Language Influence 2: Argumentation

"ANALYZING WIKIPEDIA DELETION DEBATES WITH A GROUP DECISION-MAKING FORECAST MODEL."

BACKGROUND ON WIKIPEDIA

- Articles for Deletion in Wikipedia pages.
- Nomination by voting.
- Conflict between "deletionist" and "inclusionist."

A single AfD discussion Example

- Often debate for post.
- Delete and Keep.
- Keep a long period.

The result was **keep**. Can't sleep, clown will eat me 01:21, 8 October 2007 (UTC)

Missed call [edit]

Missed call (edit | talk | history | links | watch | logs | views) – (View log)

Seems to fail WP:NOT, is essentially social commentary and no references are given for the major assertions presented. Orderinchaos 09:13, 2 October 2007 (UTC)

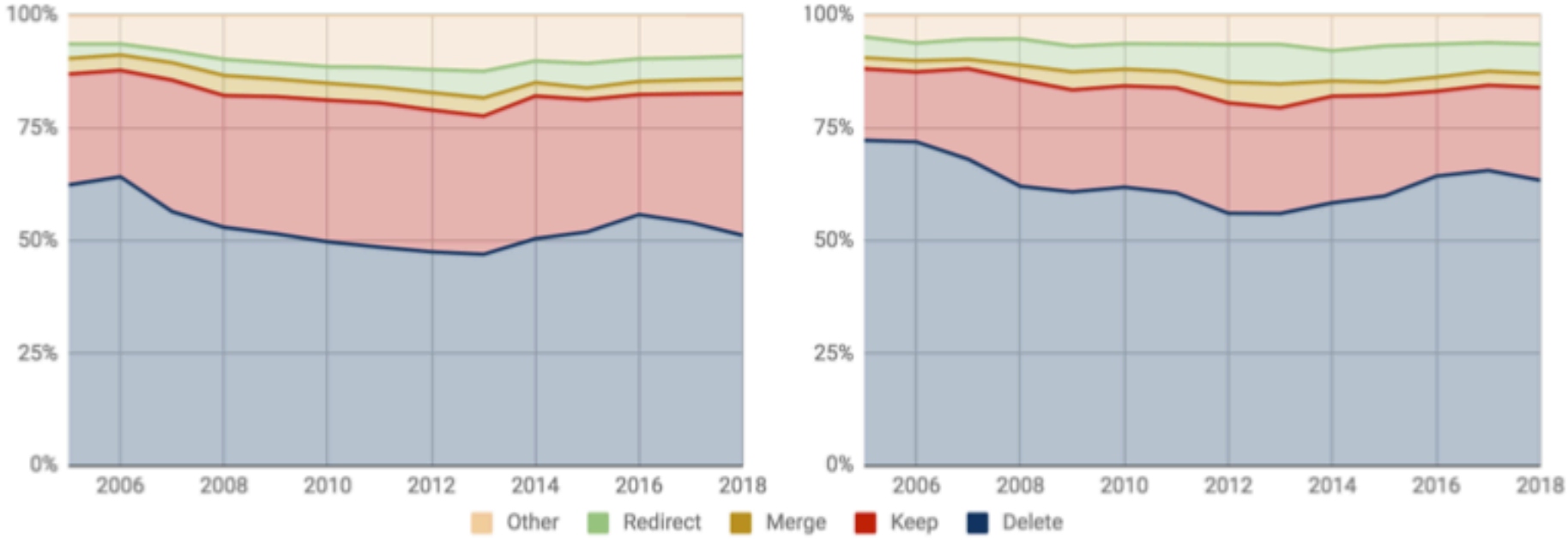
- **Delete** Just a junk article, not notable. Jmlk17 09:52, 2 October 2007 (UTC)
- **Keep**. I don't know guys, this thing is very prevalent in our culture. See [1]. I don't know for other cultures though. --Lenticel (talk) 10:18, 2 October 2007 (UTC)
  - Could possibly be mentioned in an article on Telecommunications in India? Twenty Years 15:50, 6 October 2007 (UTC)
    - Philippines is not India.--Lenticel (talk) 00:35, 8 October 2007 (UTC)
- **Keep** as above. I don't see why we should not keep this. .. Elmao 10:23, 2 October 2007 (UTC)
- **Keep**, there are bazillions of articles on cell phone etiquette out there to source this. I think the money-saving angle is only one part of it. --Dhartung | Talk 11:47, 2 October 2007 (UTC)
  - ...which could easily be covered in the article entitled Telecommunications in India. Twenty Years 15:57, 6 October 2007 (UTC)
- **Keep**, i added enough links to merit inclusion. it is not just a social commentary, it is a business, revenue and profit headache too. the apex body of indian telecom operators, coai has even instituted studies for tracking revenue loss. pls revisit the article to see the new links. Ankur Jain 12:23, 2 October 2007 (UTC)
  - **Comment:** hate to add this, but i believe there is a distinct anglo-american bias to article editing. just because you guys don't know about the widespread use of this thing, probably never having visited india or africa etc., that does not mean it does not exist. there is world beyond your countries.Ankur Jain 12:26, 2 October 2007 (UTC)

Research in Wikipedia deletion debates

- new model for forecasting outcomes of online group decision-making discussions.
- (1) Historical context of administrative tasks.
- (2) Exploratory data analysis in Corpus.
- (3) Forecasting methods
- (4) Results analysis

	Delete	Keep	Merge	Redirect	Other
Votes (2005-2018)	54.9	28.4	3.6	3.8	9.3
Outcomes (2005-2018)	63.9	20.7	3.2	6.0	6.2
Prior Work [60] (2003-2010)	63.6	23.6	3.9	1.9	7.0

OVERALL BREAKDOWNS OF LABELS ACROSS ALL DATA.



DISTRIBUTIONS BY YEAR FOR VOTES (LEFT) AND OUTCOMES (RIGHT) OVER WIKIPEDIA'S HISTORY.

Forecasting Methods

- Natural Language Processing (NLP)
- Machine Learning Predictive Models (ML)

NLP Methods

- Not like bag-of-word traditionally.
- Analysis surroundings context rather than single word.
- Using $BERT_{\mathrm{BASE}}$ models by apply consecutive words.

NLP Methods

$$\phi_d(t_i) = \frac{\sum_{j=0}^{i} \frac{\phi_j}{\ln(len(r_j))}}{i}$$

- Encode overall discussion content
- A single deletion discussion is labeled d.
- timestamp t_i , and a rationale text, r_i (which might be empty).
- The features of a single contribution ci can be extracted using arbitrary representations of language, and represented as ϕi .

ML Methods

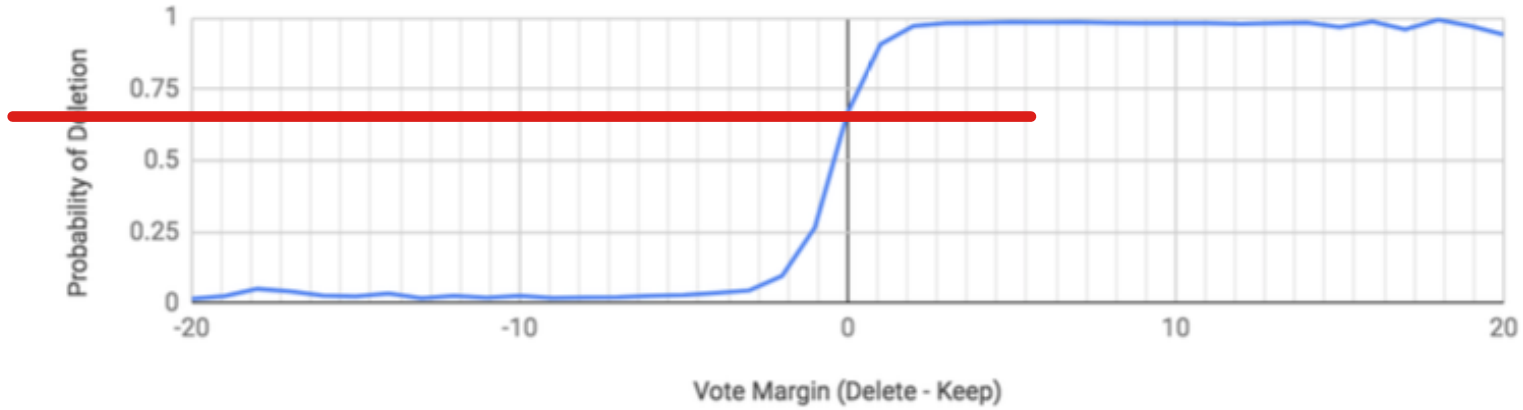
$$\|\mathbf{w}\|_2 = \left(|w_1|^2 + |w_2|^2 + \ldots + |w_N|^2\right)^{\frac{1}{2}}$$

