Back Up Your Stance: Recognizing Arguments in Online Discussions

Filip Boltužić and Jan Šnajder

Text Analysis and Knowledge Engineering Lab
FER, University of Zagreb

First Workshop on Argumentation Mining
ACL 2014
Baltimore, Maryland
26 June 2014
Should gay marriage be legal?

User comment 1
Gay marriages must be legal in all 50 states. 2 people regardless of their genders. Discrimination against gay marriage is unconstitutional and biased. Tolerance, education and social justice make our world a better place.

User comment 2
Absolutely No. Who are we to rewrite the creator of this world’s view on what marriage is? They deserve the civil union and employment security laws, but rewriting the definition of marriage is going too far!
Online discussions are a **valuable source of opinions**: Comments on news stories, social networks, blogs, discussion forums,...

**Relevant for:**
Political opinion mining, sociological studies, brand analysis,...
To really leverage this ocean of opinions, we should be able to answer the *whys of opinions*
User opinions are often backed up with arguments. Argument-Based Opinion Mining involves determining the arguments on which users base their stance.
# Task Description

## Argument Recognition

Identifying what arguments, from a set of predefined arguments, are used in a comment, and how.

**Input:**
1. Prominent arguments from past debates
2. Noisy comments from current on-line discussions

**Output:**
1. Is an argument used in a comment?
2. Does the comment support or attack the given argument?
Should gay marriage be legal?

Comment
Gay marriages must be legal in all 50 states. 2 people regardless of their genders. Discrimination against gay marriage is unconstitutional and biased. Tolerance, education and social justice make our world a better place.

-supported argument
It is discriminatory to refuse gay couples the right to marry

-attacked argument
Marriage should be between a man and a woman.
Comment

Absolutely No. Who are we to rewrite the creator of this world's view on what marriage is? They deserve the civil union and employment security laws, but rewriting the definition of marriage is going too far!

- Supported argument

Gay couples can declare their union without resort to marriage.

- Supported argument

Gay couples should be able to take advantage of the fiscal and legal benefits of marriage.

- Supported argument

Marriage should be between a man and a woman.
Related Work

- Argumentation mining [Palau and Moens, 2009]
  - Argument identification
  - Argument proposition classification
  - Argumentative parsing
- Argumentation networks [Cabrio and Villata, 2013]
  - Textual inference (support/attack relations)
  - Computation of acceptable arguments (debate overview)
- Stance classification
  - Stance on forum posts [Anand et al., 2011]
  - Support/opposition user groups [Murakami and Raymond, 2010]
- Opinion mining + Argumentation mining
  [Hogenboom et al., 2010, Grosse et al., 2012, Wyner and Schneider, 2012, Villalba and Saint-Dizier, 2012, Chesñevar et al., 2013]
We do not aim to extract the argumentation structure (within a comment nor between comments in a discussion)

Challenges:

1. Noisy input
2. Users’ arguments are often informal, ambiguous, vague, implicit, and poorly worded
3. Comment may contain several arguments as well non-argumentative text
Outline

1. Corpus of Comment-Argument Pairs
2. Argument Recognition Model
3. Evaluation
Outline

1. Corpus of Comment-Argument Pairs
2. Argument Recognition Model
3. Evaluation
**ComArg** Corpus

**ComArg**: Corpus of comments, arguments, and manually annotated comment–argument pairs

<table>
<thead>
<tr>
<th>Comment (Pro/Con)</th>
<th>Argument (Pro/Con)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Online discussions (procon.org)</td>
<td>Past debates (idebate.org)</td>
</tr>
<tr>
<td>(2) Should Gay Marriage Be Legal?</td>
<td>This house would allow gay couples to marry</td>
</tr>
<tr>
<td>Should the Words &quot;under God&quot; be in the US Pledge of Allegiance?</td>
<td>This house would remove the words &quot;under God&quot; from the American Pledge of Allegiance</td>
</tr>
<tr>
<td>(3) Manual spam filtering</td>
<td>Manually paraphrased</td>
</tr>
</tbody>
</table>
### COMARG Statistics

<table>
<thead>
<tr>
<th></th>
<th>Under God in Pledge (UGIP)</th>
<th>Gay Marriages (GM)</th>
</tr>
</thead>
<tbody>
<tr>
<td># Argument</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td># Comment</td>
<td>175</td>
<td>198</td>
</tr>
<tr>
<td># Pair</td>
<td>1,050</td>
<td>1,386</td>
</tr>
</tbody>
</table>
## Arguments for UGIP

