Influence and Belief in Congressional Hearings

Micah Clark, Adam Dalton, Tomas By, Yorick Wilks
IHMC
15 SE Osceola Ave., Ocala FL 34471
{mclark,adalton,tby,ywilks}@ihmc.us

Samira Shaikh, Ching-Sheng Lin, Tomek Strzalkowski
SUNY – University at Albany
1400 Washington Ave., Albany NY 12222
{samirashaikh,cslin612}@gmail.com
tomek@albany.edu

Abstract

We report on an exploratory analysis of the 2014 PoliInformatics congressional hearing corpus using CUBISM, a system for the analysis and deep understanding of multi-participant dialogues. CUBISM brings together two typically separate forms of discourse analysis: semantic analysis and sociolinguistic analysis. In the paper proper, we describe CUBISM and illustrate some of the major analytical components. Then, we explain how we use sociolinguistic phenomena to guide the extraction of potentially interesting beliefs. Finally, we discuss how the PoliInformatics corpus poses certain analytical challenges because of the motivations and preparations of congressional hearing participants.

1 Introduction

We hypothesize that integration of sociolinguistic analysis and semantic content analysis can enable the (semi-)automatic detection of belief-related phenomena in multiparty conversations. We analyze congressional hearings and discuss that while the format of the hearings lends itself to our approach, the political context poses additional challenges not commonly seen in informal dialogues. The belief-related phenomena of interest to us include such things as meaningful shifts in dispositions toward topics of discussion (e.g., shifts in sentiment) and toward other participants (e.g., changes in social roles) as well as changes in participants' attitudes (e.g., beliefs, intentions) about discourse topics and other participants. Such analysis may ultimately lead to detectible signatures for things like intentional deception, pandering, and successful persuasion. The above hypothesis is the basis for the CUBISM (Conversation Understanding through Belief Interpretation and Sociolinguistic Modeling) dialogue analysis system we are developing as part of DARPA's Deep Exploration and Filtering of Text (DEFT) program.

The 2014 PoliInformatics Unshared NLP Task using a corpus made up of data about the U.S. financial crisis of 2007–8 affords an opportunity to test our hypothesis and to evaluate the viability and efficacy of the analytical techniques built into CUBISM thus far. Initial results of our evaluation against congressional hearings contained in the PoliInformatics 2014 corpus are reported in the present paper.

The remainder of this paper is organized as follows. In §2 we give a highly condensed overview of CUBISM. In §3 we describe the analysis processes we applied to the transcribed congressional hearings. In §4 we discuss analysis results. Finally, in §5 we conclude by commenting on the relationship between the task and our methods of analysis.

2 Synoptic Description of CUBISM

CUBISM brings together research on dialogue understanding along two analytical dimensions: (i) participants' social roles and relationships, and (ii) participants' attitudes and beliefs about the world and each other.

With respect to social roles and relationships, such information is latent in dialogue and derivable via sociolinguistic features like topic, sentiment, and notions of "distance" between participants (Strzalkowski et al., 2013).

With respect to participants' beliefs (and other propositional attitudes), CUBISM uses a variant of the ViewGen paradigm (Wilks and Ballim, 1989) for modeling the interrelated viewpoints of multiple agents. Utterances that might express beliefs or other attitudes are identified based on dialog act information and the presence of certain verbs with modal relevance (e.g., believe, intend). The objects of belief are extracted as logical formulae. After participant viewpoints are populated with the extracted, explicit expressions of belief, the viewpoints are enriched via application of pragmatics and implicatures. At that point, CUBISM can use ascription and machine reasoning to contrast participant viewpoints, or track participant's beliefs over time.

Sociolinguistic, semantic, and linguistic data are integrated together, and with background knowledge, in a shared database. The database acts as a common repository, linking elements and annotations across analytical dimensions and levels of abstraction, and mediates information sharing with other DEFT software.

