Winning on the Merits: The Joint Effects of Content and Style on Debate Outcomes

Lu Wang[1], Nick Beauchamp[2,3], Sarah Shugars[3], Kechen Qin[1]

[1]College of Computer and Information Science

[2]Department of Political Science

[3]Network Science Institute

Northeastern University

College of Computer and Information Science

Why Do We Care about Debates?

"I felt Sanders yelled better on big banks and political corruption but Clinton yelled better on Israel and the minimum wage."

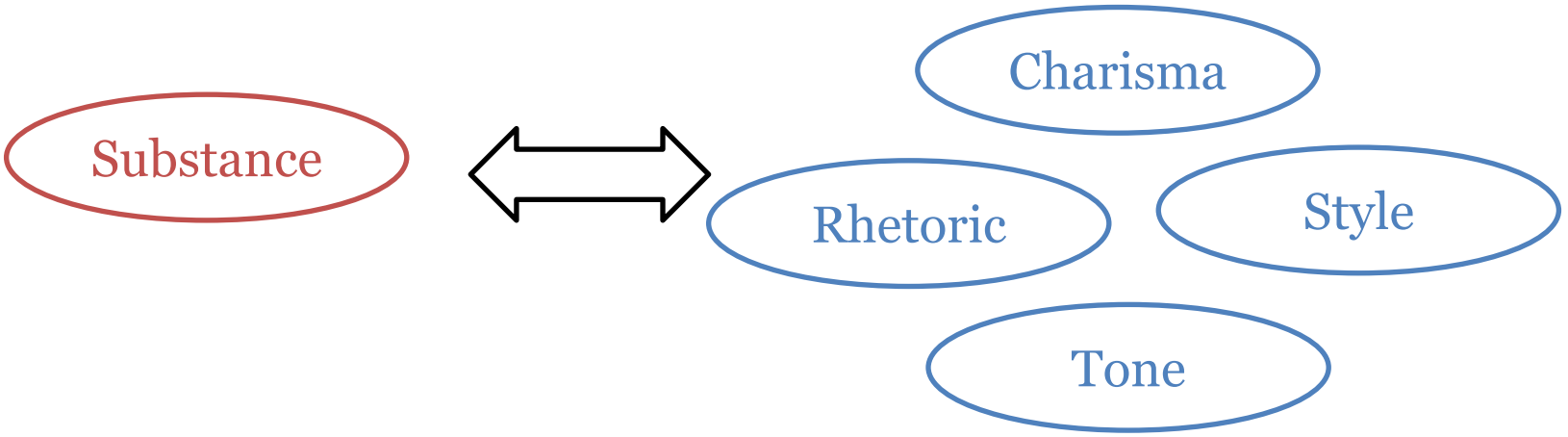
[Source: www.newyorker.com]

How Does One Win a Debate?

- Ideally, win a debate based on the merits
  - Facts
  - Reasons
  - Mutual understanding

How Does One Win a Debate?

- However, in reality…



???

I think my strongest asset, maybe by far, is my temperament. I have a winning temperament.

The Joint Effect:
A Discussion on "*Abolishing the Death Penalty*"

Pro: … When you look at capital convictions, you can demonstrate on innocence grounds a 4.1 percent error rate… I mean, would you accept that in flying airplanes? …

Con: … The risk of an innocent person dying in prison and never getting out is greater if he's sentenced to life in prison than it is if he's sentenced to death. So the death penalty is an important part of our system.

The Joint Effect:
A Discussion on "*Abolishing the Death Penalty*"

Topic: execution of the innocents

Pro: … When you look at capital convictions, you can demonstrate on innocence grounds a 4.1 percent error rate… I mean, would you accept that in flying airplanes? …

Con: … The risk of an innocent person dying in prison and never getting out is greater if he's sentenced to life in prison than it is if he's sentenced to death. So the death penalty is an important part of our system.

The Joint Effect:
A Discussion on "*Abolishing the Death Penalty*"

Topic: execution of the innocents

Numbers

Pro: … When you look at capital convictions, you can demonstrate on innocence grounds a 4.1 percent error rate… I mean, would you accept that in flying airplanes? …

Con: … The risk of an innocent person dying in prison and never getting out is greater if he's sentenced to life in prison than it is if he's sentenced to death. So the death penalty is an important part of our system.

The Joint Effect:
A Discussion on "*Abolishing the Death Penalty*"

Topic: execution of the innocents

Questions

Pro: … When you look at capital convictions, you can demonstrate on innocence grounds a 4.1 percent error rate… I mean, would you accept that in flying airplanes? …

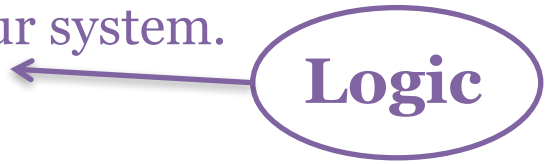
Con: … The risk of an innocent person dying in prison and never getting out is greater if he's sentenced to life in prison than it is if he's sentenced to death. So the death penalty is an important part of our system.