<table>
<thead>
<tr>
<th>Argument</th>
<th>Stance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likely to be seen as a state sanctioned condemnation of religion</td>
<td>Pro</td>
</tr>
<tr>
<td>The principles of democracy regulate that the wishes of American</td>
<td>Pro</td>
</tr>
<tr>
<td>Christians, who are a majority are honored</td>
<td></td>
</tr>
<tr>
<td>Under God is part of American tradition and history</td>
<td>Pro</td>
</tr>
<tr>
<td>Implies ultimate power on the part of the state</td>
<td>Con</td>
</tr>
<tr>
<td>Removing under God would promote religious tolerance</td>
<td>Con</td>
</tr>
<tr>
<td>Separation of state and religion</td>
<td>Con</td>
</tr>
</tbody>
</table>
Corpus Annotation

Three annotators labeled 2,436 comment-argument pairs

Five-point scale:
- **A** – comment explicitly attacks the argument
- **a** – comment vaguely/implicitly attacks the argument
- **N** – comment makes no use of the argument
- **s** – comment vaguely/implicitly supports the argument
- **S** – comment explicitly supports the argument
I believe that the statement about God in the pledge should be eliminated. In order to create unity in our nation we shouldn’t be forcing someone else’s God onto people. Also, adding the phrase Under God" was a decision made to widen the gap between us and the Soviet Union. It wasn’t put there to "honor god" or make us any better. Furthermore, we should separate church from state. Its the law.

S (explicitly supported)
Separation of state and religion.

a (vaguely/implicitly attacked)
Under God is part of American tradition and history.

N (not used)
Likely to be seen as a state sanctioned condemnation of religion.
Problematic comment-argument pairs:
1. all three annotators disagree
   OR
2. the ordinal distance between any of the labels is greater than one

- A, a, N
- A, A, s
- A, A, N
- A, A, a

515 problematic items (21%)
Each re-annotated independently by the three annotators
86 revisions
Annotation Statistics

- Average number arguments per comment: 1.9
- Fleiss’/Cohen kappa: 0.49
- Pearson’s r: 0.71

Gold annotation: majority label (3-way disagreements discarded)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>a</th>
<th>N</th>
<th>s</th>
<th>S</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td># Pair</td>
<td>137</td>
<td>159</td>
<td>1,540</td>
<td>156</td>
<td>306</td>
<td>2,298</td>
</tr>
<tr>
<td>%</td>
<td>5.96</td>
<td>6.92</td>
<td>67.0</td>
<td>6.79</td>
<td>13.3</td>
<td>100</td>
</tr>
</tbody>
</table>
1. Corpus of Comment-Argument Pairs
2. Argument Recognition Model
3. Evaluation
Argument Recognition framed as multiclass classification

Features:
1. Textual Entailment (TE)
2. Semantic Text Similarity (STS)
3. Stance Alignment (SA)
   - Binary feature: 1 if argument and comment have same stance
Stance Alignment

- **Pro comments:**
  - Usually support Pro arguments
  - May attack Con arguments

- **Con comments:**
  - Usually support Con arguments
  - May attack Pro arguments

- But exceptions are possible:
  - E.g. a Pro comment attacking a Pro argument
Should the Words "under God" be in the US Pledge of Allegiance?

Comment

I am not bothered by "under God" but by the highfalutin christians that do not realize that phrase was NEVER in the original pledge - it was not added until 1954. So stop being so pompous and do not offend my parents and grandparents who NEVER used "under God" when they said the pledge. Let it stay, but know the history of the Cold War and fear of communism.

Attacked argument

Under God is part of American tradition and history.
Textual Entailment (TE) is defined as a *directional relation* between two text fragments, called text (\(T\)) and hypothesis (\(H\)), so that a human being, with common understanding of language and common background knowledge, can infer that \(H\) is most likely true on the basis of the content of \(T\).

**T: Comment**

*Marriage should be between Adam and Eve. NOT Adam and Steve.*

**H: Argument**

*Marriage should be between a man and a woman.*
Textual Entailment: Implementation

- Excitement Open Platform (EOP) [Padó et al., 2013]
  - Seven pre-trained entailment decision algorithms (EDAs)
- Each EDA gives two outputs
  - Decision
  - Confidence
- 14 features
Comment-Argument Entailments

![Bar chart showing ratio of positive entailment decisions (%)]

- A: 0.2
- a: 0.4
- N: 0.6
- s: 0.8
- S: 1.0

26 / 43
Semantic Textual Similarity [Agirre et al., 2012]

Semantic Textual Similarity (STS) measures the degree of semantic equivalence between two texts. STS differs from TE in as much as it assumes symmetric graded equivalence between the pair of textual snippets.