3 Analytical Methods

3.1 Sociolinguistic Analysis

We employ *topical positioning* to gain sociolinguistic insight into speakers' behaviors. Topical positioning is defined as the (dispositional) attitude a speaker has toward the meso-topics of discussion. In turn, *meso-topics* are defined as the most persistent topics of discourse—topics widely cited through long stretches of dialogue. When discussing issues (especially issues of controversy), speakers express and establish their attitudes toward topics, classified here as for, against, or neutral/undecided. In so doing, speakers actively shape the agenda and outcomes of the discussion. Quantifying topical positioning allows us to identify speakers who are for, against, or neutral on a given topic or issue.

To quantify topical positioning, we first identify meso-topics (Shaikh et al., 2012). Then, for each utterance made by a speaker regarding a meso-topic, we determine the polarity, i.e., if the utterance is for (positive), against (negative), or neutral on the meso-topic. We distinguish three forms of meso-topic valuation: (a) *express advocacy/disadvocacy*, when the valuation is applied directly to the topic (e.g., "I'm for regulation."); (b) *supporting/dissenting information*, when the valuation is made indirectly by offering additional information about the topic (e.g., "He has experience with collateralized debt obligations."); and (c) *express agreement/disagreement* with a polarized statement made by another speaker.

Two measures of topical positioning are defined: *Topic Polarity Index*, which establishes the polarity of a speaker's attitude toward the topic, and *Polarity Strength Index*, which measures the magnitude of this attitude.

Topic Polarity Index (TPX). To detect the polarity of topical positioning on meso-topic *T*, we count for each speaker:

- All utterances about *T* using statements with polarity *P* applied directly to *T* using appropriate adverb or adjective phrases, or when *T* is a direct object of a verb;[1]
- All utterances that offer information with polarity *P* about topic *T*;
- All responses to other speakers' statements with polarity *P* applied to *T*.

From the above, we calculate TPX for each speaker as a proportion of positive, negative, and neutral polarity utterances made by the speaker about *T*. A speaker whose utterances are overwhelmingly positive (above 80%) has a pro-topic position (TPX = +1); a speaker whose utterances are overwhelmingly negative takes an against-topic position (TPX = –1); a speaker whose utterances are either generally neutral or vary in polarity, has a neutral topic position (TPX = 0).

Polarity Strength Index (PSX). The strength of topical positioning is calculated as the proportion of utterances on the topic made by each speaker to all utterances made about this topic by all speakers. (Speakers, who make most utterances on the topic relative to other speakers, take a stronger position on this topic.) PSX is measured on a 5-point scale corresponding to the quintiles in normal distribution.

Topical Positioning Measure (TPM). To establish the value of topical positioning for a given meso-topic, we multiply TPX by PSX. For example, a speaker who makes 25% of all utterances on the topic "regulation" (group mean is 12%) and whose most statements are positive, has the strongest pro topical positioning on regulation: +5 (for fifth quintile on the positive side).

Distance between speakers on a meso-topic, as well as across all meso-topics, is calculated using the cosine between vectors of speaker TPM values. Using this notion of distance, we can detect opinion shifts and model the impact of speakers

[1] Polarities of adjectives and adverbs are taken from the expanded ANEW lexicon (Bradley and Lang, 1999).

with specific social roles in a dialogue. For example, an influencer is a participant who introduces ideas that others adopt or support. An influencer model is generated from mid-level sociolinguistic behaviors, including Topic Control, Disagreement, and Involvement (Shaikh et al., 2012). To detect and calculate the effect of an influencer, we track *changes* in the distances between speakers, e.g., if participants move closer to, further from, or both (e.g., polarization), some particular speaker(s).

3.2 Semantic Content Analysis

We extract dialogue act information and semantic content in the form of logical formulae (currently OWL) from utterances. Our basic approach to semantic content extraction is closer to that of Information Extraction (Gaizauskas and Wilks, 1997) than to traditional NLP approaches based on independent syntactic and semantic analyses. We are currently experimenting with several NLP pipelines (each based on a different NLP toolkit) for extracting semantic content from utterances. For the present effort, we used a pipeline based on the GATE toolkit[2]. This pipeline is a combination of the GATE ANNIE information extraction system, a dialogue act tagger, and an RDF triple extractor[3]. The triple extractor was designed to select only the main content of the sentence, which is appropriate for an RDF representation.