The Joint Effect:
A Discussion on "*Abolishing the Death Penalty*"

Topic: execution of the innocents

Pro: … When you look at capital convictions, you can demonstrate on innocence grounds a 4.1 percent error rate… I mean, would you accept that in flying airplanes? …

Con: … The risk of an innocent person dying in prison and never getting out is greater if he's sentenced to life in prison than it is if he's sentenced to death. So the death penalty is an important part of our system.



Content and Style are Deeply Intertwined

- Two topic strength assumptions
- Debate topics come with intrinsic strengths for different sides.
  - E.g., "*execution of the innocents*" is stronger for Pro (supporting abolishing death penalty) than Con.
- Style may vary for strong arguments and weak arguments.

Related Work

- Style and content have been studied separately.
- Stylistic elements of arguments
  - Argument extraction and classification [Feng and Hirst, 2011; Mochales and Moens, 2011; Stand and Gurevych, 2014]
  - Persuasion effect [Tan et al., 2016; Cano-Basave and He, 2016]
- Topic control and shift
  - Self-promotion and attacks [Zhang et al., 2015]

Our Goal

- We aim to build a debate prediction model which is able to
  - identify the topics and their intrinsic strengths for different sides
  - model the interaction between topic strength and linguistic features of arguments

Data

- 118 Intelligence Squared U.S. debates
- Oxford-style

Opening Statement

⇩

Moderated Discussion

⇩

Closing Statement

Data

- Who is the winner
  - Recording votes before and after debate
    - pro, con, undecided
  - Winner: the side that gains more votes