2-norm (also known as L2 norm or Euclidean norm)

$$\text{logistic}(\eta) = \frac{1}{1 + exp(-\eta)}$$

LOGISTIC REGRESSION WITH L2 REGULARIZATION

administrators choose Delete in 66.9% of these cases

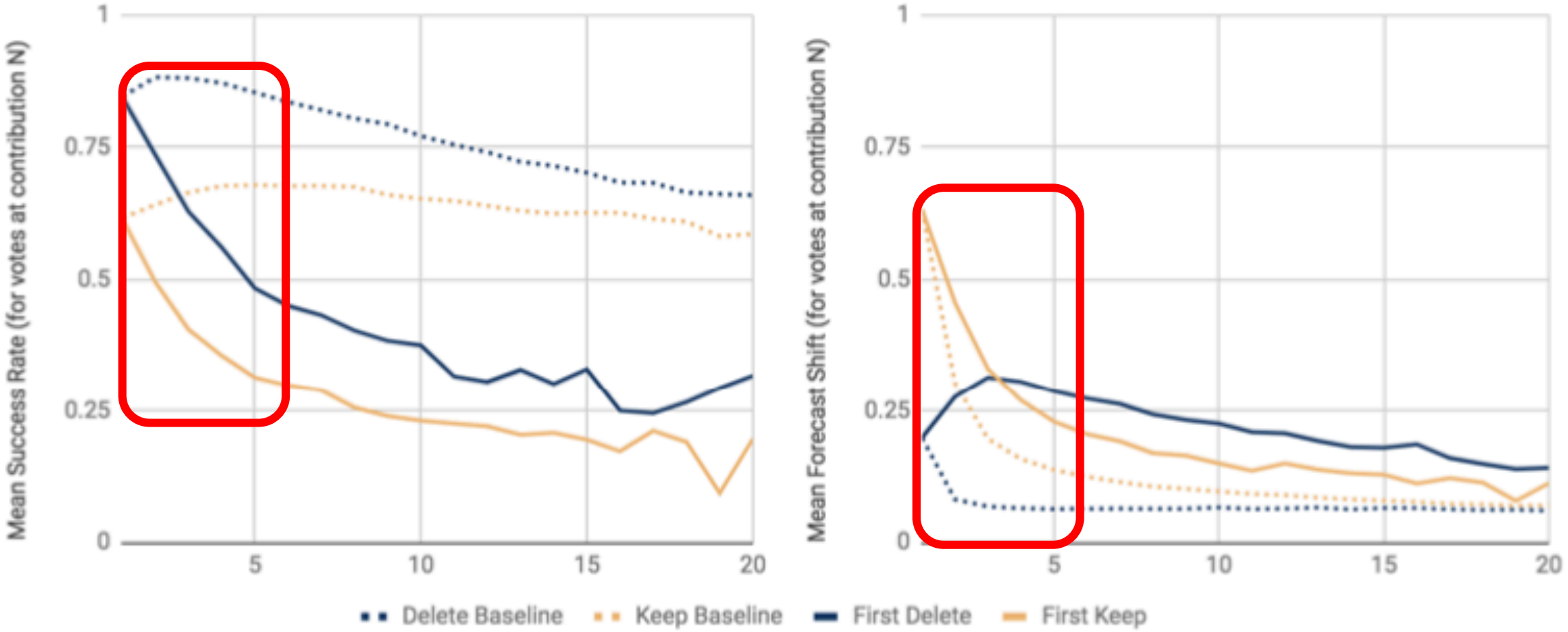


Representation	Full Debate		Incremental	
	%	κ	%	κ
Majority Class Baseline	74.0	0.00	62.1	0.00
GloVe	81.7	0.49	69.1	0.31
Bag-of-Words	84.2	0.58	72.4	0.39
BERT	85.8	0.62	73.4	0.41
BERT + Vote Labels	93.5	0.83	79.7	0.55

FORECASTING LAYOUT

Forecasting models

- 10-fold cross validation (95% training data)
- 5% of the corpus (20,000 discussion) in testing data
- Logistic regression predicting outcomes (D or K)
- Measuring Forecast Shifts in predictive models

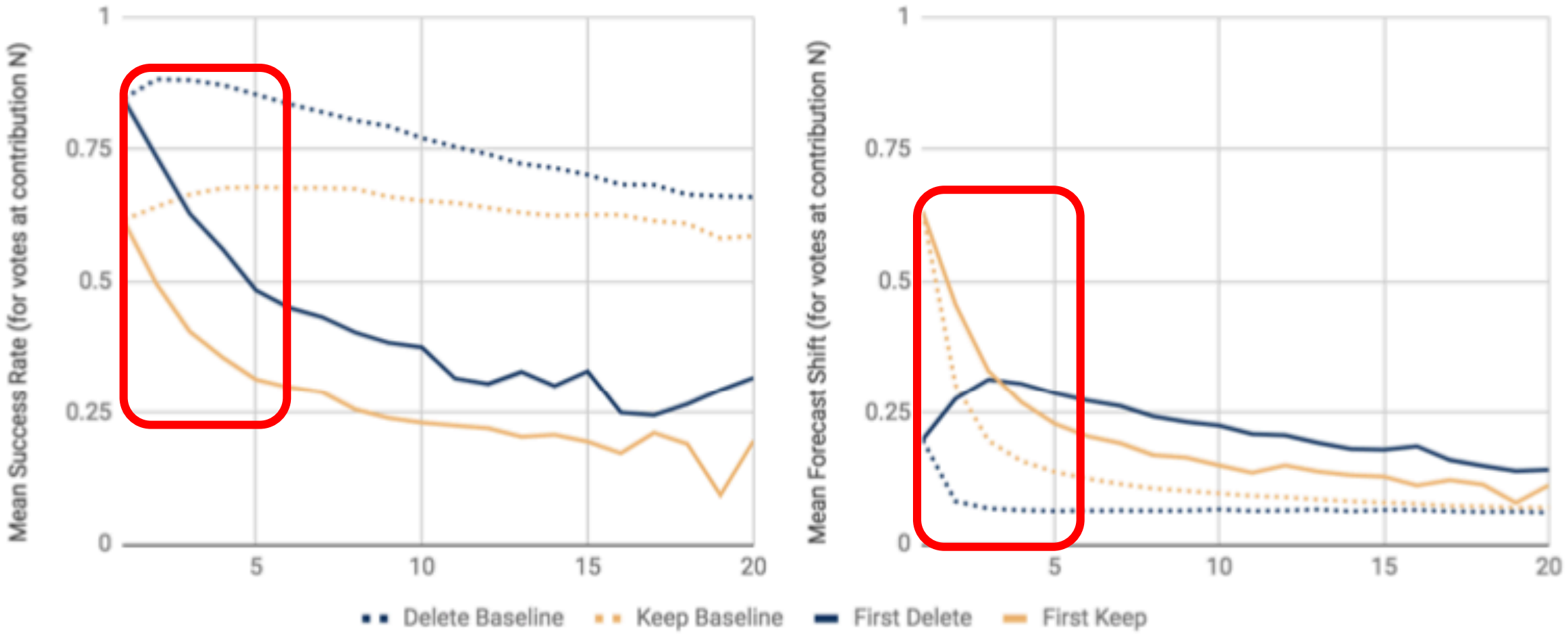

Mean Success Rate (for votes at contribution N)
1
0.75
0.5
0.25
0
5
10
15
20
Mean Forecast Shift (for votes at contribution N)
1
0.75
0.5
0.25
0
5
10
15
20
Delete Baseline
Keep Baseline
First Delete
First Keep

Forecasting Accuracy

- Best-performing model with 97.3% accuracy(less than 5).
- Best-performing model with 85.3% accuracy(higher than 5).
- 1.6% higher than models in other papers

Limitation

- Size of corpus for testing (5%)
- Did NOT include article texts themselves in our work.
- Unclear/insignificantly predictive results in shift change.
- Models only perform well in certain period.
- Fewer discussion in ML methods


Mean Success Rate (for votes at contribution N)
1
0.75
0.5
0.25
0
5
10
15
20
Mean Forecast Shift (for votes at contribution N)
1
0.75
0.5
0.25
0
5
10
15
20
Delete Baseline
Keep Baseline
First Delete
First Keep

Further Discussion

- High forecasting accuracy among other predictive models.
- Both implement NLP and ML algorithm.
- Search and analyze surrounding textural information.
- High reliability and volatility.

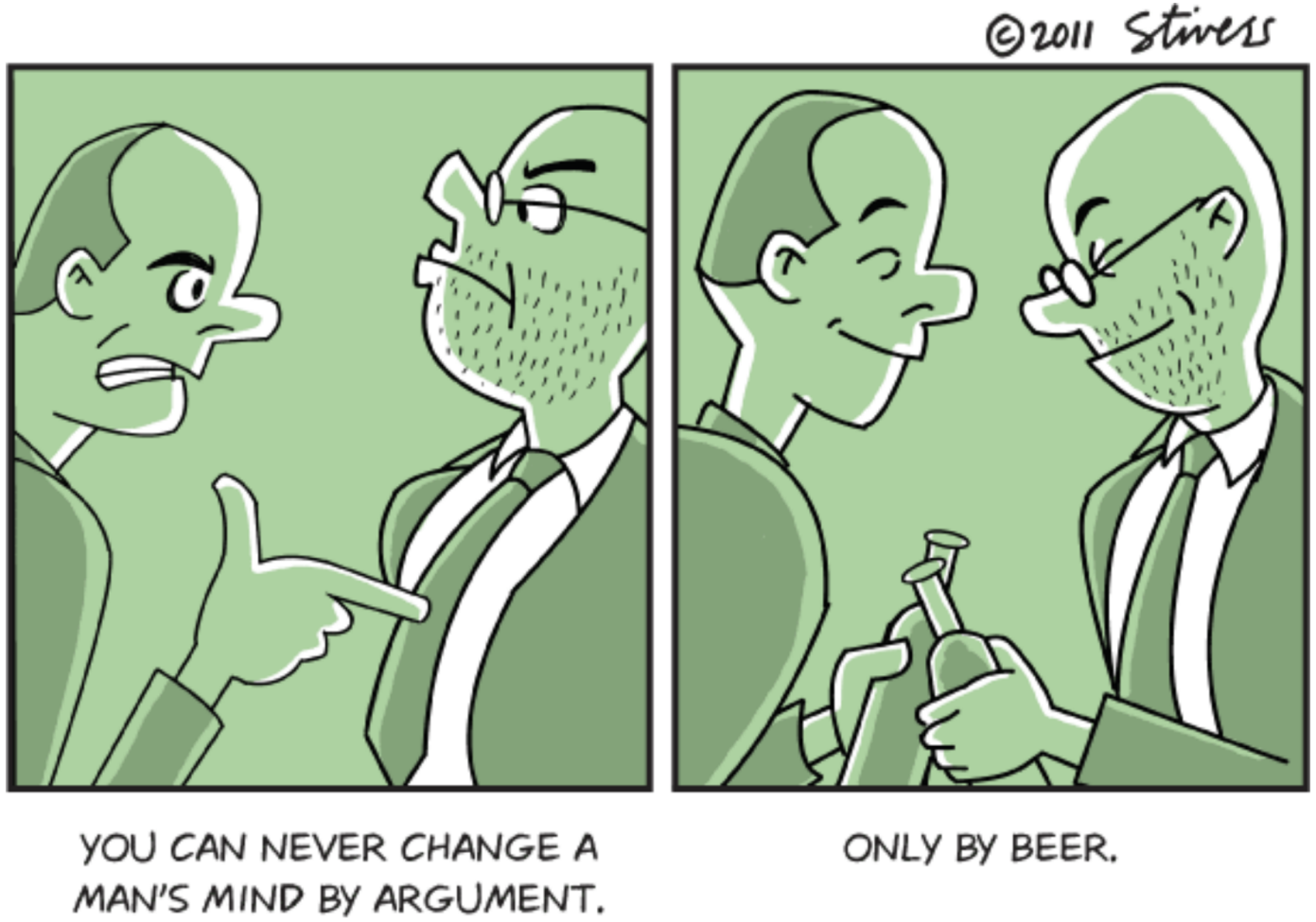
Q&A

- Any Questions for the presentation?

