Computational Predictors in Online Social Deliberations

Woolf, Beverly,[1] Murray, Thomas,[1] Xu, Xiaoxi [1]

Osterweil, Leon,[1] Clarke, Lori,[1] Wing, Leah [2] Katsh, Ethan [2]

[1]Department of Computer Science; [2] Department of Legal Studies

University of Massachusetts-Amherst

(bev, murray, xiaoxi, ljo, clarke@cs.umass.edu); lwing@legal.umass.edu; ethan.katsh@gmail.com

Abstract

This research seeks to identify online participants' disposition and skills. A prototype dashboard and annotation scheme were developed to support facilitators and several computational predictors were identified that show statistically significant correlations with dialogue skills as observed by human annotators.

Detecting Communication Skills.

This paper describes efforts to identify time-based trends and interaction patterns in online deliberations and to evaluate participants' disposition state and communication skills. Deliberation is when people carefully examine a problem and arrive at a "well-reasoned solution after a period of inclusive, respectful consideration" (Gastil & Black, 2008). The Internet is not always a harmonious place; it fosters large numbers of transactions and interactions and inevitably a percentage of these interactions encounter problems (Katsh & Wing, 2006). A growing need for novel dialogue, deliberation and dispute resolution processes has become apparent. The current study seeks diagnostic predictors using text analysis tools (e.g., Coh-Metrix and LIWC) to structure dialogue and provide useful feedback to human facilitators. We tested the hypothesis that a statistically significant correlation exists between *computational features* and *annotation-based skills* that might help facilitators predict the state and skills of deliberators. In other words, computer generated measures of online linguistic and psychological features were correlated with human annotations of the same dialogue. One goal is to see which annotated states and skills we can train machine learning algorithms to categorize with reasonable accuracy in support of effective dialogue.

Figure 1. The Draft Facilitator Dashboard.
The dashboard records deliberations (top, left) and text analysis tools (top right) suggest participants' skills along with above threshold activities that may require attention (bottom row).

A facilitator dashboard (Figure 1) developed at UMass supports the definition, execution and analysis of social deliberation. It records participants' online discourse through MEDIEM[1] (top left) and displays text analysis alerts (top right). The dashboard will be used only by facilitators and will ultimately indicate when dialogues are proceeding well or when disputants are in trouble and an intervention (e.g., private caucus, verbal response) might help.

Online Dispute Resolution and "Fourth Party"

Through the generosity of collaborators like JuriPax (Workplace Grievances) and Net Neutrals (eBay)[2] we acquired authentic online deliberations. Online participants

[1] Mediem is available at http://idealogueinc.com/contact.html

[2] http://www.juripax.com/ and http://netneutrals.com/

often discuss different options, make decisions, collaborate and in general go through processes that are analogous to face-to-face deliberative discussion (Black et al., 2011).
Mediation, in which parties resolve their problem under the guidance of an acceptable third party, is an alternative way to solve conflicts. Online Dispute Resolution explores ways in which computer and communication technologies can facilitate dispute resolution while also decreasing the degree of human involvement (Katsch, 1996).[3] The "third party" delivers instructions, questions, and other types of messages that resembled the manner in which a face-to-face third party would interact with disputants. Technology is seen as a "fourth party" that focuses attention on capabilities that machines have and that humans might not, e.g., to model and monitor dialogue and provide data (Katsh & Rifkin, 2002; Katsh b& Gaitenby, 2003).

Annotation	Deliberation Behavior
Question topic	Ask questions to discover more.
Reflect back	Explicit reflection of words or ideas
Back Reference	Reference what another said, including quoting, summarizing.
Solution	A proposed solution, or suggest what participants should do
Negative emotion	Negative emotion about interlocutors, or dialog process or topic; disrespect, insult; anger.
Mediate	Suggestions about how interlocutors should communicate.
Meta statement	Highlight or summarize, e.g., consensus, disagreement.

Table 1. Social Deliberative Coding Scheme

A sample of the 46 annotations for online deliberation, some correlate with computational predictors (bold).