Preprocessing: Argument Identification

Deterrent effect

⇩

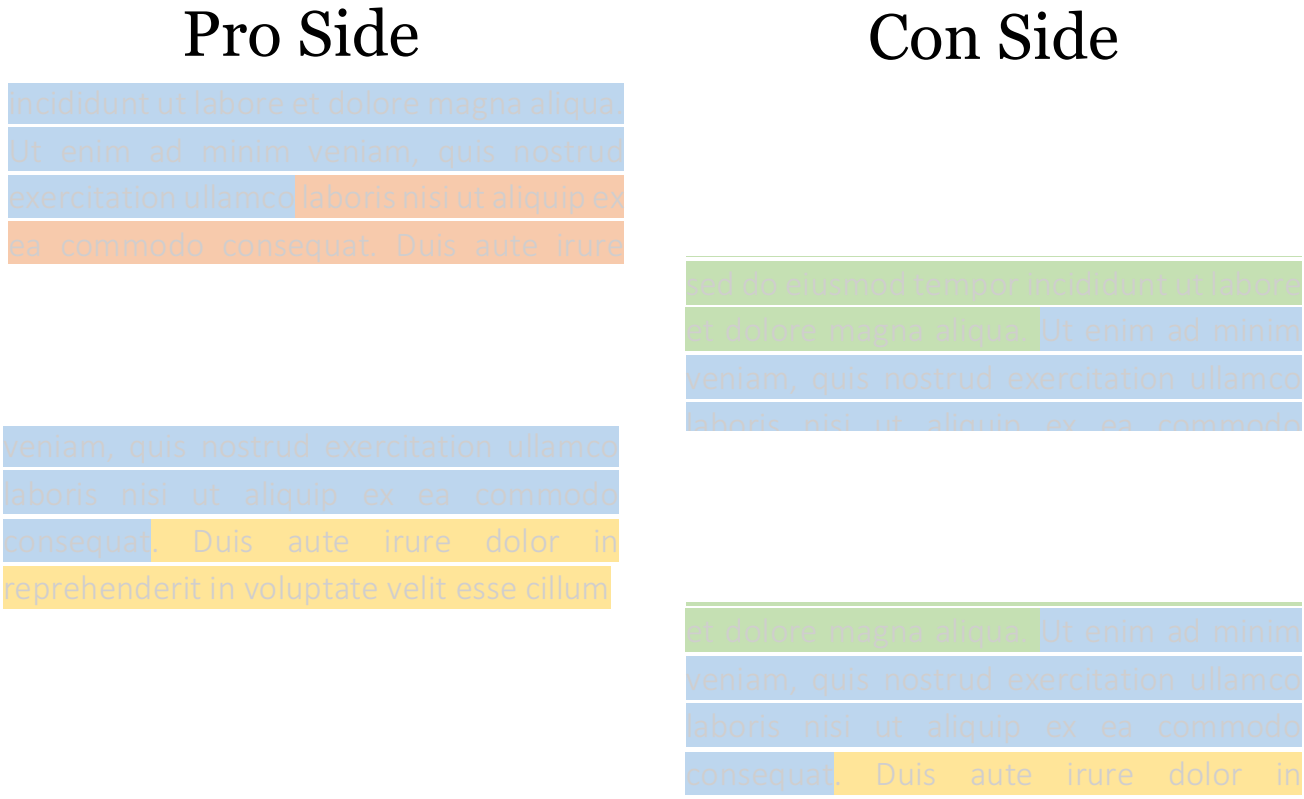
Execution of innocents

Pro: The death penalty does not deter. The National Academy of Sciences recently reviewed all of the studies and found no evidence of a deterrent effect. … The death penalty is administered arbitrarily. … When you look at capital convictions, you can demonstrate on innocence grounds a 4.1 percent error rate…

Hidden topic Markov model (HTMM) [Gruber et al., 2007]: A topic modeling approach that models topics and topic transitions

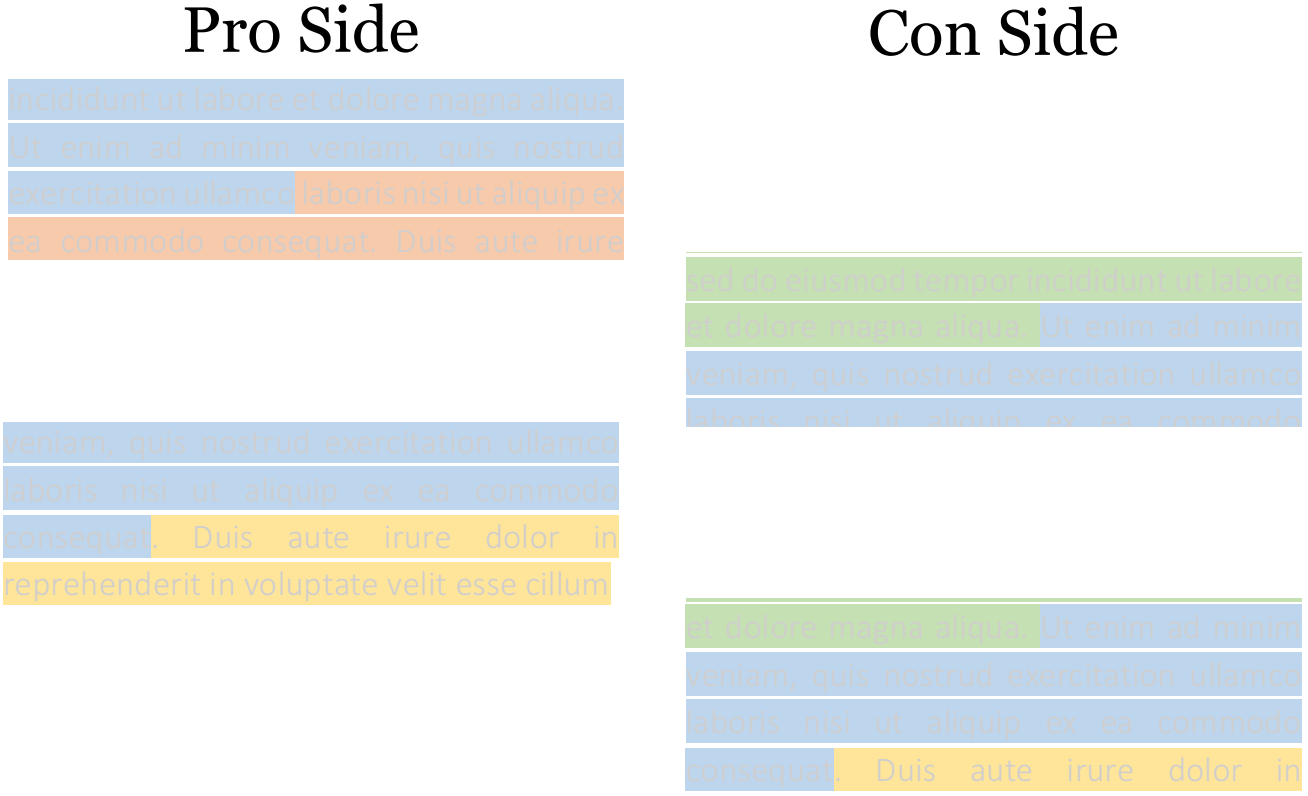
The Debate Prediction Model

- For each debate d_i , it consists of a sequence of arguments, $\mathbf{x}_i$, from two sides.