Language Influence 2: Argumentation

"WINNING ARGUMENTS:
INTERACTION DYNAMICS AND
PERSUASION STRATEGIES
IN GOOD-FAITH ONLINE DISCUSSIONS"

We are always trying to change someone else's opinion!



Original Opinion

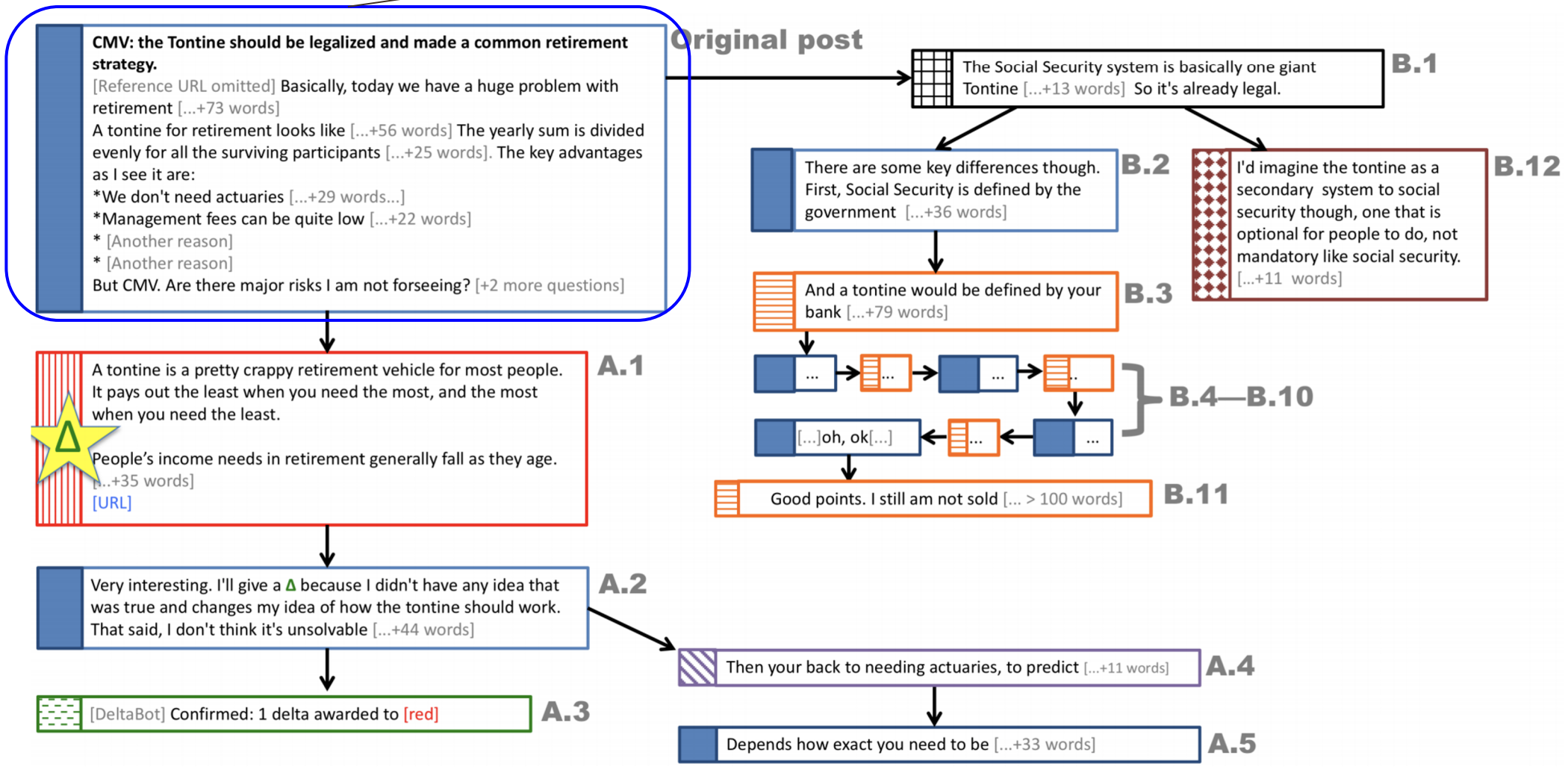
- Certainty of the holder (Pomerantz et al.1995)
- Importance of the belief (Petty et al. 1997)

Argument

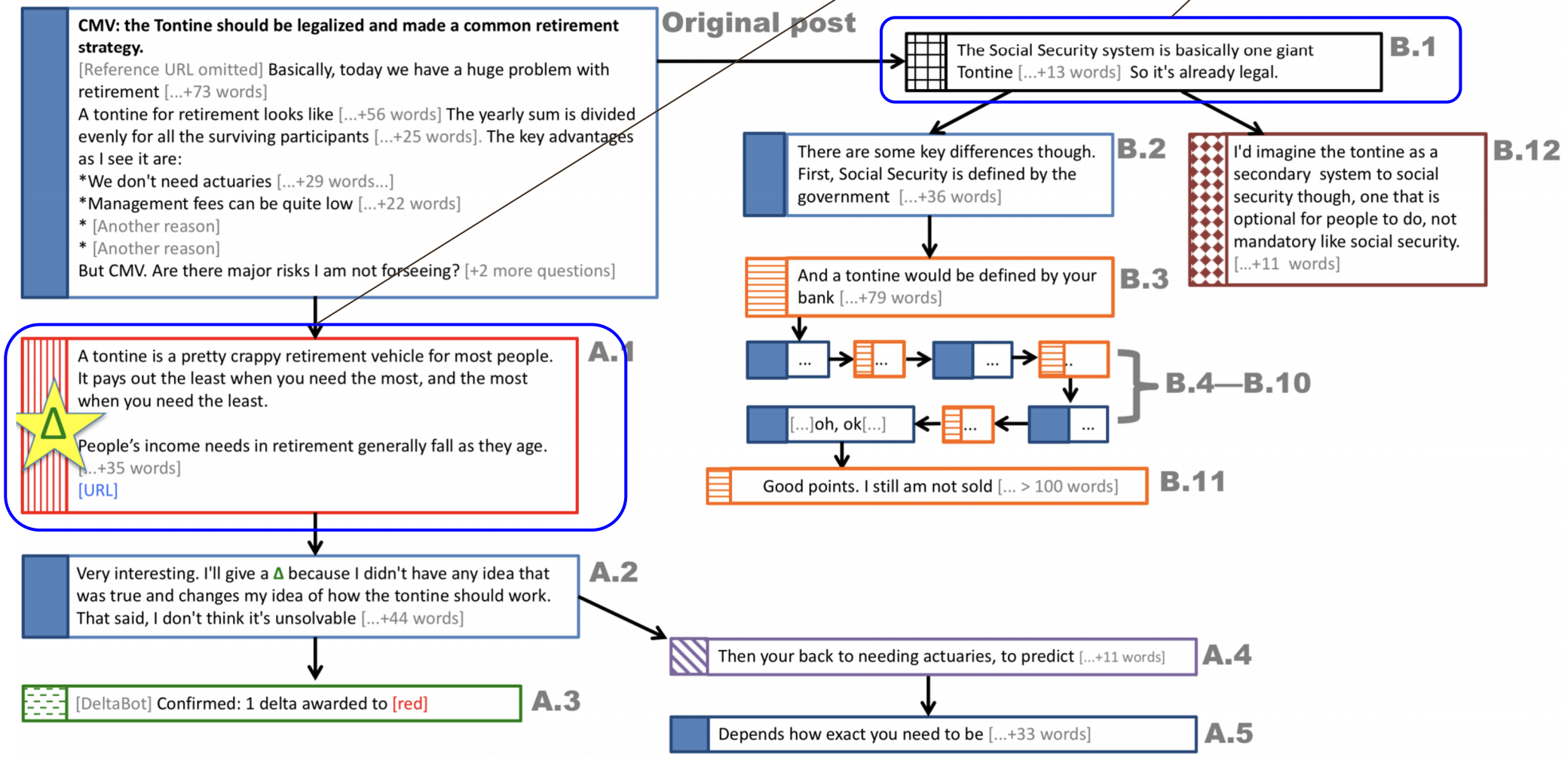
- Properties such as intensity, valence, and framing (Althoff et al. 2014, Bailey et al. 2014, Bryan et al. 2013, etc)
- Social aspects such as social proof and authority (Chaiken 1987, Cialdini et al. 1999, etc)

