Linguistic Department

The Ohio State University

CLIPPERS Group

What: This is a weekly forum open to anyone with an interest in computational linguistics.

Who: Faculty from CSE and Linguistics, and PhD students.

The instigators are the CLLT faculty (de Marneffe, Elsner, Fosler-Lussier, Ritter, Schuler, Sun, White, Xu)

Naming: Following the example of the other spirited discussion groups in the Linguistics department: Changelings, Commies, Lacqueys, Phonies, Pragmatics, the Psycholinguistics Lab Meeting, So Mean and Synners.

Research Interest

Linguistics :

Computational Psycholinguistics, Incremental Parsing and Interpretation, Spoken Dialog Systems, Paraphrasing, Natural Language Generation, Computational pragmatics & semantics, veridicality assessment, grounding meaning from Web data, textual entailment, coreference resolution, language acquisition.

Computer Science Engineering:

Machine learning, natural language processing and social media, human behavior understanding, Machine learning and knowledge extraction, Statistical NLP, Spoken Dialog Systems, Speech Recognition.

I do not disagree: Leveraging monolingual alignment to detect disagreement in dialogue

Ajda Gokcen, Marie-Catherine de Marneffe

ACL - 2015

Linguistic Department

The Ohio State University

The problem of interest

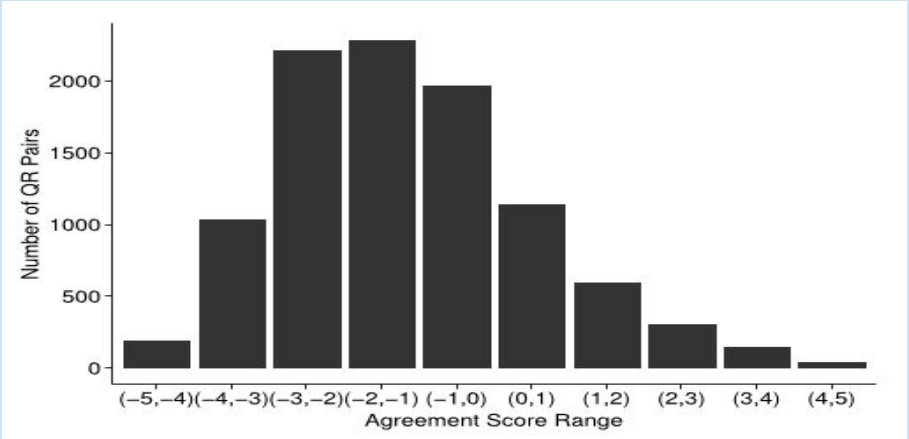
What : NLP model to automatically identify agreement and disagreement in debate corpuses by extracting information that is expressed indirectly.

Input: Quote-Response (QR) pairs from Internet Argument Corpus (IAC). The corpus is annotated for agreement via Mechanical Turk.

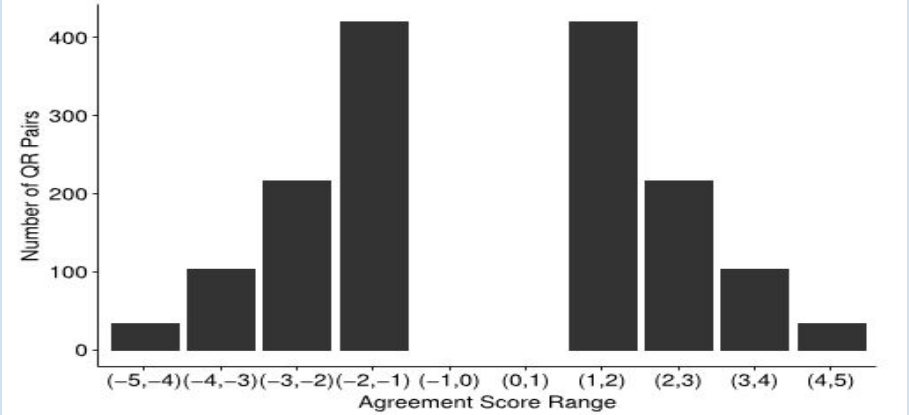
Output: Classification of the response in each QR pair as agree/disagree with the paired quote using a scoring range.