1. Research Method

We analyzed the Workplace Grievance and Faculty Deliberations described in Table 2. These dialogues were no longer being argued, meaning that they did not change during the time of our analysis, making the annotation process manageable.

During the **Annotation Phase** we used a newly developed Social Deliberative Coding Scheme, see Table 1 (Murray et al., 2012) to assess how well the parties achieved problem-solving tasks (analytic dimensions) and then analyzed aspects of relational communication (social dimensions) (Black et al., 2011). We collected 202 annotated segments from the Faculty and 217 from the Workplace discourse and performed an inter-rater reliability test for two human annotations, using Cohen's Kappa statistics. The fair inter-rater agreement (around 25%) highlights the difficulty of annotating deliberative skills from online disputation.

[3] Hundreds of thousands of disputes are handled online (see http://odr.info/nodr/32 and http://www.odr.info/node/78).

During the **Analysis Phase** each discourse was further analyzed using software packages, Linguistic Inquiry and Word Count (LIWC) and Coh-Metrix, to identify linguistic and rhetorical relations. LIWC (Pennebaker and Francis, 1999) reports the percentage of words in a given text devoted to grammatical (e.g., articles, pronouns), or content categories (e.g., home, occupation, religion). Coh-Metrix provides an array of measures based on levels of words, syntax, situation model, rhetorical structure, genre and pragmatic communication (Graesser et al., 2004; McNamara et al., 2006).

Workplace Grievance: 524 segments, 60 posts, 2 parties, 1 mediator

The director of a company can no longer work with an employee. She says he must be terminated and she can't come to a solution with him. The employee says his boss wants to get rid of him and the place makes him sick. A human mediator supports both parties.

Example Dispute:

"There are numerous examples where the employee loudly disagreed with proposed measures." (*Provide evidence*.)

"Ryker does not need to leave, but agreements need to be made about how to proceed further." (*State issues*)

Faculty Deliberation: 507 segments, 71 posts, 15 parties, no mediator

Faculty from one academic community scheduled a conference that interferes with another conference. The dialogue involves both communities to decide whether to reschedule the conference. By the end of the discussion four people have resigned.

Example Dispute:

"I'm pleased to see this initiative." (*Acknowledge*)

"I have no desire to be ruled." "We do not want to be a part of that type of community." (*Negative emotion*)

"I am very disappointed in your decision;" "I will elect to spend my energies elsewhere and resign." (*Negative emotion*)

Table 1. Authentic online disputes and human annotations.

When participants are in agreement, more informal language may be used with features associated with every day oral conversation as in the beginning of the Faculty Dialogue when participants were appreciative and hopeful. By the end of the dialogue, many people had resigned and the rate of emotion, self-reflection and meta-summary had increased. At the beginning of the conversation, Coh-Metrix detected high narrativity correlated with participants questioning the topic and providing acknowledgement. However, when the conversation became difficult, as in the case of participants not agreeing, the language shifted to being more stilted with a serious slant and persuasion pitch.

During the **Correlation Phase**, we evaluated the hypothesis that features from LIWC and Coh-metrix will predict online participants' skills as identified by human annotators, Figure 2. The computational discourse features included 25 features selected from 80 LIWC indices and 8 features from over 100 Coh-Metrix indices. We carried out a correlation analysis between text analysis features and

