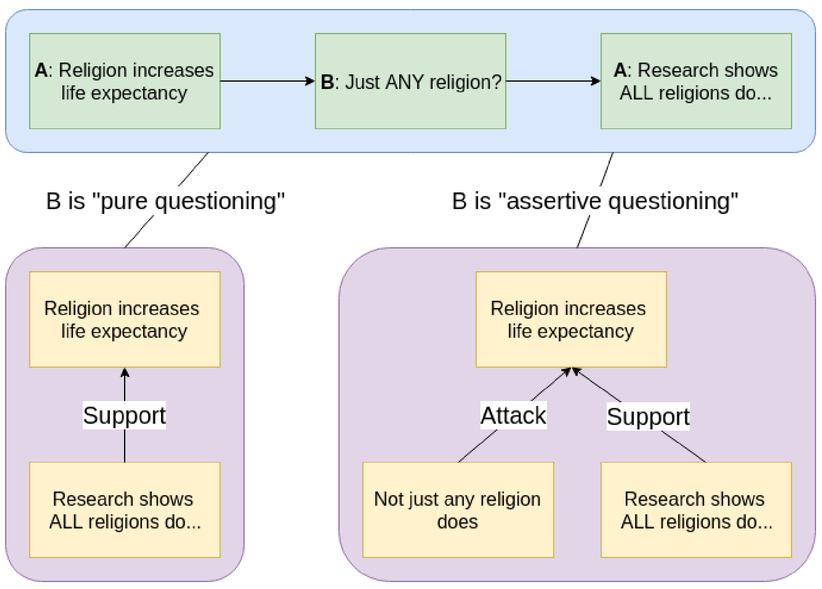
Detecting Argumentative Discourse Units with Linguistic Alignment

Timothy Niven and **Hung-Yu Gao**
Intelligent Knowledge Management Lab
National Cheng Kung University, Taiwan
tim.niven.public@gmail.com

Discourse acts and argument maps



- Pure questioning is not an argumentative discourse unit
- But assertive questioning is
- Judging pure versus assertive questioning requires complex pragmatic judgments

Linguistic alignment

A's message	B's reply: has pronoun	B's reply: no pronoun
has pronoun	8	2
no pronoun	5	5

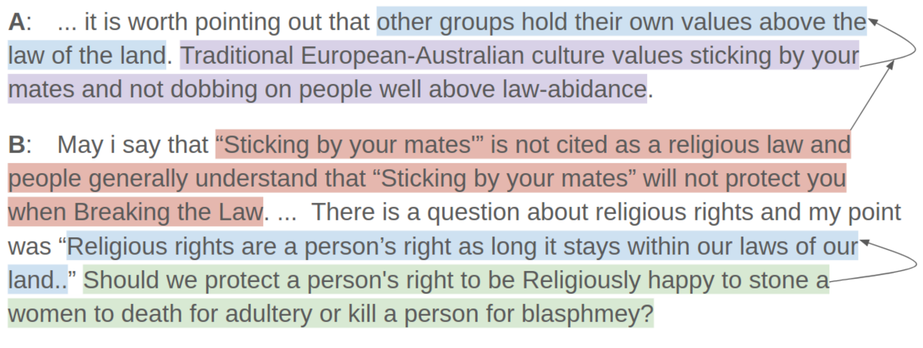
$$P(B|\neg A) = 0.5$$
$$P(B|A) = 0.8$$
$$P(B|A) - P(B|\neg A) = 0.3 > 0$$

- An increase in the probability of using a *function word category* (e.g. pronouns), relative to a baseline probability (where the message replied to does not contain the category)
- B aligns to A in this conversation

Research questions

- Are different discourse acts associated with different alignment effects?
- Are alignment calculations useful for predicting argumentative discourse acts?