	Quote	Response	Score
1	CCW LAWS ARE FOR TRACKING GUN OWNERS WHO EXERCISE THIER RIGHTS!!!	I agree. What is the point? Felons with firearms do not bother with CCW licenses.	2.5
2	God doesn't take away sinful desires. You've never had sinful desires? I know I have. People assume that when you become a Christian some manner of shield gets put up around you and shields you from "worldly" things. I believe that's wrong, I actually believe that life as a Christian is very hard. We often pawn it off as the end of our troubles to "convert" people. I don't believe it.	Yes, God does take away sinful desires. (If you ask Him.) I'm not saying that it doesn't take any work on your part, though. When you have a sinful desire, you allow a thought to become more than just a stray idea. You foster and encourage the thought and it becomes a desire. God takes away the desires, helps you deal with your "stray thoughts", and shows you how to keep them from becoming desires.	-1.7
3	Your idea about science is a philosophy of science. [...] *The Apostles saw Jesus walk on water.* There was no 'measure' by your version of science, but what they saw remains true.	Many people once believed that the earth is flat: perhaps some still do. [...] *The apostles may have believed that Jesus walked on water: that does NOT make it true.*	-2
4	*does life end here?*	*end where?* ambiguously phrased. if "here" = "death", then yes! by definition, yes!	-1.4
5	*Is even 'channel' sufficiently ateleological a verb?*	Yes. It describes an action without ascribing its form to its end result, outcome, whatever but strictly to a cause's force's in action. [...] *But since it is understood that mechanical forces can also 'channel', unintentional, out of simple mechanics, the word channel cannot be called teleological.* In the same way, 'sorting' can be considered non-teleological, hence mechanical, and thus suited to your glossary, because things can be sorted by mechanical forces alone.	2.8

Table 1: Quote-Response (QR) pairs from Internet Argument Corpus(IAC).



(a) Full dataset.



(b) Balanced training set.

Figure 1: Agreement score distribution of the dataset, before and after balancing.

Scoring range:
-5 is strong disagreement,
+5 is strong agreement , and
[-1, +1] neutral.

	Full Data Set	Balanced Training Set
Disagree	5741	779
Neutral	3125	0
Agree	1113	779
Total	9980	1158

Table 2: Category counts in training set.

Approach

- Previous work mostly relied on n-gram and grammatical dependencies features taken from respondent text.
- Approach of this paper is to introduce semantic environment features derived by comparing quote and response sentences which align well.
- To evaluate, the generated model (Alignment + Features) is compared to a previous model (Baseline Features).

Features of the Model

Common features between both models

1. N-Grams: unigrams, bigrams, trigrams from each response
2. Discourse Markers: *Oh, so, really.*
3. Response Typed Dependencies and MPQA sentiment:
4. *(agree, I) yields (positive, I) and (wrong, you) yields (negative, you)*
5. Strings of Repeated Punctuations : *!!, ??*

Features specific to New Model

Post Length: length features like word count, sentence count, average sentence length: shorter correlates with agreement.

Emoticons: popular of communicating sentiments, using RE :-D

Speech Acts: count of imperative and interrogative instances (please read carefully, try again.

Personal pronouns: inclusion of personal pronoun in response tend to indicate emotional and personal argument , esp. 2nd person.

Features specific to New Model

Explicit Truth Value: polarity and modality. Track context of instances of (agree, disagree, right, true , false) in the response. Check polarity if negation modifiers exits (never, not), modality if modal (might, could) and adverbs

Sentiment scoring: used positive/negative/neutral and strong/weak classifiers to to calculate the score of post and focal sentence, well-aligned sentences from quote and response as well as the first sentence of the response.

Factuality Comparison: given well aligned sentences from quote and response, analyze polarity, modality, and any subsequent contradictions of both quote and response.

Results

	Accuracy	Agreement			Disagreement		
		P	R	F1	P	R	F1
Baseline	71.85	70.64	74.77	72.65	73.21	68.92	71.00
Alignment+	75.45	76.04	74.32	75.17	74.89	76.58	75.73

Table 3: Accuracy, precision(P), recall(R) and F1 scores for both categories (agreement and disagreement) on the test set.