/r/changemyview

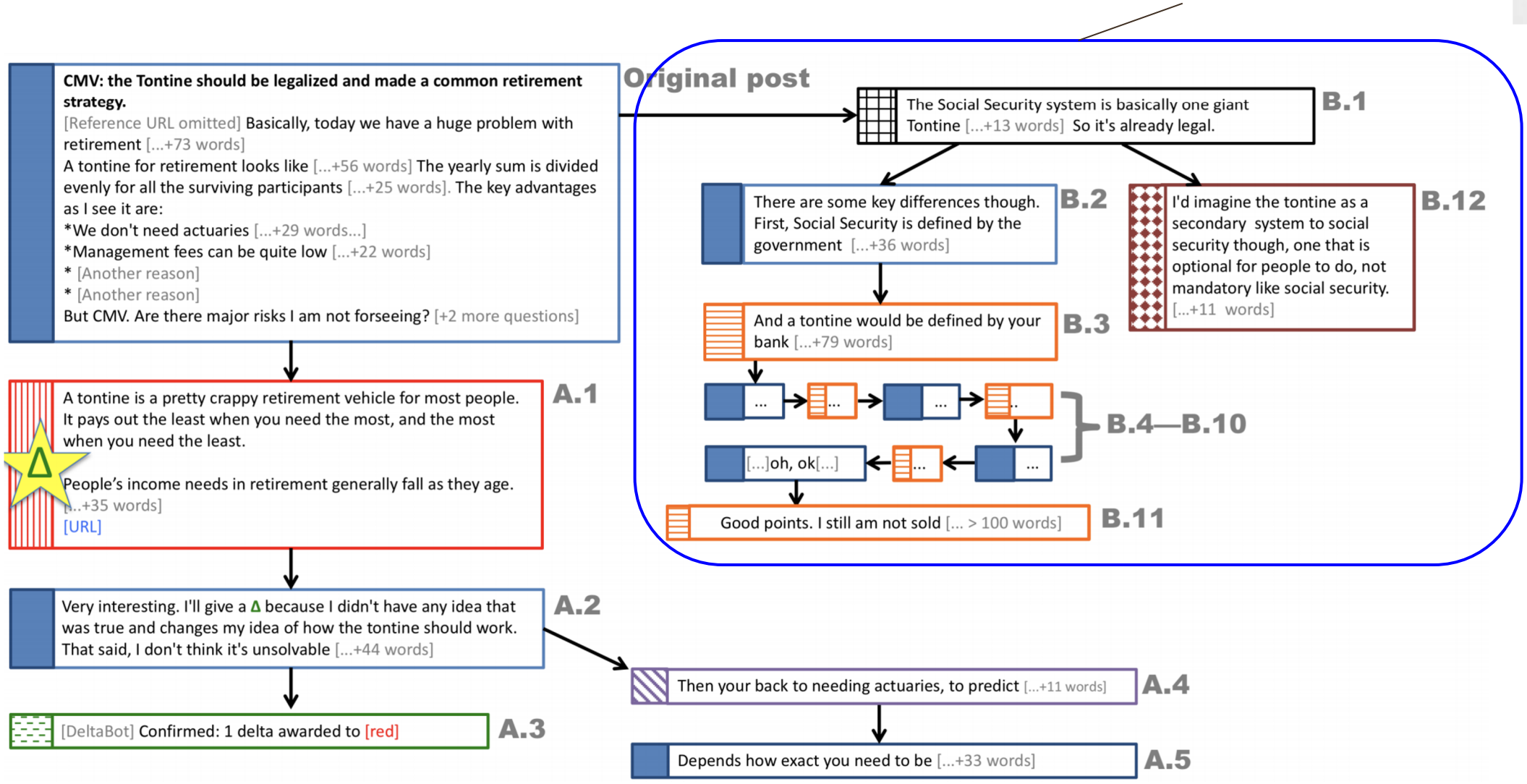
Original Post



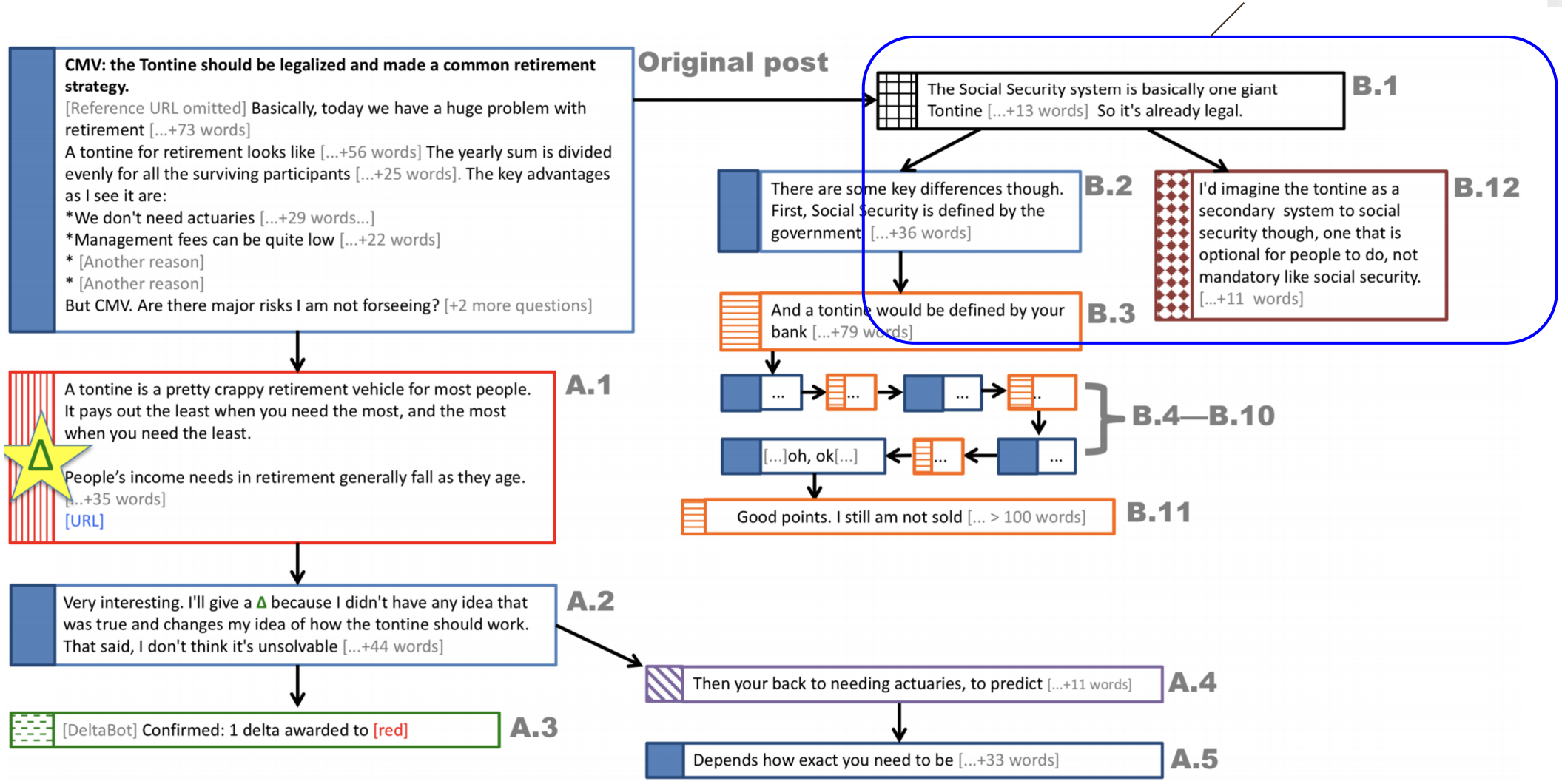
Root Replay


CMV: the Tontine should be legalized and made a common retirement strategy.
[Reference URL omitted] Basically, today we have a huge problem with retirement [...+73 words]
A tontine for retirement looks like [...+56 words] The yearly sum is divided evenly for all the surviving participants [...+25 words]. The key advantages as I see it are:
*We don't need actuaries [...+29 words...]
*Management fees can be quite low [...+22 words]
* [Another reason]
* [Another reason]
But CMV. Are there major risks I am not forseeing? [+2 more questions]
Original post
A tontine is a pretty crappy retirement vehicle for most people. It pays out the least when you need the most, and the most when you need the least.
People's income needs in retirement generally fall as they age. [...+35 words]
[URL]
A.1
Very interesting. I'll give a Δ because I didn't have any idea that was true and changes my idea of how the tontine should work. That said, I don't think it's unsolvable [...+44 words]
A.2
[DeltaBot] Confirmed: 1 delta awarded to [red]
A.3
Then your back to needing actuaries, to predict [...+11 words]
A.4
Depends how exact you need to be [...+33 words]
A.5
The Social Security system is basically one giant Tontine [...+13 words] So it's already legal.
B.1
There are some key differences though. First, Social Security is defined by the government [...+36 words]
B.2
And a tontine would be defined by your bank [...+79 words]
B.3
...
[...]oh, ok[...]
B.4—B.10
Good points. I still am not sold [... > 100 words]
B.11
I'd imagine the tontine as a secondary system to social security though, one that is optional for people to do, not mandatory like social security. [...+11 words]
B.12