The Debate Prediction Model

- The debate outcome is y_i.
  - y_i=1 means Pro wins and y_i=-1 means Con wins.



The Debate Prediction Model

- Topic system: debaters issue arguments from K topics.
- Each topic has an intrinsic persuasion strength, which may vary between sides.

The Debate Prediction Model

- Topic system: debaters issue arguments from *K* topics.
- Each topic has an intrinsic persuasion strength, which may vary between sides. For example,

Debate: *"Abolishing the Death Penalty"*

	T1: execution of innocents	**T2: deterrent effect**	**T3: morality**
Pro	Strong	Weak	Strong
Con	Weak	Strong	Weak

The Debate Prediction Model

- Topic system: debaters issue arguments from K topics.
- Each topic has an intrinsic persuasion strength, which may vary between sides.
- The topic strength system is represented as $\mathbf{h}_i$.
  - Unknown, and need to be inferred for both training and test

Stylistic Features × Topic Strength

Each argument *x* is represented as a feature vector:

$$\phi(x, \mathbf{h_i})$$

Argument
-> *represented with linguistic/stylistic features*

Topic Strength

Stylistic Features × Topic Strength

$$\phi(x, \mathbf{h}_i)$$

		Stylistic Feature	Topic strength
T1		# “you” = 1	T1 = Strong
T2		# “you” = 2	T2 = Weak
T1		# “you” = 1	T1 = Strong
T3		# “you” = 1	T3 = Weak
T1		# “you” = 0	T1= Strong

Stylistic Features × Topic Strength

$$\phi(x, \mathbf{h}_i)$$

		Stylistic Feature	Topic strength
T1	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo	# "you" = 1	T1 = Strong
T2		+	
T1		# "you" = 1	T1 = Strong
T3		+	
T1		# "you" = 0	T1= Strong
		= 2	

Stylistic Features × Topic Strength

$$\phi(x, \mathbf{h}_i)$$

		Stylistic Feature	Topic strength
T1	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo	# "you" = 1	T1 = Strong
T2		+	
T1		# "you" = 1	T1 = Strong
T3		+	
T1		# "you" = 0	T1= Strong

Pro: $\Phi(\mathbf{x}_i^{\mathbf{p}}, \mathbf{h}_i) = \sum \phi(x, \mathbf{h}_i)$ Con: $\Phi(\mathbf{x}_i^{\mathbf{c}}, \mathbf{h}_i) = \sum \phi(x, \mathbf{h}_i)$

The Debate Prediction Model

- Compute scores for two sides
  - **Pro**: $f^p = \max_{\mathbf{h_i}} \mathbf{w} \cdot [\Phi(\mathbf{x_i^p}, \mathbf{h_i}) - \Phi(\mathbf{x_i^c}, \mathbf{h_i})]$
  - **Con**: $f^c = \max_{\mathbf{h_i}} \mathbf{w} \cdot [\Phi(\mathbf{x_i^c}, \mathbf{h_i}) - \Phi(\mathbf{x_i^p}, \mathbf{h_i})]$
  - **w** contains the feature weights, and is learned from training data.
  - $\mathbf{h_i}$ is inferred topic strengths.

The Debate Prediction Model

- Compute scores for two sides
  - **Pro**: $f^p = \max_{\mathbf{h_i}} \mathbf{w} \cdot [\Phi(\mathbf{x_i^p}, \mathbf{h_i}) - \Phi(\mathbf{x_i^c}, \mathbf{h_i})]$
  - **Con**: $f^c = \max_{\mathbf{h_i}} \mathbf{w} \cdot [\Phi(\mathbf{x_i^c}, \mathbf{h_i}) - \Phi(\mathbf{x_i^p}, \mathbf{h_i})]$
  - If $f^p > f^c$, then y=1 (**Pro** wins);
  - otherwise, y=-1 (**Con** wins).

Training

- To learn the feature weights **w**, we use the large margin training objective:

$$\min_{\mathbf{w}} \frac{1}{2}\|\mathbf{w}\|^2 + C \cdot \sum_i l(-y_i \cdot \max_{\mathbf{h_i}} \mathbf{w} \cdot [\Phi(\mathbf{x_i^p}, \mathbf{h_i}) - \Phi(\mathbf{x_i^c}, \mathbf{h_i})])$$

Features

- Basic Features
  - Personal pronouns
    - Implication of communicative goals [Brown and Gilman, 1960 Wilson, 1990]
  - Sentiment and emotion words
    - Subjective language usage is prevalent.

