DISCOVERING SOCIO-CULTURAL PHENOMENA IN DISCOURSE

Tomek Strzalkowski[^], George Aaron Broadwell[^], Jennifer Stromer-Galley[^], Samira Shaikh[^], Sarah Taylor[*], Umit Boz[^], Alana Elia[^], Laura Jaio[^], Ting Liu[^] and Nick Webb[^].
[^]ILS Institute, University at Albany, SUNY
[*]Lockheed Martin Corporation

ABSTRACT

We describe a novel approach to computational modeling and understanding of social and cultural phenomena in multi-party online dialogues. We developed a two-tier approach in which we first detect and classify *social language uses* (LU) in discourse, including *topic control*, *task control*, *disagreement*, and *involvement*. These languages uses are the socio-linguistic devices deployed by discourse participants in order to construe certain higher-level *social constructs* (SC), such as leadership or group stability. In our approach, LU's are detected per participant by tracking observable and computable linguistic features in speakers' utterances, such topical references, dialogue acts or sentiment markings. Combinations of LU assignments yield models for assessing social constructs. We developed an automated system DSARMD-1 that can compute language use assignments with high degree of accuracy, and can make initial predictions of social constructs. This paper focuses on English online chat conversations; however, our approach is equally applicable to other conversational situations: informal face-to-face interactions, formal meetings, moderated discussions, as well as interactions conducted in other languages, e.g., Urdu and Mandarin. The range of applications for this technology is also very broad, including, among others, intelligence, influence operations, training, human effectiveness and crime prevention.

PRIMARY TRACK

Application of Social Cultural Methods, Models, and Tools (MMT)

SECONDARY TRACK

Understanding and Modeling Human Behavior

DESCRIPTION

Social interaction in an increasingly online world is an important resource through which we can study and understand the vast landscape of human terrain, across cultures and languages. In a reduced-cue environment, such as online chat, interaction necessitates the deployment of more explicit linguistic devices to convey social and cultural nuances than is typical in face-to-face or even voice conversations. The resulting highly informal dialect presents additional challenges in understanding social dynamics among the participants.

Our objective is to develop computational models of how certain social phenomena such as leadership and group stability are signaled and reflected in language through the choice of lexical, syntactic, semantic and conversational forms by discourse participants. Given a representative segment of multiparty task-oriented dialogue, our prototype system called DSARMD-1 (Detecting Social Actions and Roles in Multiparty Dialogue) performs the following: (1) automatically classifies all discourse participants by the degree to which they deploy selected *language uses* (*LU*), such as topic control, task control, involvement, and

disagreement; (2) based on the distribution of *LU*'s among the participants, predicts emergence of higher-level *social constructs*, including leadership and group stability [1], [2], [3]. In this work we adopted a two-tier empirical approach where social language uses are modeled through *observable* linguistic features that can be automatically extracted from dialogue. The high-level social constructs are then inferred from a combination of language uses attributed to each discourse participant [4]. For example, a high degree of influence and a high degree of involvement by the same person may indicate a leadership role.

We conducted a series of experiments in which recruited subjects were invited to participate in a series of on-line chat sessions in a specially designed secure chat-room. The experiments were carefully designed around topics, tasks, and games for the participants to engage in so that appropriate types of behavior, e.g., disagreement, power play, persuasion, etc. may emerge spontaneously. The resulting multi-party chat corpus [5] was annotated at four levels: (1) Communication links – which determine who is speaking to whom or what the threads of conversations are; (2) Dialogue acts – a hierarchy of 15 dialogue acts tuned significantly towards dialogue pragmatics and away from surface characteristics; (3) Local topics – nouns or noun phrases introduced into discourse that are subsequently mentioned again via repetition, synonym, or pronoun; and (4) Meso-topics – topics that persist through a number of turns in conversation and are of significance as users make polarized statements about them.

The DSARMD-1 system comprises of a series of modules that create automated annotation of the source dialogue for all the language elements. These include, extraction of mentions of local topics, topic co-references, communicative links between utterances, classification of dialogue acts, identification or meso-topics, and assignment of polarity markers. We refer readers to [6] for a detailed description of the computation models involved in DSARMD-1. Automatically annotated dialogue is then used to compute index values from which LU degree claims are derived. We have a total of 16 indices, which map into a degree of language use, based on quintiles in normal distribution. For example, one hypothesis is that topic control is indicated by the rate of local topic introductions (LTI) per participant [6], [7]. Thus, one measure of topic control is the number of local topics introduced by each participant as percentage of all local topics in a discourse.