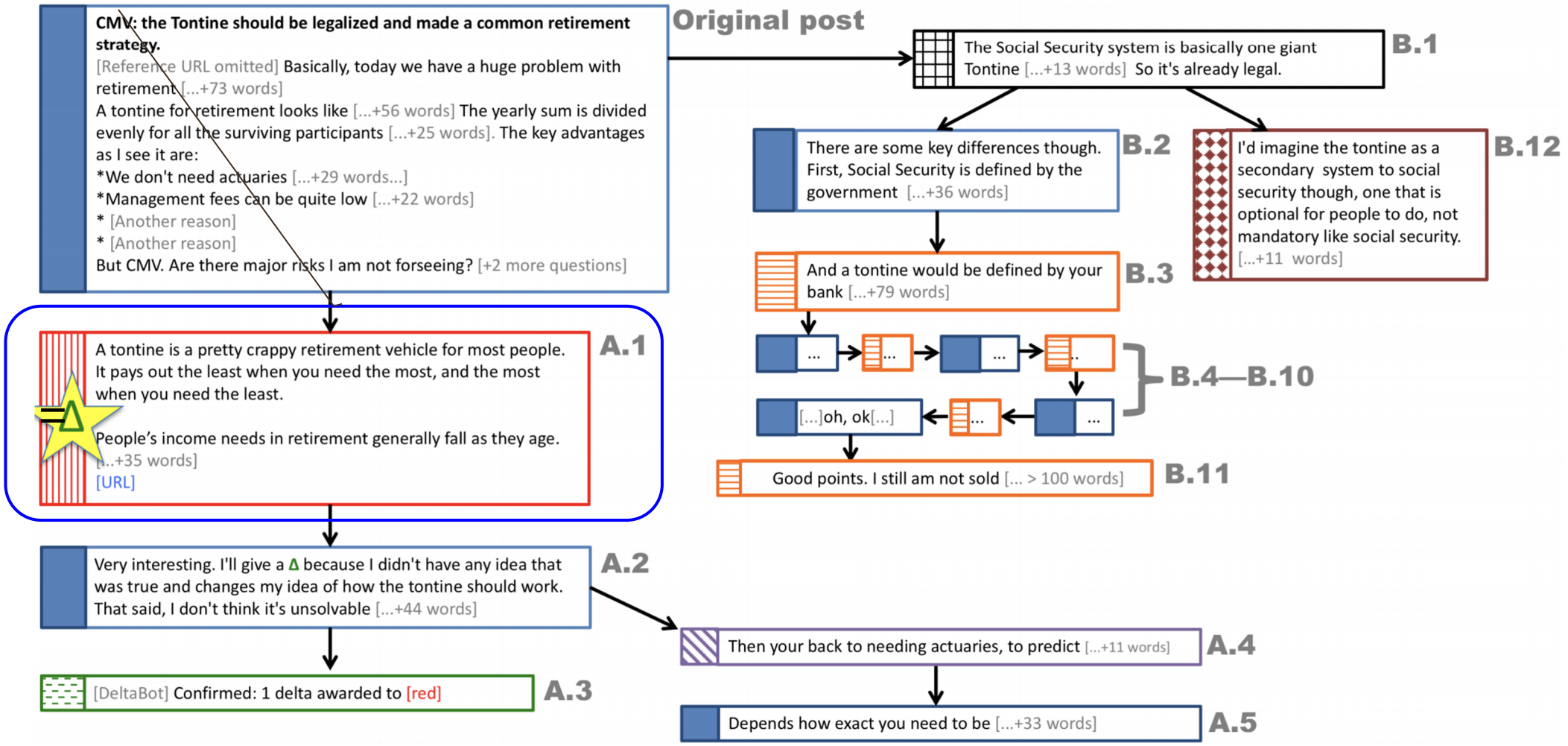
Subtree


Original post
CMV: the Tontine should be legalized and made a common retirement strategy.
[Reference URL omitted] Basically, today we have a huge problem with retirement [...+73 words]
A tontine for retirement looks like [...+56 words] The yearly sum is divided evenly for all the surviving participants [...+25 words]. The key advantages as I see it are:
*We don't need actuaries [...+29 words...]
*Management fees can be quite low [...+22 words]
* [Another reason]
* [Another reason]
But CMV. Are there major risks I am not forseeing? [+2 more questions]
A tontine is a pretty crappy retirement vehicle for most people. It pays out the least when you need the most, and the most when you need the least.
People's income needs in retirement generally fall as they age. [...+35 words]
[URL]
A.1
Very interesting. I'll give a Δ because I didn't have any idea that was true and changes my idea of how the tontine should work. That said, I don't think it's unsolvable [...+44 words]
A.2
[DeltaBot] Confirmed: 1 delta awarded to [red]
A.3
Then your back to needing actuaries, to predict [...+11 words]
A.4
Depends how exact you need to be [...+33 words]
A.5
The Social Security system is basically one giant Tontine [...+13 words] So it's already legal.
B.1
There are some key differences though. First, Social Security is defined by the government [...+36 words]
B.2
I'd imagine the tontine as a secondary system to social security though, one that is optional for people to do, not mandatory like social security. [...+11 words]
B.12
And a tontine would be defined by your bank [...+79 words]
B.3
...
[...]oh, ok[...]
B.4—B.10
Good points. I still am not sold [... > 100 words]
B.11

Path


CMV: the Tontine should be legalized and made a common retirement strategy.
[Reference URL omitted] Basically, today we have a huge problem with retirement [...+73 words]
A tontine for retirement looks like [...+56 words] The yearly sum is divided evenly for all the surviving participants [...+25 words]. The key advantages as I see it are:
*We don't need actuaries [...+29 words...]
*Management fees can be quite low [...+22 words]
* [Another reason]
* [Another reason]
But CMV. Are there major risks I am not forseeing? [+2 more questions]
Original post
A tontine is a pretty crappy retirement vehicle for most people. It pays out the least when you need the most, and the most when you need the least.
People's income needs in retirement generally fall as they age. [...+35 words]
[URL]
A.1
Very interesting. I'll give a Δ because I didn't have any idea that was true and changes my idea of how the tontine should work. That said, I don't think it's unsolvable [...+44 words]
A.2
[DeltaBot] Confirmed: 1 delta awarded to [red]
A.3
Then your back to needing actuaries, to predict [...+11 words]
A.4
Depends how exact you need to be [...+33 words]
A.5
The Social Security system is basically one giant Tontine [...+13 words] So it's already legal.
B.1
There are some key differences though. First, Social Security is defined by the government [...+36 words]
B.2
I'd imagine the tontine as a secondary system to social security though, one that is optional for people to do, not mandatory like social security. [...+11 words]
B.12
And a tontine would be defined by your bank [...+79 words]
B.3
B.4—B.10
[...]oh, ok[...]
Good points. I still am not sold [... > 100 words]
B.11

Successful Argument



Original Opinion  **Argument**

- Interaction dynamics
- Stylistic choices (language indicators) in arguments
- Resistance to persuasion

Original Opinion ⟺ **Argument**

Interaction Dynamics

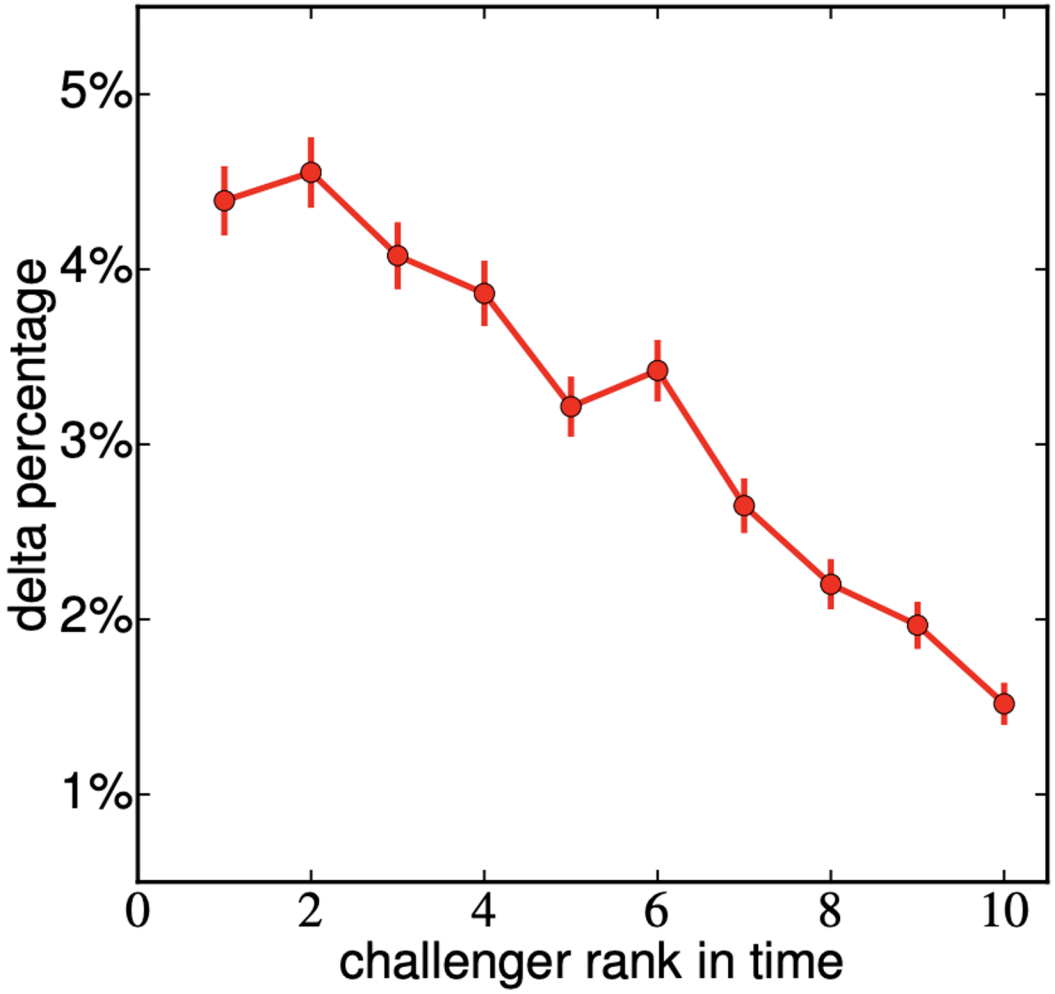
- The entry time of the challenger
- The number of back-and-forth exchanges