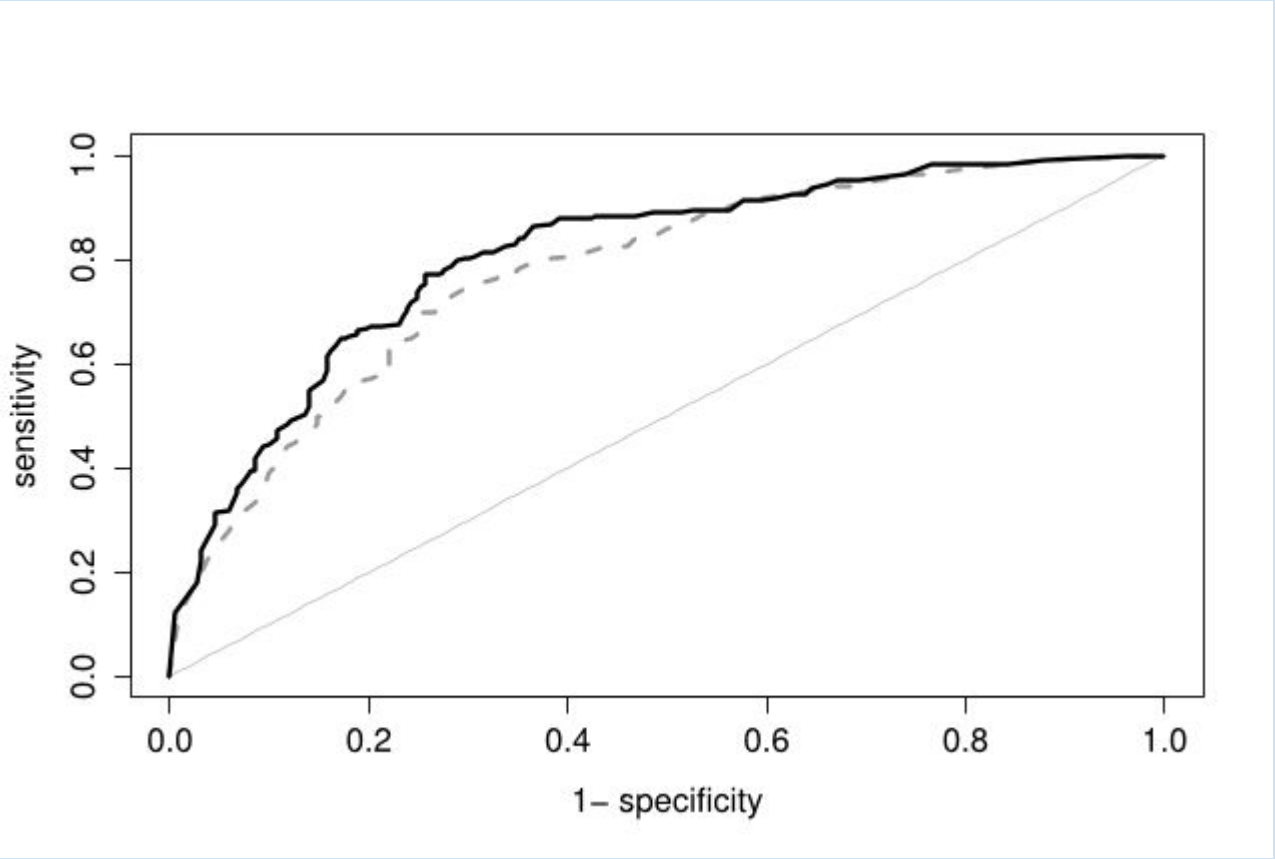


Figure 2: ROC curves. The gray dotted lines represents the baseline feature set, while the solid black line represents the alignment+ feature set.

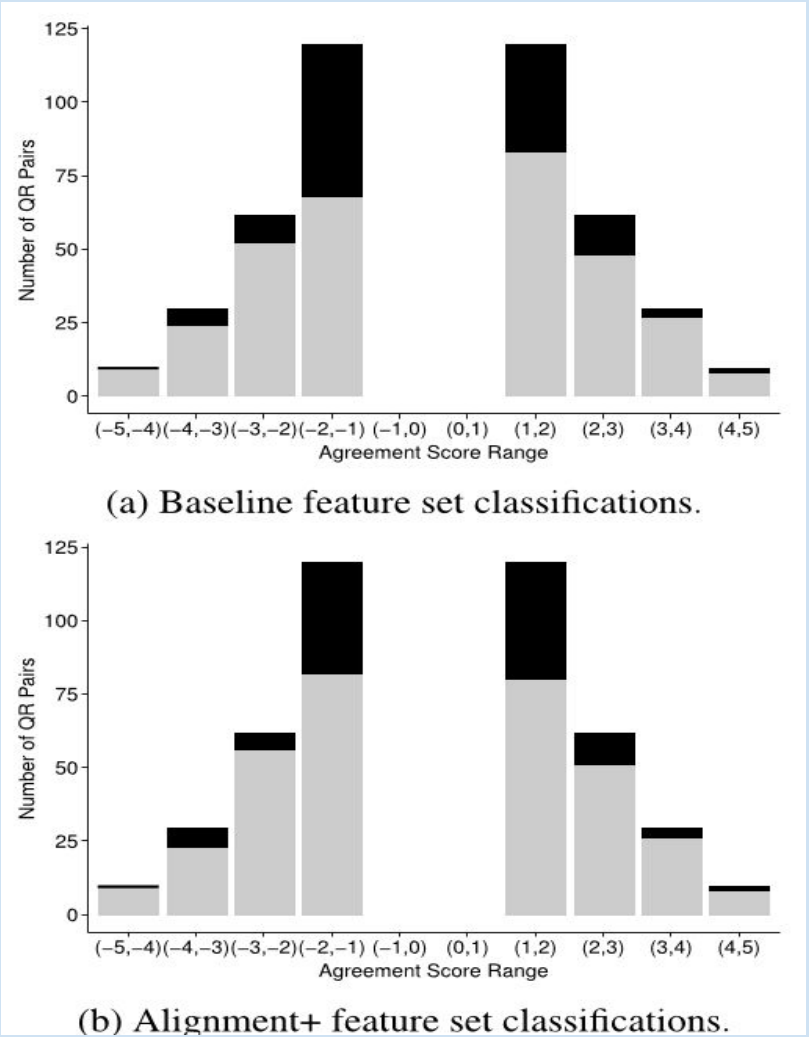


Figure 3: Correct and incorrect classifications on the test set given the corpus agreement scores, for both feature sets. The gray area represents correct classification, while the black area represents incorrect classification.