- Outputs real valued score [0,5]
TakeLab Semantic Textual Similarity [Šarić et al., 2012]
  * Two levels
    * Sentence level similarity
      (29-dimensional similarity vector, max, mean)
    * Comment level similarity

32 features
<table>
<thead>
<tr>
<th>Comment</th>
<th>Score: 2.906</th>
<th>Gold label: A</th>
</tr>
</thead>
<tbody>
<tr>
<td>The argument that legalizing gay marriage will destroy traditional religious marriages is a red herring.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Score: 1.969</th>
<th>Gold label: N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gay couples should be able to take advantage of the fiscal and legal benefits of marriage.</td>
<td></td>
</tr>
</tbody>
</table>
Comment-Argument Similarities (scaled)
Tools:
- Baselines – majority class (MCC), Bag of Words Overlap (BoWO)
- SVM with RBF (5×3 cross-validation)

Setups:
- 5-way: A-a-N-s-S
- 3-way: Aa-N-sS
- 3-way: A-N-S
- Within-topic / Combined / Cross-topic
Results: Within-Topic Argument Recognition

Micro-averaged F1-score

<table>
<thead>
<tr>
<th>Model</th>
<th>A-a-N-s-S</th>
<th>Aa-N-sS</th>
<th>A-N-S</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UGIP</td>
<td>GM</td>
<td>UGIP</td>
</tr>
<tr>
<td>MCC baseline</td>
<td>68.2</td>
<td>69.4</td>
<td>68.2</td>
</tr>
<tr>
<td>BoWO baseline</td>
<td>68.2</td>
<td>69.4</td>
<td>67.8</td>
</tr>
<tr>
<td>TE</td>
<td>69.1</td>
<td>81.1</td>
<td>69.6</td>
</tr>
<tr>
<td>STS</td>
<td>67.8</td>
<td>68.7</td>
<td>67.3</td>
</tr>
<tr>
<td>SA</td>
<td>68.2</td>
<td>69.4</td>
<td>68.2</td>
</tr>
<tr>
<td>STS+SA</td>
<td>68.2</td>
<td>69.5</td>
<td>67.5</td>
</tr>
<tr>
<td>TE+SA</td>
<td>68.9</td>
<td>72.4</td>
<td>71.0</td>
</tr>
<tr>
<td>TE+STS+SA</td>
<td><strong>70.5</strong></td>
<td>72.5</td>
<td>68.9</td>
</tr>
</tbody>
</table>

- STS or STS+SA not good
- TE outperforms baseline from 0.6% to 11.7% F1
- TE+SA overall best
- SA helps distinguish entailment/contradiction
## Results: Combined topics

Macro-averaged F1-score

<table>
<thead>
<tr>
<th>Model</th>
<th>A-a-N-s-S</th>
<th>Aa-N-sS</th>
<th>A-N-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCC baseline</td>
<td>68.9</td>
<td>68.9</td>
<td>77.9</td>
</tr>
<tr>
<td>TE+SA</td>
<td>71.1</td>
<td>73.3</td>
<td>81.6</td>
</tr>
<tr>
<td>STS+TE+SA</td>
<td><strong>71.6</strong></td>
<td>71.4</td>
<td>80.4</td>
</tr>
</tbody>
</table>

- STS+TE+SA best on **A-a-N-s-S**
- Slight improvement when discarding vague/implicit cases
COMARG corpus of comments and arguments

Argument Recognition task
- TE-based models reach 70.5–81.8% micro-F1, outperform baseline
- (Marginally affected on unseen topic)

Improvements: Corpus
- Annotation of argumentative segments
- Topic expansion

Improvements: Model
- Linguistically-inspired features
- Argument interactions
- Stance classification
Thanks!

Get the COMARG corpus from:

takelab.fer.hr/comarg


*Natural Language Engineering*, 1(1):000–000.

In *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, pages 98–107. ACM.


**Comment**

*Marriage isn’t the joining of two people who have intentions of raising and nurturing children. It never has been. There have been many married couples who have not had children. (...) If straight couples can attempt to work out a marriage, why can’t homosexual couple have this same privilege?*

**Argument**

*It is discriminatory to refuse gay couples the right to marry.*

- Best model says S, annotators say s
(...)
There are no legal reasons why two homosexual people should not be allowed to marry, only religious ones (...)

Argument
Gay couples should be able to take advantage of the fiscal and legal benefits of marriage.

STS+SA: N ✓
TE+SA: S X