} Challenger's perspective

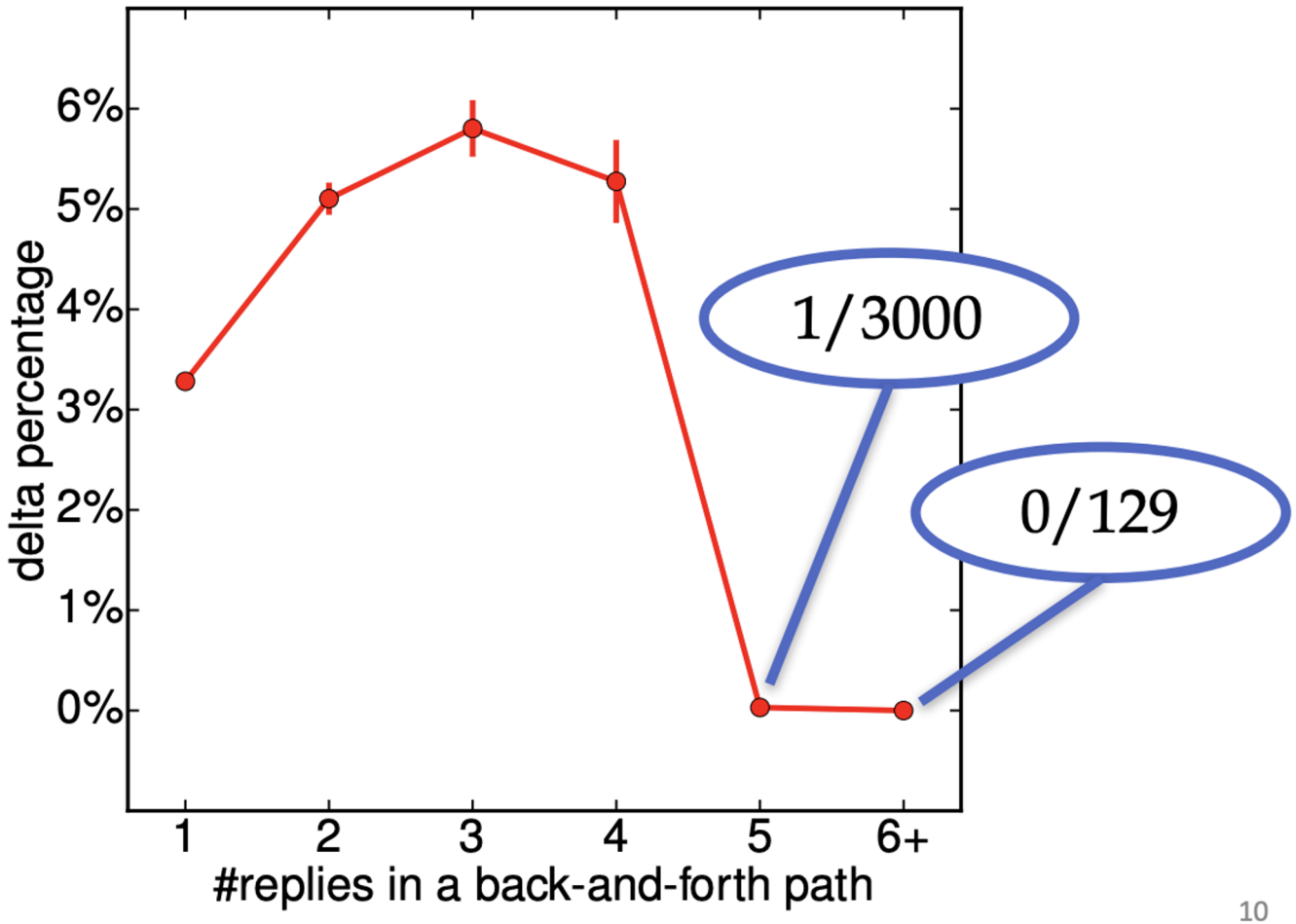
- The number of challengers
- and possibly more …

Original Post

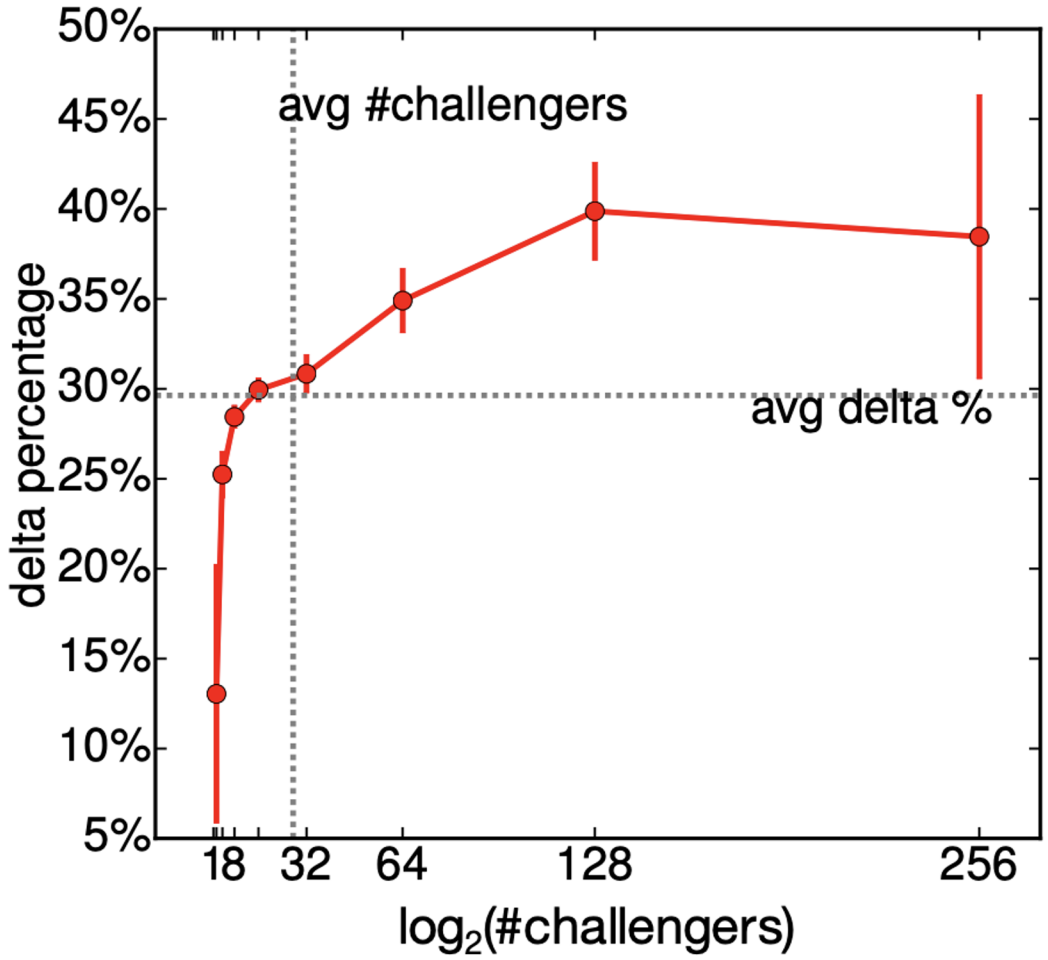
Success of a comment vs. the challenger’s entry time



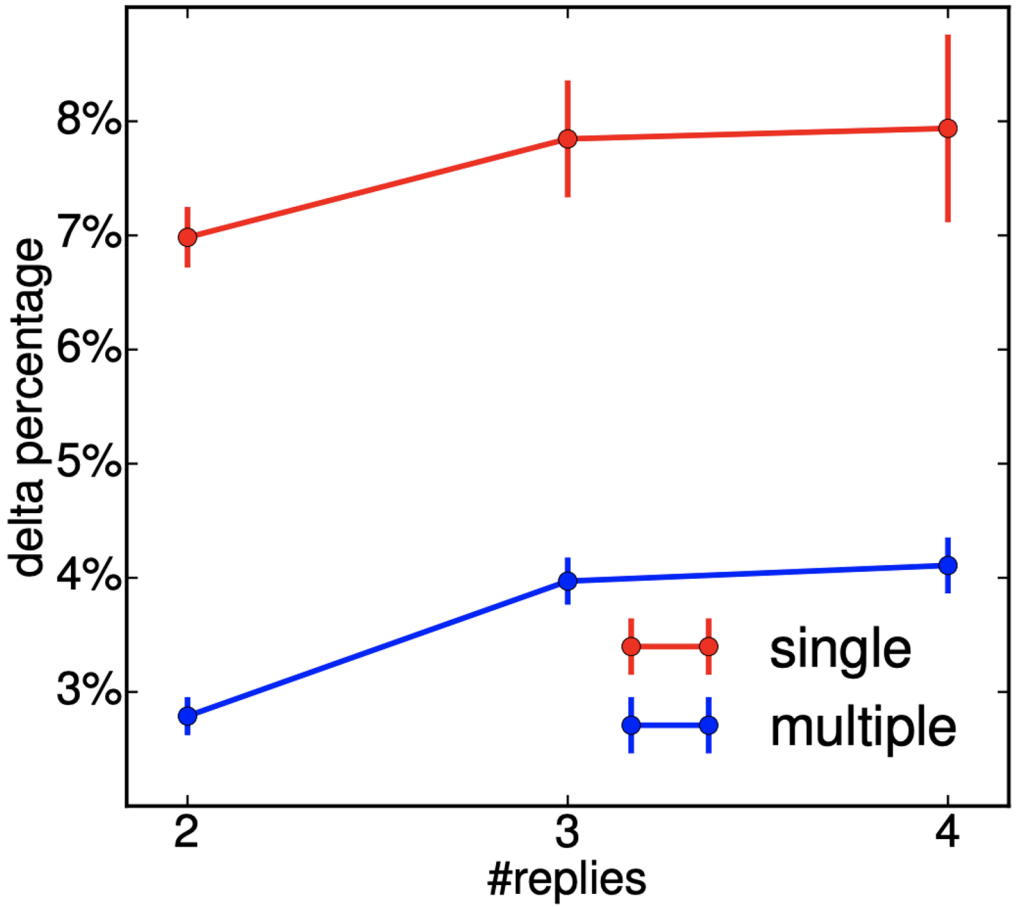
Success of a comment vs. the number of back-and-forth exchanges



Probability that opinion was changed vs. the number of challengers



Sheer number of challengers or diversity of counterarguments?