We also extract information regarding expressed propositional attitudes and the individuals to whom they belong. This information is extracted from verbs with modal relevance, coreference, and entity linking; it takes the form of a sequence of normalized agent and attitude pairs that represents the nested scoping of attitudes. Agent and attitude information is used to attribute extracted OWL formulae as the appropriate attitude held by the appropriate individual (including individuals mentioned in, but not part of, the dialogue). Thus, we handle nested attitude reports such as, "I think you want Carla to get the job."

Extracted semantic content populates a variant of the ViewGen system (Wilks and Ballim, 1989; Ballim and Wilks, 1991), which we have extended to propositional attitudes in general.[4] Using ViewGen, we are able to represent and reason over arbitrarily nested agent viewpoints (e.g., what *X* believes that *Y* intends for *X* to believe). The material content of an agent's beliefs, desires, etc. are categorized according to the meso-topic(s) of the source utterances, thus linking semantic content to sociolinguistic indices.

Our ViewGen variant supports three types of reasoning: (1) Rule-based pragmatic reasoning. For example, an explicitly expressed desire such as, "I wish the market wasn't dropping," warrants attribution to the speaker a belief that the market is dropping. (2) Default ascription, where each agent ascribes its own beliefs to others unless there is evidence to the contrary (evidence such as introduction of a new contradiction in belief). (3) Logical inference within a viewpoint using background knowledge.[5] These reasoning mechanisms are the building blocks of higher-level detection and analysis algorithms (such as tracking changing beliefs).

4 Initial Results

Our analysis focused on the congressional hearing transcripts on Monetary Policy, TARP, Dodd-Frank, and fifteen others related to financial reform. Each transcript was converted to collections of XML <turn> elements with attributes for turn number and speaker name. The resulting XML documents were then concurrently processed by our sociolinguistic and semantic content analyzers and their results stored in a central database. A high-level overview of extraction statistics is shown in Table 1.

As mentioned in the introduction, our modest aim was to explore the potential and efficacy of our hybrid (sociolinguistic and semantic) approach to dialogue understanding on "real data," rather than an attempt at full-scale automatic analysis of the corpora. (Our initial results are encouraging and we will continue to explore the congressional hearing corpus.) With that in mind, some of the more interesting initial results are the extremely high sentiment of Mr. Rush toward "the chair" (i.e., himself) in congressional hearing CHRG-111hhrg67816. Mr. Rush referred to himself in the third person as *the chair* 36 times in total, and only one reference could be deemed to have a negative sentiment, compared to 28 that were positive. This is perhaps an idiosyncra-

[2] http://www.gate.ac.uk
[3] Extracted triples are mapped to OWL ObjectPropertyAssertions; a more complex mapping is in development.
[4] The original ViewGen system only deals with beliefs.
[5] Of course, the quality and value of logical inferences depends on the availability of shared background knowledge. The difficult problem of acquiring such knowledge (especially in an ever-changing world), is being tackled by our collaborators at the University of Florida.

sy of parliamentary procedure but may have interesting sociolinguistic ramifications.

Table 1

Number of Hearings	113
Number of Participants	628
Number of Turns	38057
Unique Meso-topics	4419
Polarized Statements	
…in support of a topic	114121
…in opposition to a topic	84006
…neutral on a topic	121246
Beliefs extracted from statements	332735

On the other end of the sentiment spectrum, Senator Levin mentioned "standards" (credit, public, and other) 22 times in congressional hearing CHRG-111shrg57321, each time with a negative sentiment.