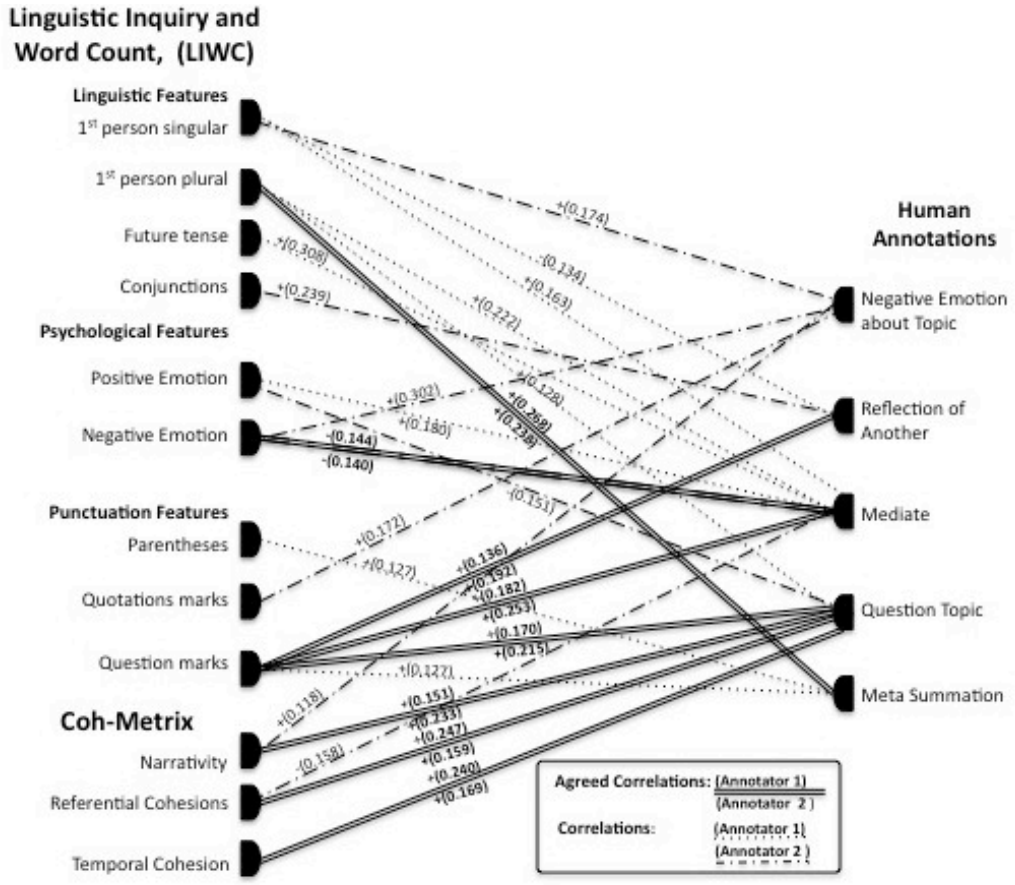

Figure 2. Workplace Dispute. Statistically significant correlations between text analysis (left) and human annotations (right) are indicated. Agreed correlation values are listed on the top/bottom of each double line and single correlations by dashed lines.

each set of human annotations, in part because of the low inter-rater agreement between annotators. Using annotations from a single annotator to identify correlations can sometimes find significant correlations by chance. However, if significant correlations were present between discourse features and both sets of annotations, we would have more confidence to claim that those significant correlations are not due to sampling errors and judgment biases. We call these correlations *agreed correlations*.

Results

We calculated statistically significant correlations between features detected by the text analysis tools and human annotations at the confidence level of 95%. We used a P value of 0.05 for the one-tailed significant test, meaning that we are 95% confident about rejecting the null hypothesis (i.e., no significant correlations exists). In Figure 2, double lines indicate *agreed correlations,* in that both inter-rater agreement between annotators agreed about the annotations and dashed lines indicate significant correlations, but only for a single annotator. Text analysis features for the Workplace Deliberation (Figures 2, left) can be seen as diagnostic predictors for the corresponding human annotated skills (right). In both Workplace and Faculty dialogues we found agreed correlations between discourse features and the human annotations of *negative emotion about topic*, *reflect others*, *mediate* and *question topic*.

In the Workplace Deliberation, Figure 2, we found significant agreed correlations between text analysis results and the human annotations of ***reflection of another*** (right), which were negatively correlated with *question marks*. Question marks are rarely used to quote and summarize messages from others. The human annotations of ***mediate*** were negatively correlated with *negative emotion* and *question marks*. Professional mediators are trained to be

neutral and impersonal and would use few negative emotions. The human annotations of **question topic** were positively correlated with four predictors, *question marks*, *narrativity*, *referential cohesion*, and *temporal cohesion*. The characteristics of narrativity are often present in oral language where sentence constructions are easy for the audience to comprehend, such as shorter noun phrases, fewer words before the main verbs of main clauses, and fewer passive constructions. Referential cohesion characterizes to what extent content words and ideas are connected with each other. Questions about topics inevitably reference back ideas mentioned previously and thus can have a positive correlation with referential cohesion and temporal cohesion. And finally the human annotations of **meta-summation** were positively correlated to *1st person plural*; clearly the word "we" is used when participants summarize a consensus or conflict.