Features

- Style Features
  - Formality [Brooke et al., 2010]
    - Revealing speakers' opinions or intentions [Irvine, 1979]
    - E.g., digest vs. imbibe, add vs. affix
  - Hedging [Hyland, 2005]
    - E.g., probably, somewhat

Features

- Discourse Features
  - Discourse structure has been shown effective for detecting argumentative structure [Stab and Gurevych, 2014]
  - Usage of discourse connectives
    - E.g. *however*, *moreover*, *therefore*
  - Collected from Penn Discourse Treebank [Prasad et al., 2007]

Features

- Argument Features
  - Readability scores
    - Flesch reading ease score

Features

- Interaction with Opponents
  - Whether the debater addresses opponent's point, i.e., arguments of the same topic
  - Number of words used to address opponent

Experimental Setup

- Leave-one-out
- Baselines:
  - Ngrams: unigrams + bigrams
  - Audience feedback: applause + laughter

Main Results

Features	SVM
Unigrams + Bigrams	61.0
Audience Feedback (applause and laughter)	56.8

	Without Topic Strength	With Topic Strength
Basic (unigrams, sentiment words, etc)	57.6	59.3
+ Style, Semantic, Discourse	59.3	65.3
+ Argument	62.7	69.5
+ Interaction	66.1	**73.7**

[Note: our features do not contain bigrams or above.]

Main Results

Features	SVM
Unigrams + Bigrams	61.0
Audience Feedback (applause and laughter)	56.8

	Without Topic Strength	With Topic Strength
Basic (unigrams, sentiment words, etc)	57.6	59.3
+ Style, Semantic, Discourse	59.3	65.3
+ Argument	62.7	69.5
+ Interaction	66.1	**73.7**

[Note: our features do not contain bigrams or above.]

Discussions

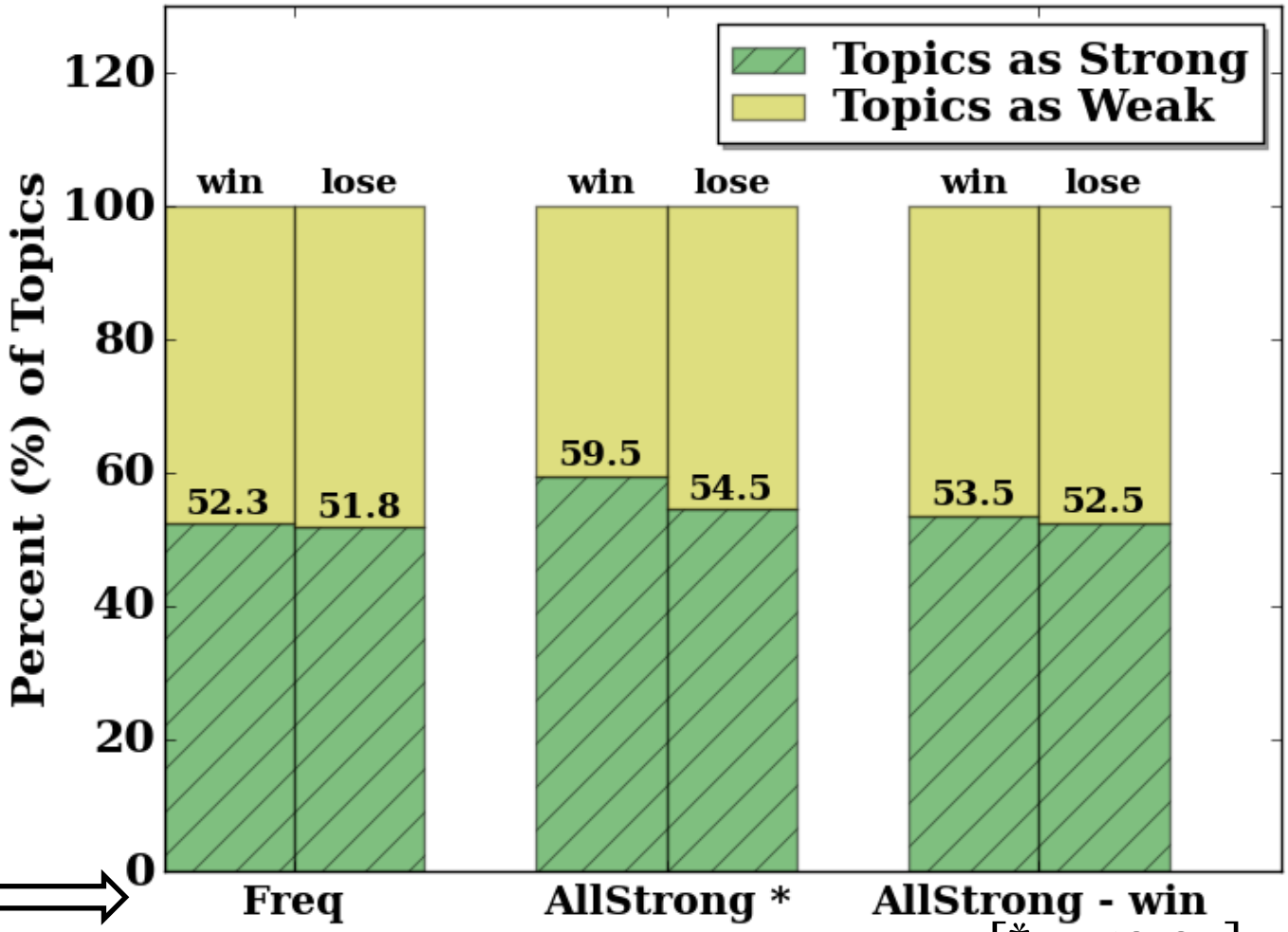
- Argument Usage:
  - Do winning sides use more strong arguments?
- Topic Shift:
  - Do debaters change topics to ones that benefit them?
- Salient Features:
  - Do strong arguments and weak arguments have different indicative features?