Finally, we identified *system* (i.e., the financial system) as a topic on which Mr. Geithner shifted on over the course of CHRG-111hhrg54867. Mr. Geithner began the hearing by speaking positively or neutrally about *system* with 8 out of 9 utterances. His final 30 utterances on *system* included only one statement of positive sentiment.

Reasoning over beliefs and detecting changes to beliefs is computationally expensive and potentially unbounded in the general case. Our stratagem is to use detectible changes in sociolinguistic features to focus belief analysis on topics and dialogue segments of greatest interest. Specifically, shifts in sentiment can indicate internal conflicts in belief or revision of belief, and we have some preliminary evidence that such shifts prefigure future changes in belief. (It is too early to pass judgment on this stratagem, as our analysis of explicit changes in belief is still underway.) Our belief extraction process does appear to give good results. For example, some of the beliefs extracted then attributed to Mr. Geithner regarding the financial system include following (in English, rather than OWL).

- The [financial] system has changed a lot;
- The [financial] system has strengths;
- The [financial] system is overseen by a patchwork diffused across agencies;
- The [financial] system is global;
- The [financial] system is through a period.

All of these are reasonable to attribute to Mr. Geithner based on his statements, except perhaps the last wherein extraction did not preserve the adjective qualifying "period."

5 Concluding Remarks

The unstructured task, directed at the economic crisis of the last decade, would require for its solution a causal explanation, and on an issue about which there is still strong difference of opinion: Was it the bankers who did it, or the politicians, for example? The task's performance might require something with the same power and scope in social science as DARPA's future Big Mechanism project on scientific explanation. To this point, texts of the kind used in the task, prepared speeches and responses in semi-dialogue form, might well not be the most appropriate texts to search for explicit events and claims to form part of a coherent causal explanation of the crisis. The reason for this is that political speeches are not normally directed at revelation and truth but at concealment, justification and the evasion of blame (though they may still reveal something about the supposed mind-sets of the individuals involved in the crisis).

However, and that said, it is worth an attempt to see if a system like CUBISM, designed to detect and link semantic and social content associated with beliefs in normal dialogue, could also pick up something of interest in these materials. Future work can then take into account additional sources such as press releases or speeches with votes on committee or on the floor. Future analysis could also occur over a longer time span in order to cover changes in a political actor's role. Our key insight is that quantitative social metric computations can help filter areas of dialogue to which more intensive qualitative belief and content computations can be subsequently applied.

Acknowledgements

This work was supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. FA8750-12-2-0348.

References

Afzal Ballim and Yorick Wilks, (1991). Artificial Believers: The Ascription of Belief. Hillsdale, NJ: Lawrence Erlbaum Associates.

Margaret M. Bradley, and Peter J. Lang. (1999). Affective Norms for English Words (ANEW): Stimuli, instruction manual, and affective ratings (Tech. Report C-1). Gainesville: University of Florida, Center for Research in Psychophysiology.

Robert Gaizauskas and Yorick Wilks. (1998). Information extraction: beyond document retrieval, *Journal of Documentation*, 54(1): 70–105.

Samira Shaikh, Tomek Strzalkowski, Jenny Stromer-Galley, Geoge Aaron Broadwell, Sarah Taylor, Ting Liu, Veena Ravishankar, Xiaoai Ren and Umit Boz. (2012). Modeling Influence in Online Multi-Party Discourse. In *Proc. of 2nd Int. Conf. on Social Computing and Its Applications* (SCA 2012), Xiangtan, China.

Tomek Strzalkowski, Samira Shaikh, George Broadwell, Jenny Stromer-Galley, Sarah Taylor, Umit Boz, and Xiaoai Ren. (2013). Influence and Power in Group Interactions. In *Proc. of the 2013 Int. Conf. on Social Computing, Behavioral-Cultural Modeling, & Prediction* (pp. 19–27), LNCS Vol. 7812, Springer.

Yorick Wilks. and Afzal Ballim. (1989). Shifting the belief engine into higher gear. In *Proc. of the Int. Conf. on AI Methodology Systems Applications* (pp. 11–20), Elsevier.