Original Opinion  **Effective/Ineffective Argument**

Stylistic choices (language indicators) in arguments

- Variations of setup
  - Root reply
  - Full path
  - Truncated

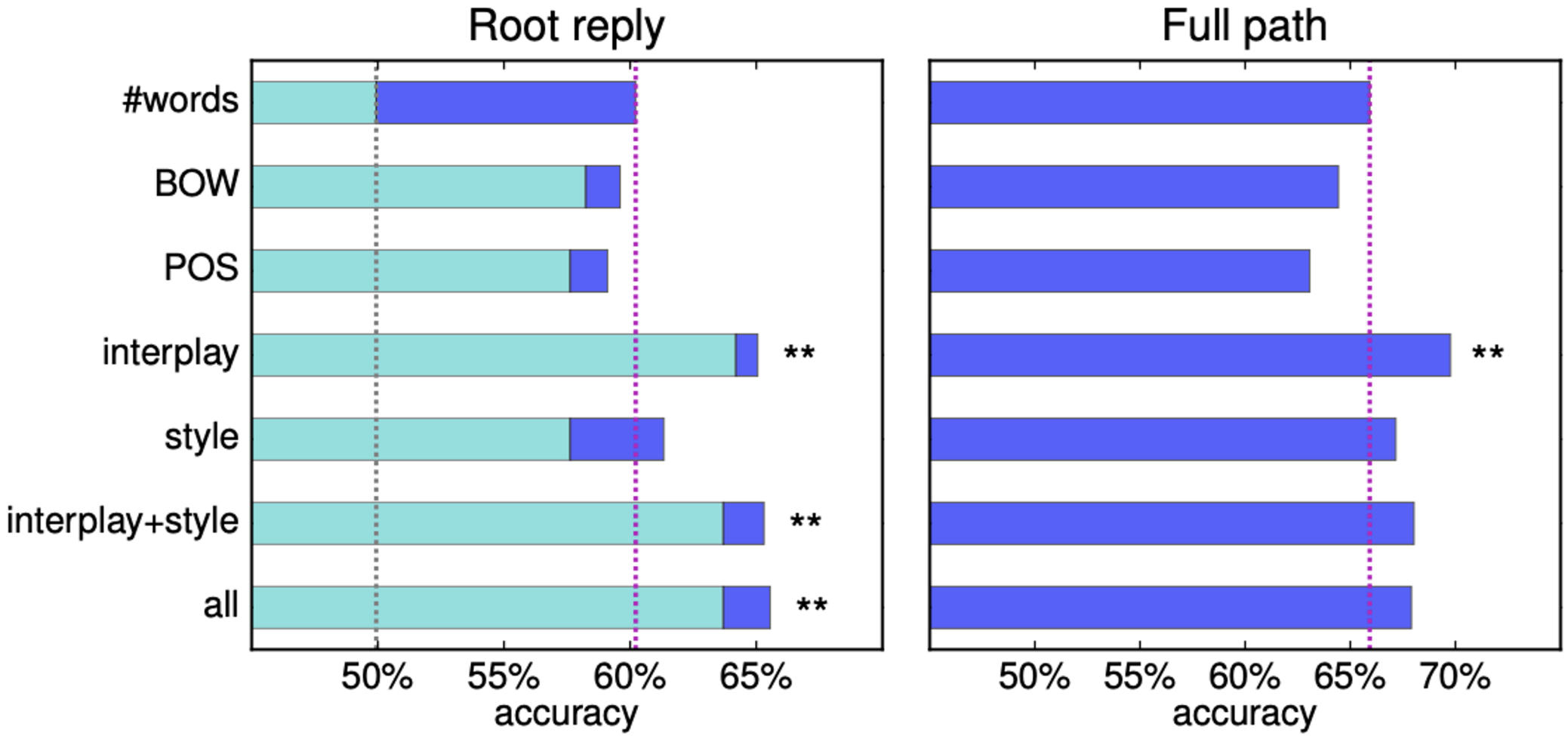
Original Opinion

Argument

- Interplay with original opinion:
  - similarity in different vocabulary sets

- Argument-only features:
  - Stylistic features
  - Bag-of-words
  - Part-of-speech tags

Performance Overview: Interplay with the OP plays an essential role.



Interplay with the original opinion

Metrics

- Number of common words
- Reply fraction
- OP fraction
- Jaccard

Types of words

- Stopwords
- Content
- All

Interplay with the original opinion

Feature name	*root reply*	*full path*
reply frac. in all	↓↓↓↓ (T)	↓↓↓↓
reply frac. in content	↓↓↓↓ (T)	↓↓↓↓
OP frac. in stopwords	↑↑↑↑ (T^R)	↑↑↑↑
#common in stopwords	↑↑↑↑ (T^R)	↑↑↑↑
reply frac. in stopwords	↓↓↓↓	↓↓↓↓
OP frac. in all	↑↑↑↑ (T^R)	↑↑↑↑
#common in all	↑↑↑↑ (T^R)	↑↑↑↑
Jaccard in content	↓↓↓↓ (T)	↓↓↓↓
Jaccard in stopwords	↑↑↑↑ (T^R)	↑↑↑↑
#common in content	↑↑↑↑ (T^R)	↑↑↑↑
OP frac. in content	↑ (T^R)	↑↑↑↑
Jaccard in all	↓ (T)	

Argument-only: stylistic features

- Number of words
- Word category-based features
- Word score-based features
- Characteristics of the entire argument
- Formatting

Number of words

Feature name	*root reply*	*full path*
#words	↑↑↑↑	↑↑↑↑

Word category-based features

- (In)definite articles:
  - General (a/an) vs specific (the)
- Positive and negative words:
  - Positive/negative from LIWC
- Arguer-relevant personal pronouns:
  - me, you, us
- Links:
  - Citing external evidence
- Hedging:
  - Hedges indicate uncertainty
  - ("It could be the case" .)
- Examples:
  - "For example", "for instance", "e.g.
- Question Marks
- Quotations

Feature name	*root reply*	*full path*
Word category–based features		
#definite articles	↑↑↑↑	↑↑↑↑
#indefinite articles	↑↑↑↑	↑↑↑↑
#positive words	↑↑↑↑ (T^R)	↑↑↑↑
#2nd person pronoun	↑↑↑↑	↑↑↑↑
#links	↑↑↑↑ (T)	↑↑↑↑
#negative words	↑↑↑↑	↑↑↑↑
#hedges	↑↑↑↑	↑↑↑↑
#1st person pronouns	↑↑↑↑	↑↑↑↑
#1st person plural pronoun	↑↑↑↑	↑↑↑↑
#.com links	↑↑↑↑ (T)	↑↑↑↑
frac. links	↑↑↑↑ (T)	↑↑↑↑
frac. .com links	↑↑↑↑ (T)	↑↑↑↑
#examples	↑	↑↑↑↑
frac. definite articles	↑ (T)	↑↑
#question marks	↑ —	↑↑↑↑
#PDF links	↑	↑↑↑
#.edu links		↑
frac. positive words	↓	
frac. question marks	—	↓
#quotations		↑↑↑↑

Word score-based features

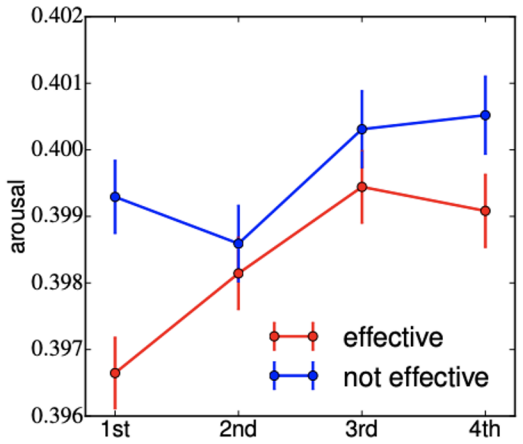
- Arousal:
  - the intensity of an emotion
  - *Dull vs terrorism*
- Concreteness:
  - the degree to which a word denotes something perceptible
  - *Hamburger vs justice*
- Dominance:
  - the degree of control expressed by a word.
  - *Earthquake vs completion*
- Valence:
  - how pleasant the word’ s denotation is.
  - *Murder vs sunshine*

Feature name	*root reply*		*full path*
Word score–based features			
arousal	↓	(*T*)	↓↓↓
valence	↓		

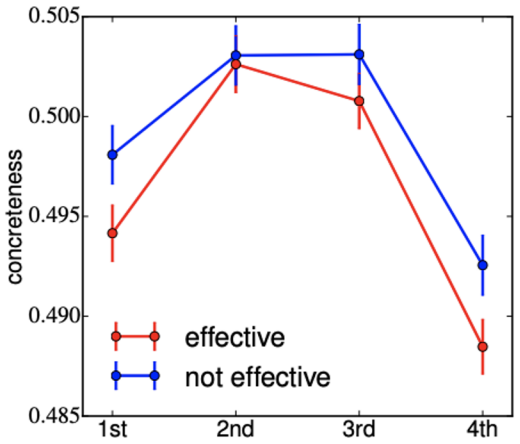
Formatting

Feature name	*root reply*	*full path*
Markdown formatting		
#italics	↑↑↑↑ —	↑↑↑↑
bullet list	↑↑↑↑ —	↑↑↑↑
#bolds	↑↑ —	↑↑↑↑
numbered words	↑	↑↑↑↑
frac. italics	↑ —	↑

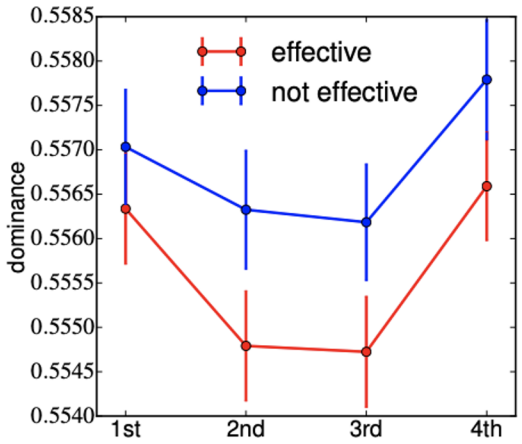
Structure of an argument



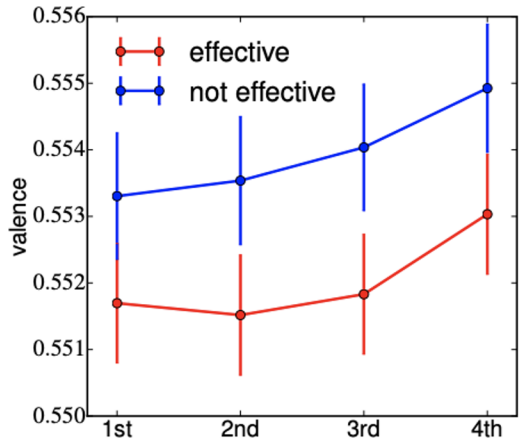
(a) Arousal in root replies.