Agreed correlations in the Faculty Deliberations (not shown) found significant correlation between ***negative emotion about topic***, which were positively correlated to *negative emotion* and *tentative*, and *quotation marks*. Negative emotion, uncertain statements, and referencing back are likely to be used by participants when they present negative accounts of a topic. The human annotations of ***reflection of another*** were positively correlated to *conjunction*, which is frequently used to quote and summarize what others say. The human annotation of ***mediate*** was positively correlated with *1st person plural* and *cognitive process*. Mediators inevitably use the word "we" and employ cognitive processes to learn the conflicts of disputant parties and make suggestions about how a conversation should proceed.

The agreed correlations described above are domain dependent, with only one predictor (question marks) positively correlated with question topics in both domains. We will use datasets from more domains to see if this phenomenon still persists.

Discussion and Future Work

We found significant correlation between computational deliberation predictors and human annotations for the same dialogue. Our intention is to integrate such text analysis tools into a real-time dashboard that might enable facilitators to manage several deliberations and gain a high level view of the nature of the discussion. One goal is to provide a high level view of the nature of deliberation and to deliver alerts when deliberations are not proceeding well and when interventions might be useful. We have tested the effectiveness of a mock-up dashboard in a Wizard-of-Oz experiment with college students (Murray et al., 2012).

Acknowledgments

This research was funded by an award from the NSF #0968536, The Fourth Party: Improving Computer-Mediated Deliberation through Cognitive, Social and Emotional Support, Beverly Woolf (PI), Lori Clarke, Leon Osterweil, Thomas Murray and Ethan Katsh. Any opinions, findings, conclusions or recommendations are those of the authors and do not necessarily reflect the views of the funding agencies.

References

Black, L., Welser, H., Cosley, D., and DeGroot, J., (2011) Self-Governance Through Group Discussion in Wikipedia Measuring Deliberation in Online Groups, *Small Group Research* 2011 42

D'Mello, S., Dowell, N., and Graesser, A., Cohesion Relationships in Tutorial Dialogue as Predictors of Affective States.

Gastil, J., & Black, L. W. (2008). Public deliberation as the organizing principle in political communication research. *Journal of Public Deliberation, 4*.

Graesser, A.C., McNamara, D.S., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36*, 193-202.

Katsh E. and Wing L. (2006) Ten Years of Online Dispute Resolution: Looking At The Past and Constructing The Future, 38 U. Tol. Law Rev. 19 (2006-2007)

Katsh, E. and Rifkin (2002) J., *Online Dispute Resolution: Resolving Disputes in Cyberspace*. San Francisco: Jossey-

Katsh, E., and Gaitenby A. (2003) Introduction: Technology as the "Fourth Party" Proceedings of the UNECE Forum on ODR 2003 http://www.odr.info/unece2003.

Katsh, E., (1996) *Dispute Resolution in Cyberspace*, 28 CONN. L. REV. 953 (1996); E. Casey Lide.

McNamara, D., Ozuru, Y., Graesser, A, Louwerse M., (2006) Validating Coh-Metrix. In R. Sun & N. Miyake (Eds), Proceedings of the 28th Annual Conference of the Cognitive Science Scociety (pp. 573). Mahwah, NJ: Eribaum.

McNamara, D.S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology, 55*, 51-62.

Murray, T., Woolf, B.P., Xu, X., Shipe S., Howard S. and Wing, L. (submitted) Toward supporting social deliberative skills in online classroom dialogues and beyond, International Conference on Intelligent Tutoring Systems, June 2012.

Pennebaker JW, Francis ME. (1999) *Linguistic Inquiry and Word Count (LIWC)*. Mahwah, NJ: Erlbaum.