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

- 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,...

The Whys of Opinions

- To really leverage this ocean of opinions, we should be able to answer the **whys of opinions**

Opinions and Arguments

- Users often back up their opinions with **arguments**...

Argument-Based Opinion Mining

Determining the arguments on which the 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.

Should gay marriage be legal?

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]

Argument Recognition?

- 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

Outline

COMARG Corpus

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

	Comment (Pro/Con)	Argument (Pro/Con)
(1)	Online discussions (procon.org)	Past debates (idebate.org)
(2)	*Should Gay Marriage Be Legal?*	*This house would allow gay couples to marry*
	Should the Words "under God" be in the US Pledge of Allegiance?	*This house would remove the words "under God" from the American Pledge of Allegiance*
(3)	Manual spam filtering	Manually paraphrased

	Under God in Pledge **(UGIP)**	Gay Marriages **(GM)**
# Argument	6	7
# Comment	175	198
# Pair	1,050	1,386

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

- 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

Annotation Example

Comment

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 seperate 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.

Annotation Revision

- 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)

	A	**a**	**N**	**s**	**S**	Total
# Pair	137	159	1,540	156	306	2,298
%	5.96	6.92	67.0	6.79	13.3	100

Outline

Features

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

Textual Entailment [Dagan et al., 2006]

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.
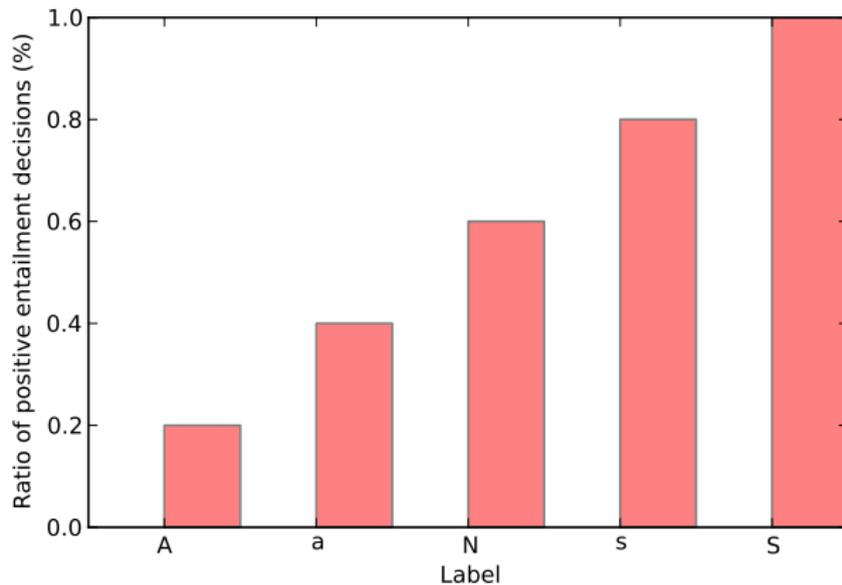
H: Argument

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

Semantic Textual Similarity

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]

Semantic Textual Similarity: Implementation

- TakeLab Semantic Textual Similarity [Šarić et al., 2012]
  - Two levels
    - Sentence level similarity
      (29-dimensional similarity vector, max, mean)
    - Comment level similarity
- 32 features

Semantic Textual Similarity: Example

Comment

The argument that legalizing gay marriage will destroy traditional religious marriages is a red herring.

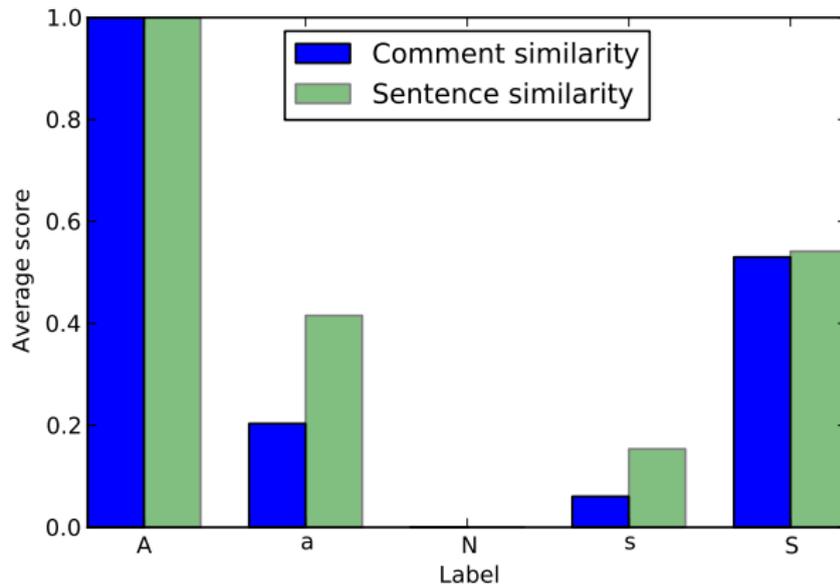## Score: 2.906 — Gold label: **A**

Gay marriage undermines the institution of marriage, leading to an increase in out of wedlock births and divorce rates.