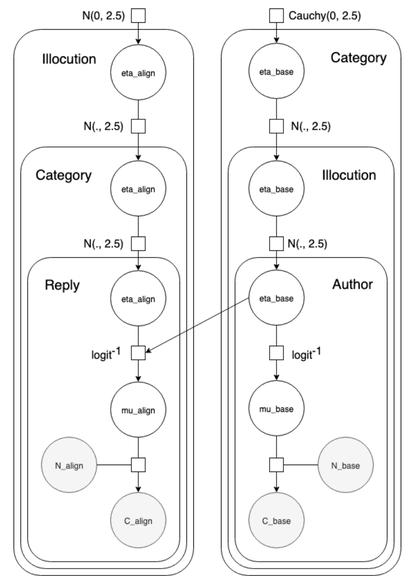
Data

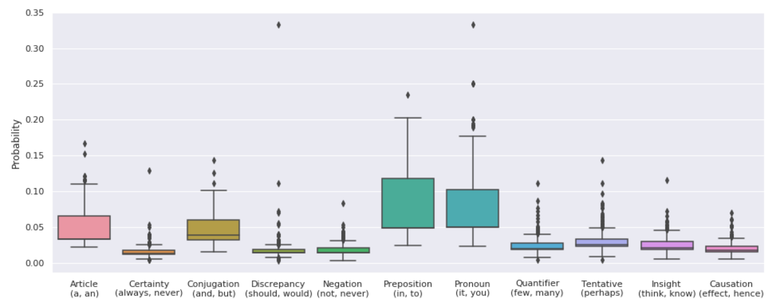


- We annotate replies (just B) in message-reply pairs, taken from below-the-line comments on an academic news website, The Conversation
- This data point is annotated as {*Asserting* (blue), *Disagreeing* (red), *Assertive Questioning* (green)}
- In this preliminary work we only have 800 data points marked by a single annotator

Bayesian Model

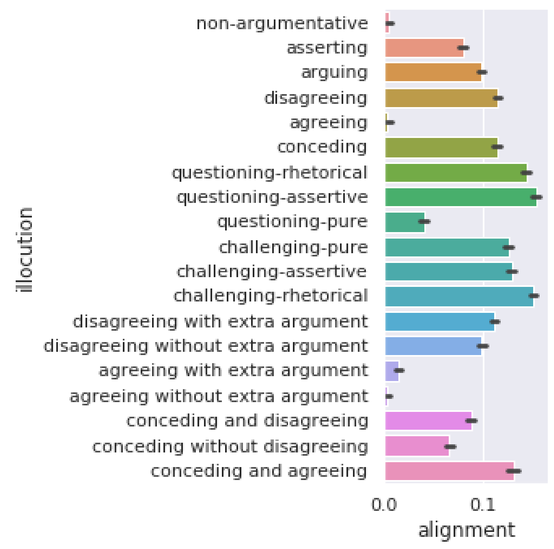
- Message is treated as a bag of words and category usage as a binomial draw
- We cannot aggregate statistics over messages in a conversation like WHAM
- Average comment length is 82.5 words ($\sigma = 66.5$)
- Lowest category usage frequency is 0.8% ($\mu = 3.6\%$, $\sigma = 2.2\%$)
- Therefore this likely systematically underestimates alignment
- However we do have very reliable baseline estimates from $\sim$ 1M comments





Different discourse acts have different alignment effects

- Error bars are 95% credible intervals on the Bayesian posterior
- Pairwise t-test for difference in means yields only 6.4% (22/342) insignificant at $p < 0.05$
- Given the limitations of our preliminary study we cannot draw general conclusions about the specific patterns observed



Alignment is useful for predicting argumentative discourse acts

- Logistic regression with BoV
- Concatenating an 11-dim vector of alignment scores over categories improves ROC AUC
- Let $\mathbb{C}_a$ be the set of categories in the original message
- $\eta^{(c)}$ is the replier's logit-space baseline usage of category c
- $r^{(c)}$ is the observed rate of usage of category c in the reply
- Each category alignment score $s^{(c)}$ is $\text{logit}(r^{(c)}) - \eta^{(c)}$ if $c \in \mathbb{C}$ else 0

Future work

- Expanding the dataset and using multiple annotators
- Cross-linguistic comparison with Mandarin
- What is the relationship between alignment and different argumentation schemes?