For example, in a multi-party discussion *dialogue-1* where 90 local topics are identified:

1. *Speaker LE introduces 23/90 (25.6%) of local topics in this dialogue.*
2. *The mean rate of local topic introductions is this dialogue is 14.29%, and standard deviation is 8.01.*
3. *LE is in the top quintile of participants for introducing new local topics.*

We can now claim the following, with a degree of confidence (to be determined):

$TopicControl_{LTI}$ (LE, 5, *dialogue-1*)

We read this as follows: speaker LE exerts the highest degree (5) of topic control in *dialogue-1*. Of course, LTI is just one source of evidence and we developed other indices to complement it.

Table 1 shows the performance of precision of DSARMD-1 system on two of our language uses, Topic Control and Involvement. Each participant in the dialogue is given a single score, based on combining the index measures of Topic Control and Involvement respectively. This is compared against human annotated dialogue. Based on our strict metric, for which each exact match is counted as 1 and mismatches are counted as 0, our performance is 71.43 on both Topic

Control and Involvement. An alternative binary classification where scores 5 and 4 are considered High, while the remaining scores are considered Low is also of interest. Under this metric, the Topic Control and Involvement score quite high.

TABLE 1. Performance of DSARMD-1 against human annotated dialogue

	Topic Control		Involvement	
Participants	system	human	system	human
JR	**5**	**4**	3	**4**
LE	**4**	**5**	**4**	**4**
KN	3	3	**5**	**5**
KI	**4**	**4**	3	2
CS	2	2	**4**	**4**
KA	2	2	2	2
JY	1	1	1	1
System performance				
Strict	71.43		71.43	
High-Low	100		85.71	

The DSARMD-1 system was also tested over other types of data, including face-to-face meetings and dialogues in virtual world games, such as Second Life. For example, in Second Life experiments, participants filled in exit questionnaires assessing their leadership qualities displayed in the game just completed. These questionnaires were converted into leadership ranking and compared to the ranking computed by DSARMD-1, based on the chat communications between participants. Using only the Topic Control measures, the system was able to achieve 71% prediction accuracy.

The work reported here is still in the early stages, as many challenges remain, including more dependable extraction of basic discourse features upon which the language use indicators are built: topic references, dialogue act classification, polarity assignments, part-of-speech tagging, among others. Nonetheless, the current prototype already shows a remarkable accuracy at picking out key players in group interactions, on-line as well as in traditional meetings.

ACKNOWLEDGEMENTS

This research was funded by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), through the U.S. Army Research Lab. All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of IARPA, the ODNI or the U.S. Government.

BIOGRAPHY

Dr. Tomek Strzalkowski is Director of the Institute for Informatics, Logics and Security Studies (www.ils.albany.edu) and an Associate Professor of Computer Science at SUNY Albany. Prior to joining SUNY, he was a Natural Language Group Leader and a Principal Scientist at GE

R&D, and before that a faculty at the Courant Institute of New York University. He has done research in computational linguistics, including socio-cultural aspects of language, information retrieval, human-computer dialogue, formal semantics, and reversible grammars. He has directed research projects sponsored by IARPA, DARPA, EC and NSF. Dr. Strzalkowski has published over a hundred scientific papers, and is the editor of several books, including *Advances in Open Domain Question Answering*.

REFERENCES

[1] Fairhurst, G. 2007. Discursive leadership: In conversation with leadership psychology. Los Angeles, SAGE Publications.

[2] Ellis, D. G. and B. A. Fisher. 1994. Small group decision making: Communication and the group process. New York, McGraw-Hill

[3] Bales, R. F. 2001. Social Interaction Systems: Theory and Measurement. Transaction Publishers.

[4] Broadwell, G. A et al. 2010. Social Phenomena and Language Use. ILS Technical report. Available at http://ils.albany.edu/tech_reports.html

[5] Shaikh S. et al. 2010. MPC: A Multi-Party Chat Corpus for Modeling Social Phenomena in Discourse, Proc. LREC-2010, Malta.

[6] Strzalkowski T. 2010 et al. 2010. Modeling Socio-Cultural Phenomena in Discourse. Proceedings COLING-2010, Beijing, China.

[7] Givon, Talmy. 1983. Topic continuity in discourse: A quantitative cross-language study. Amsterdam: John Benjamins.