Winners Own More Strong Topics

Freq: one side that uses more arguments is assigned as strong

AllStrong: both sides are assigned as strong

AllStrong - win: winning side is assigned as strong

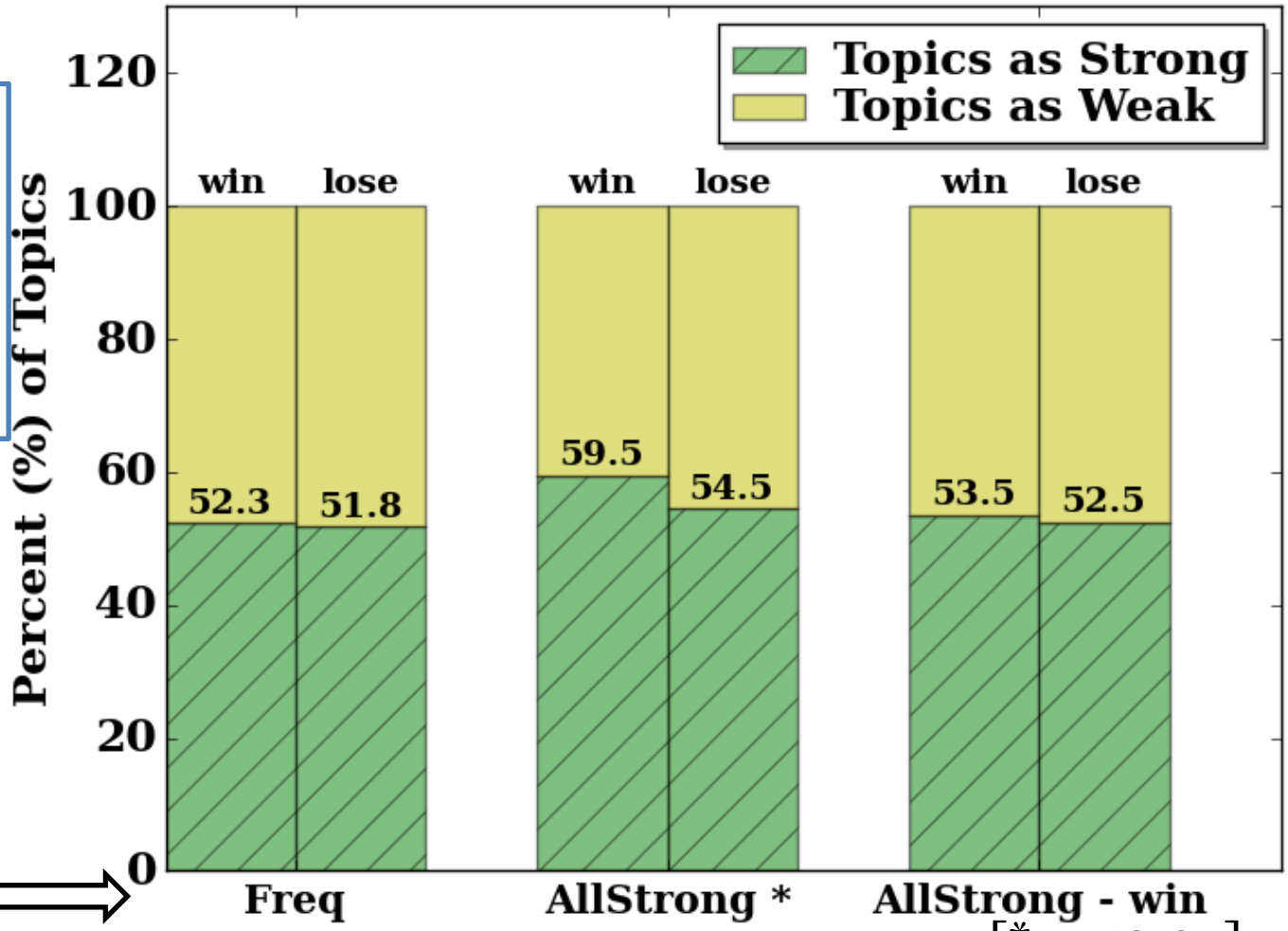


Topic strength initialization (training)