Score: 1.969 — Gold label: **N**

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

Comment-Argument Similarities (scaled)

Evaluation Setup

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

Model	A-a-N-s-S		Aa-N-sS		A-N-S	
	UGIP	GM	UGIP	GM	UGIP	GM
MCC baseline	68.2	69.4	68.2	69.4	79.5	76.6
BoWO baseline	68.2	69.4	67.8	69.5	79.6	76.9
TE	69.1	**81.1**	69.6	72.3	80.1	73.4
STS	67.8	68.7	67.3	69.9	79.2	75.8
SA	68.2	69.4	68.2	69.4	79.5	76.6
STS+SA	68.2	69.5	67.5	68.7	79.6	76.1
TE+SA	68.9	72.4	**71.0**	**73.7**	**81.8**	**80.3**
TE+STS+SA	**70.5**	72.5	68.9	73.4	81.4	79.7

- 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

Model	**A-a-N-s-S**	**Aa-N-sS**	**A-N-S**
MCC baseline	68.9	68.9	77.9
TE+SA	71.1	**73.3**	**81.6**
STS+TE+SA	**71.6**	71.4	80.4

- STS+TE+SA best on **A-a-N-s-S**
- Slight improvement when discarding vague/implicit cases

Wrap Up

- 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

📄 Agirre, E., Diab, M., Cer, D., and Gonzalez-Agirre, A. (2012).
Semeval-2012 task 6: A Pilot on Semantic Textual Similarity.
In *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*, pages 385–393. Association for Computational Linguistics.

📄 Anand, P., Walker, M., Abbott, R., Tree, J. E. F., Bowmani, R., and Minor, M. (2011).
Cats rule and dogs drool!: Classifying stance in online debate.
In *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis*, pages 1–9. Association for Computational Linguistics.

📄 Cabrio, E. and Villata, S. (2013).
A Natural Language Bipolar Argumentation Approach to Support Users in Online Debate Interactions†.
Argument & Computation, 4(3):209–230.

📄 Chesñevar, C. I., González, M. P., Grosse, K., and Maguitman, A. G. (2013).
A first approach to mining opinions as multisets through argumentation.
In *Agreement Technologies*, pages 195–209. Springer.

📄 Dagan, I., Glickman, O., and Magnini, B. (2006).
The pascal recognising textual entailment challenge.
In *Machine learning challenges. evaluating predictive uncertainty, visual object classification, and recognising tectual entailment*, pages 177–190. Springer.

📄 Grosse, K., Chesñevar, C. I., and Maguitman, A. G. (2012).
An argument-based approach to mining opinions from Twitter.
In *AT*, pages 408–422.

📄 Hogenboom, A., Hogenboom, F., Kaymak, U., Wouters, P., and
De Jong, F. (2010).
Mining economic sentiment using argumentation structures.
In *Advances in Conceptual Modeling – Applications and
Challenges*, pages 200–209. Springer.

📄 Murakami, A. and Raymond, R. (2010).
Support or oppose?: Classifying positions in online debates from
reply activities and opinion expressions.
In *Proceedings of the 23rd International Conference on
Computational Linguistics: Posters*, pages 869–875. Association
for Computational Linguistics.

Padó, S., Noh, T., Stern, A., Wang, R., and Zanoli, R. (2013).
Design and Realization of a Modular Architecture for Textual
Entailment.
Natural Language Engineering, 1(1):000–000.

Palau, R. M. and Moens, M.-F. (2009).
Argumentation mining: The Detection, Classification and Structure
of Arguments in Text.
In *Proceedings of the 12th International Conference on Artificial
Intelligence and Law*, pages 98–107. ACM.

Šarić, F., Glavaš, G., Karan, M., Šnajder, J., and Bašić, B. D.
(2012).
Takelab: Systems for Measuring Semantic Text Similarity.
Introduction to SEM 2012*, page 441.

Villalba, M. P. G. and Saint-Dizier, P. (2012).
Some facets of argument mining for opinion analysis.
In *COMMA*, pages 23–34.

Wyner, A. and Schneider, J. (2012).
Arguing from a point of view.
In *AT*, pages 153–167.

Error Analysis: Ex. 1

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 whos 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**

Error Analysis: Ex. 2

Comment

(...) 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** ✗