(b) Concreteness in root replies.

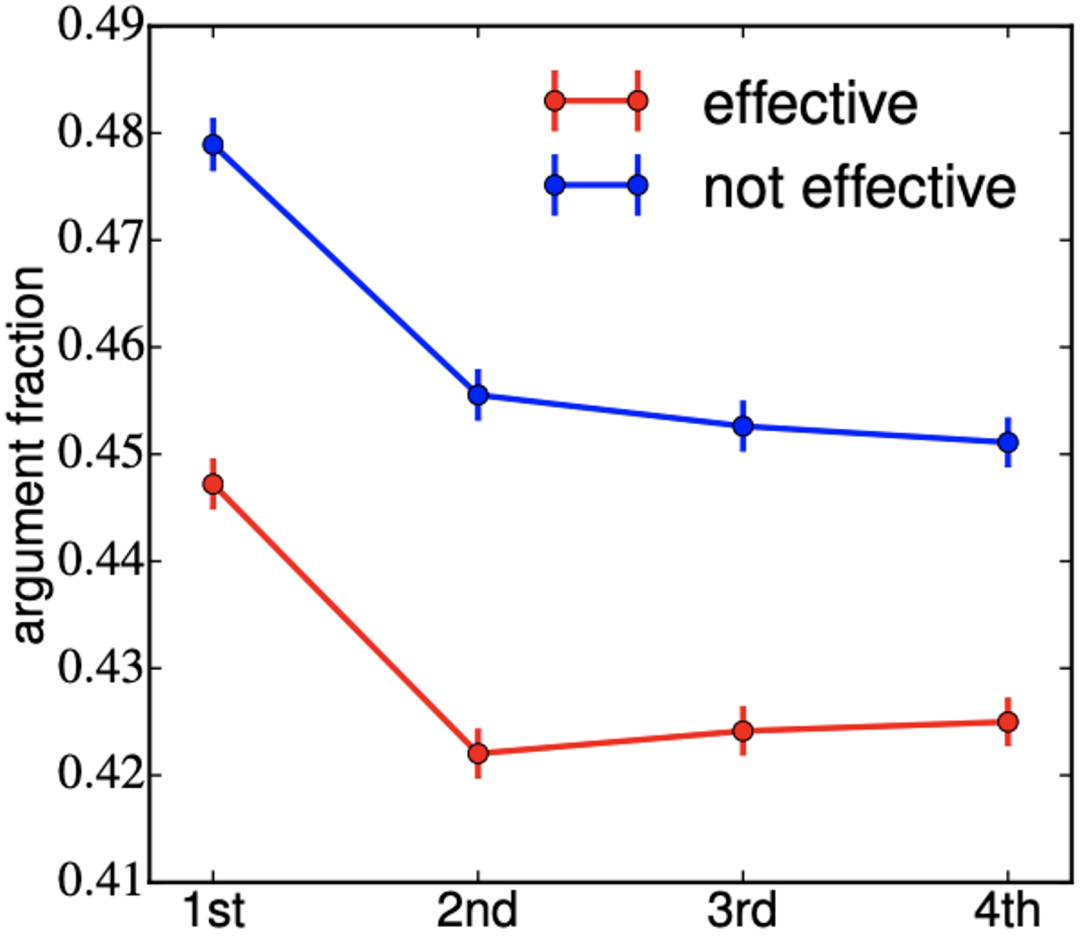


(c) Dominance in root replies.



(d) Valence in root replies.

Interplay with original posts



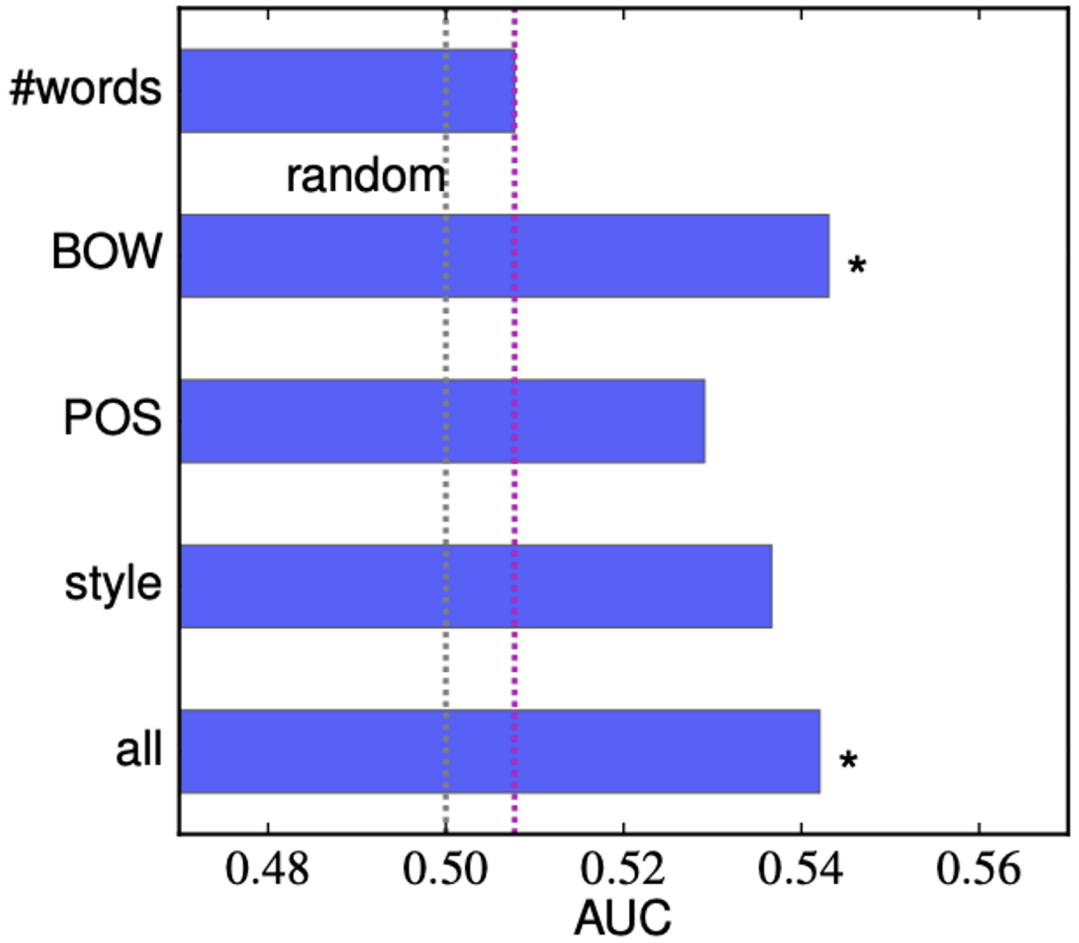
Original Opinion  **Argument**

Resistance to persuasion

Original opinion features:

- Stylistic features
- Bag-of-words
- Part-of-speech tags

Performance Overview: A much harder task

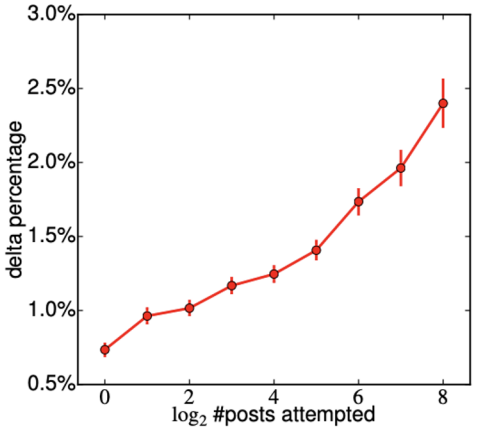


Malleability of the original opinion

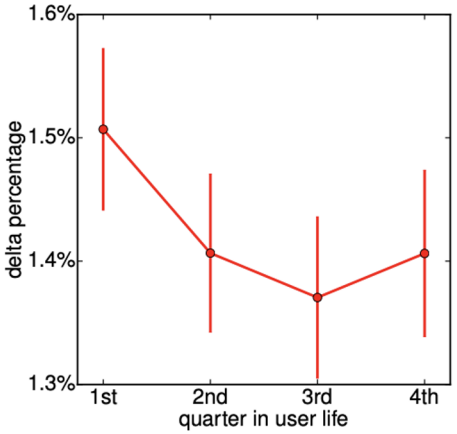
Feature name	More malleable?
#1st person pronouns	↑↑↑↑
frac. 1st person pronoun	↑↑↑↑
dominance	↑↑↑↑
frac. 1st person plural pronoun	↓↓↓
#paragraphs	↑↑
#1st person plural pronoun	↓↓
#bolds	↑
arousal	↓
valence	↑
bullet list	↑

Discussions

- Experience level of challengers?
- Topics of original opinions and arguments?
  - topic: food, eat, eating, thing, meat
  - topic: government, state, world, country, countries
- Structure of arguments?
  - "First0–but1–because2"
  - "Now1–then2–instead3"
- Persuasion strategies?



(a) Delta ratio vs. $\log_2$ #posts.



(b) No growth over life.