[*: p<0.05]

Winners Own More Strong Topics

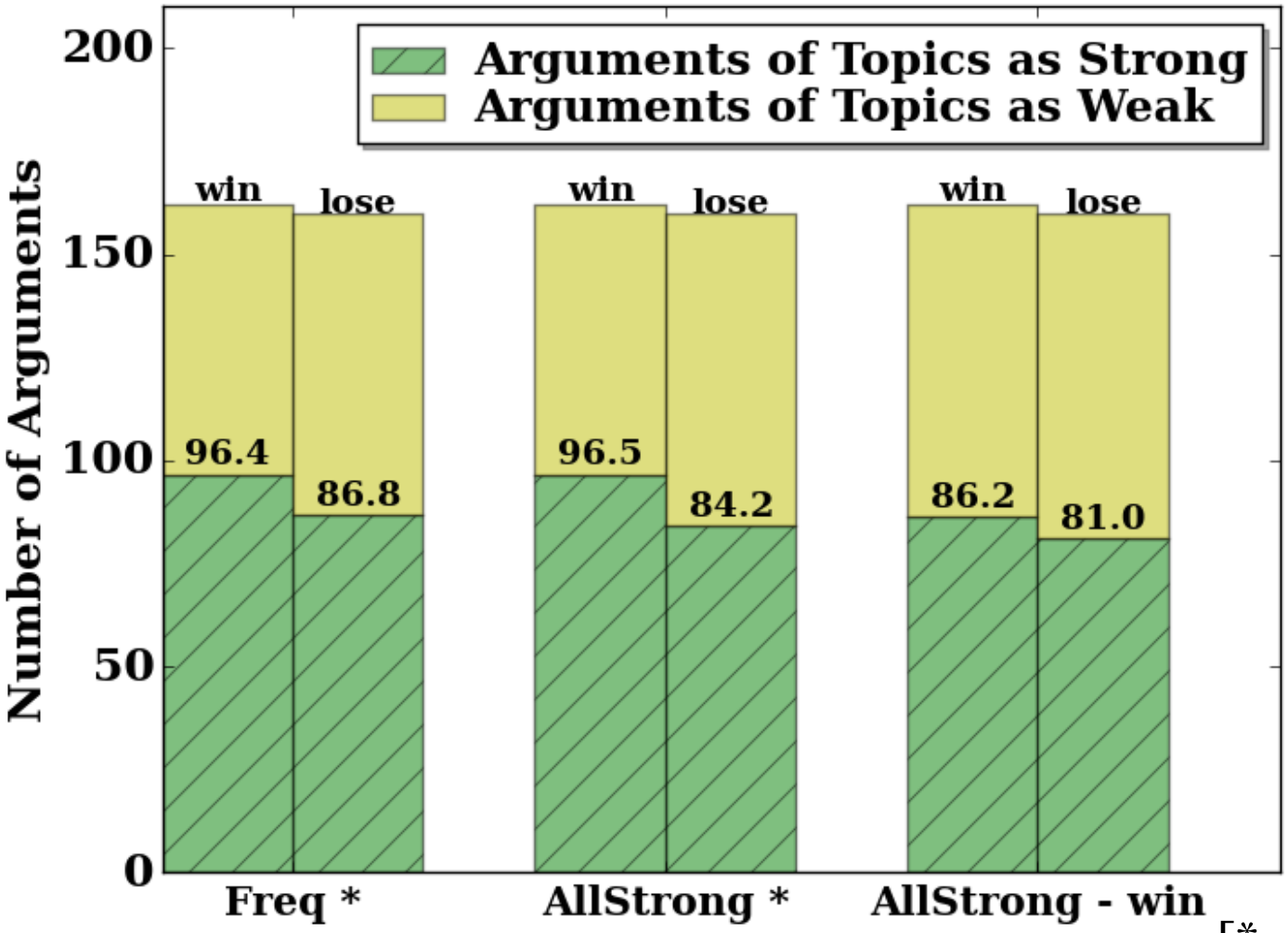
Human annotators labeled 44.4% of topics as strong for winners, compared to 30.1% for losers.



Topic strength initialization (training)

[*: p<0.05]

Winners Uses More Strong Arguments



[*: p<0.05]

Discussions

- Argument Usage:
  - Do winning sides use more strong arguments?

