who assists another country that is in worse shape instead of assisting their own can still be a patriot, but also recognize significant need in other nations and decide to assist them as well] PREMISE_{LOGOS/PATHOS}



A

[This is true]CLAIM_{AGREEMENT}, but, [I think, supporting the common good is also more important than supporting your country]CLAIM_{RATIONAL EVALUATION}



B

[Yes]CLAIM_{AGREEMENT}, but [the two are often one the same]CLAIM_{INTERPRETATION}, [especially when you live in a country as large as the U.S. most acts which serve the common good generally support your country]PREMISE_{LOGOS}.

Two-stage annotation:

- 1 Label propositions as claims and premises using *expert* annotators
- 2 Label claims (choose one of five) and premises (multi-label) for their semantic types using *crowdsourcing* with Amazon Mechanical Turk

Two-stage annotation:

- 78 discussion threads comprising 278 turns of dialogue with 2615 propositions in 2148 sentences
- 36.5% of sentences contain a claim, 49.7% contain a premise, 22% do not contain either
- 15.8% of sentences contain rational evaluations, 2.3% contain disagreement and 44% contain logos, 3% contain ethos

Analysis

Two-stage annotation:

- 78 discussion threads comprising 278 turns of dialogue with 2615 propositions in 2148 sentences
- 36.5% of sentences contain a claim, 49.7% contain a premise, 22% do not contain either
- 15.8% of sentences contain rational evaluations, 2.3% contain disagreement and 44% contain logos, 3% contain ethos

Task	Kripendorff's α
Claims	0.63
Premises	0.65
Claim Types	0.46
Premise Types	0.73

Annotator Confusion

- The problem isnt always bad parenting, though that can play a role, the problem is a black and white educational system (Gold: **Interpretation** Annotation: **Evaluation-Rational**)
- This is the best argument I have seen (**Agreement** or **Evaluation-Rational?**)

Goal: Understanding aspects of persuasion that occur at the intersection of discourse and conceptual form

Research questions:

- Can we empirically validate an annotation scheme for conceptual aspects of argumentation?
- **What relationship exists between the semantic type of argumentative components and persuasion?**

- Transitions - sequences of claim/premise semantic types
- Claims/Premises - pairs of premises and their most recent claim
- Persuasion - features of winning arguments

- Transitions
 - rational evaluations, agreements, and interpretations (claims) appear at the start of the post
 - pathos and unannotated sentences appear at the end

$p < 0.01$ by chi-squared test of independence with Bonferroni correction

Analysis

- Transitions
 - rational evaluations, agreements, and interpretations (claims) appear at the start of the post
 - pathos and unannotated sentences appear at the end
- Claims/Premises
 - logos follows rational evaluations and pathos follows emotional evaluations

$p < 0.01$ by chi-squared test of independence with Bonferroni correction

Analysis

- Transitions
 - rational evaluations, agreements, and interpretations (claims) appear at the start of the post
 - pathos and unannotated sentences appear at the end
- Claims/Premises
 - logos follows rational evaluations and pathos follows emotional evaluations
- Persuasion
 - pathos and logos are more likely to be used jointly
 - agreements are more likely to be used at the beginning
 - winning arguments are more likely to have consecutive premises of the same type

$p < 0.01$ by chi-squared test of independence with Bonferroni correction

Examples:

- Enjoying the moment is possible, but doesn't make life have a point. For once I die, all memories and all point is gone. (**Evaluation-Emotional** and **Pathos**)
- I do agree that today's moderates are potentially tomorrow's conservatives. However this isn't about being just a bit conservative. (**Agreement** and **Interpretation**)

Ongoing Work

- Label arguments for support/attack relations between claims and premises
- Build models to predict claims and premises and their semantic types
- Build models that incorporate claim and premise types for persuasiveness prediction
- Generate structured dialogues with argumentative components

Conclusions

- Moderate agreement for argumentative components and the semantic type of premises but not for claim types
- Sequences of semantic components are correlated
- Winning arguments show significant differences in semantic types compared to non-winning arguments

Questions

?

Thanks to DARPA-DEFT, SNFS, and all the annotators!