- Topic Shift:
  - Do debaters change topics to ones that benefit them?

- Salient Features:
  - Do strong arguments and weak arguments have different indicative features?

Topic Shifting Behavior

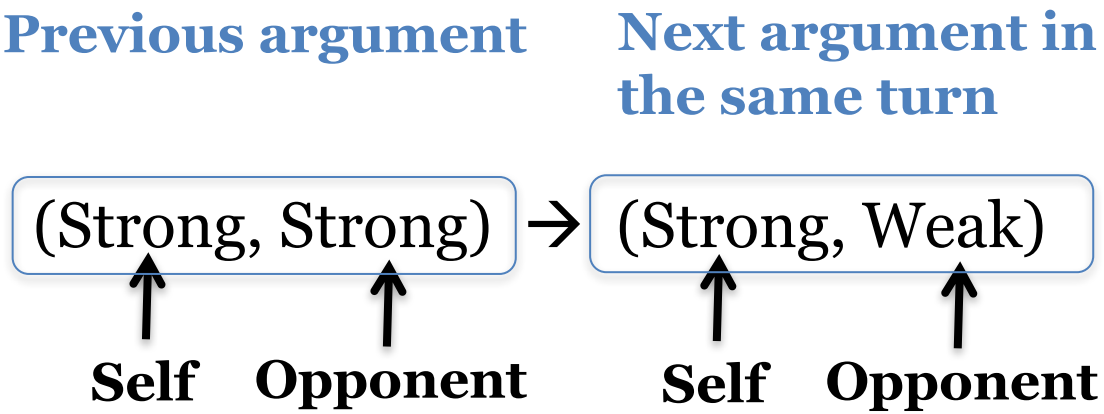
Debaters make 1.5 topic shifts in each turn on average.

	Winners		**Losers**	
Shift-to	STRONG	WEAK	STRONG	WEAK
	61.4%	38.6%	53.6%	46.4%

Topic Shifting Behavior

	Winners		Losers	
Shift-to	STRONG	WEAK	STRONG	WEAK
	61.4%	38.6%	53.6%	46.4%

Especially, one of top shifting behavior for winners:



Discussions

- Argument Usage:
  - Do winning sides use more strong arguments?
- Topic Shift:
  - Do debaters change topics to ones that benefit them?
- Salient Features:
  - Do strong arguments and weak arguments have different indicative features?

Salient Features with Topic Strength

	STRONG Topics	WEAK Topics
Basic Features	# “we”	# “you”
	# “they”	# “I”
	# “emotion:sadness”	# “emotion:joy”
	# “emotion:disgust”	# “emotion:trust”

Salient Features with Topic Strength

	STRONG Topics	WEAK Topics
Basic Features	# “we”	# “you”
	# “they”	# “I”
	# “emotion:sadness”	# “emotion:joy”
	# “emotion:disgust”	# “emotion:trust”
Style, Semantic, Discourse Features	# formal words	# “discourse:contrast”
	# “frame:capability”	# “frame:certainty”
	# “frame:information”	

Salient Features with Topic Strength

	STRONG Topics	WEAK Topics
Basic Features	# “we”	# “you”
	# “they”	# “I”
	# “emotion:sadness”	# “emotion:joy”
	# “emotion:disgust”	# “emotion:trust”
Style, Semantic, Discourse Features	# formal words	# “discourse:contrast”
	# “frame:capability”	# “frame:certainty”
	# “frame:information”	
Argument Features	# sentiment:negative	# sentiment:neutral

Salient Features with Topic Strength

	STRONG Topics	WEAK Topics
Basic Features	# “we”	# “you”
	# “they”	# “I”
	# “emotion:sadness”	# “emotion:joy”
	# “emotion:disgust”	# “emotion:trust”
Style, Semantic, Discourse Features	# formal words	# “discourse:contrast”
	# “frame:capability”	# “frame:certainty”
	# “frame:information”	
Argument Features	# sentiment:negative	# sentiment:neutral
Interaction Features	# words addressing opponent’s argument	if addressing opponent’s argument
	# common words with opponent’s argument	

Conclusion

- We present a debate prediction model that learns latent persuasive strengths of topics, and their interaction with linguistic style of arguments.

- We find that
  - winners tend to use more stronger arguments;
  - debaters tend to strategically shift topics to stronger ground;
  - strong and weak arguments differ in their language usage.

Future Work

- Better representation of topics and arguments

- Argumentation process in other types of debates, e.g., online debates, Supreme Court oral arguments

Thank you!

More information:
www.ccs.neu.edu/home/luwang/
luwang@ccs.neu.edu

Find Northeastern NLP at nlp.ccis.northeastern.edu/