REU Supplement: Annotating and Recognizing Dialog Structure in Meetings

This is a request for a supplement to the award: “ITR/PE+SY: Mapping Meetings: Language Technology to Make Sense of Human Interaction,” award number 0121396. We first provide a summary of the funded project to provide a context for the proposed work, followed by a brief summary of the two projects involved, the student selection mechanism and other justifications.

Overall Project Summary

Meetings are an essential part of every enterprise. Unlike other communicative activities, however, meetings have experienced little impact from information technology developments. Combining verbal communication with group social dynamics provides a major challenge that could greatly benefit from automatic information processing systems. This work investigates a solution based on the metaphor of a “meeting map,” a structured representation that supports the presentation of multiple views of a meeting at different scales; its primary purpose is to aid in navigation of the represented material. The project will focus on two broad map categories: content maps, portraying topics discussed and decisions made; and interaction maps, identifying the roles and relationships of the participants and the level of concurrence. Building content and interaction maps will involve automatic classification of information from topic changes and salience to disagreement/consensus. These maps will be used for generating simple indicative summaries, and off-the-shelf visualization tools will be used for map presentation. The project will build on analyses of 100 hours of meetings. Evaluations will use objective recognition accuracies and expert assessments of automatic summaries. Meeting maps respect the diversity of information present in meeting scenarios, and should greatly impact technology for human-to-human interaction support.

Proposed REU Project Descriptions

Project 1: Recognizing Consensus and Disagreement in Meetings (University of Washington)

An important sub-problem that we have identified for building meeting maps is automatic detection of consensus and disagreements, which can serve as a building block for topic segmentation and information extraction for summarization, as well as for analysis of social interactions. In a UW senior project, we have begun looking at the problem of automatic recognition of four classes of utterances: agreement, disagreement, backchannel and other. The student working on this project has shown that these simple categories can be reliably labeled (kappa statistic greater than 0.6), and that combining n-gram and other language cues in a decision tree classifier give performance significantly above chance. In addition, he has shown that unsupervised clustering techniques can be used to provide important language features for classification. By the end of the academic year, he hopes to improve results by incorporating turn-taking and prosodic cues.

The proposed REU Supplement will extend this work in two important ways. First, we plan to evaluate performance on recognizer outputs (the current work uses hand transcriptions) and (re)assess the relative importance of wording and prosodic cues using hypothesized word sequences and word confidences. Second, the work will further investigate unsupervised clustering and co-training techniques to minimize the need for costly hand labeling data of agreement/disagreement for training.

The student will be supervised directly by Prof. Ostendorf, but will also work with graduate students associated with this project. In addition, collaborators on this project from SRI and ICSI (particularly Elizabeth Shriberg) will be working with the student on the prosody modeling aspects.

Project 2: Dialog Structure Annotation and Modeling (ICSI)

A second, crucial area in which we would like to supervise undergraduate research, is the annotation and modeling of dialog structure in meetings. Dialog structure annotation will include the marking of “dialog acts”, which indicate the function of an utterance and its relation to others in the conversation. For example, common dialog acts include statements, questions, backchannels (like “uh-huh”), agreements, topic openers, closings, tangential
comments and jokes, side-talk, invitations, requests, and so on.

The hand annotations will allow us to train models that will use words as well as prosodic features, to automatically tag such dialog acts in new data. We will use the labeled data to explore to aid in mapping and summarizing the meetings. By modeling the dialog acts we can, for example, find topic boundaries, locations of decision points or “hot spots”, locations of tangential or side information, and so on. This work is related partly to the work proposed for the UW REU project above (Project 1), but differs in that rather than focusing on a particular type of dialog act pair (such as agreements), here we aim to label all utterances with a dialog act, to enable modeling of the entire meeting, at different levels of granularity.

In previous NSF work (and work at the 1997 Johns Hopkins Summer Workshop), we developed a method for annotating dialog acts in conversational data; this system was applied to the Switchboard corpus of two-party telephone conversations between strangers. It has also been used by other research groups around the world. The current REU project would extend the system to adapt it to meetings in two important ways. First, it would extend the dialog act categories to the case of multi-party, face-to-face meetings among participants who generally know each other. This difference in context gives rise to types of dialog acts that are not seen in Switchboard. Second, we will focus on annotation of interactions during simultaneous talk (or speaker overlaps), which, while present in Switchboard, were not previously modeled due to a focus on each individual channel. Rather than “linearize” a sequence of overlapping dialog acts, we will annotate them and model them taking into account their actual time relationships.

The work will be supervised directly by Dr. Shriberg at ICSI, and will involve a collaboration with Dr. Osten-dorf and students at UW. Dr. Shriberg has supervised five undergraduate students in other types of annotation tasks (including the annotation of emotion) previously at ICSI. ICSI provides an ideal environment for such projects, offering the students exposure to many of the other ongoing research projects in speech and language at the Institute.

Student Selection Method

We are requesting funding for three students, one at the University of Washington and two at ICSI. Project 2 includes two students, because in our experience it is crucial to have at least two different students on any large labeling tasks. This provides a means for assessing inter-labeler reliability, and for developing (via discussion) the extended labeling system.

We have already selected two of the candidates, Dustin Hillard (UW) and Raj Dhillon (ICSI, UC Berkeley); their resumés are included in the supplementary documentation. She has excellent background in linguistics, and a keen interest in the dialog annotation task, and would like to work on this project at ICSI this summer. As described above, Dustin has already been working on the consensus problem, and he has made good progress so far as well as contributed to this proposal. Both students are clearly outstanding candidates.

The third student (and other students in the unlikely event that either of the above two cannot accept the position) will be selected based on interest in the topic, academic standing, basic Unix experience, and on interviews with the senior investigators, after advertising to several departments in the associated universities (via undergrad mailing lists and notices to the undergrad research coordinators).

Budget Justification

The budgets include summer salary for the students and travel for two of the students, as described in the attached budget justifications. The travel will be used for students to attend the International Conference on Spoken Language Processing, which will be held in Colorado September 2002. There are currently no other funds for these students to attend this workshop.

We also note that this award was cut by 70% relative to the original proposal, so the REU supplement will provide a means of restoring the originally proposed involvement of undergraduates in research.
Previous PI Experience with Undergraduates

The PI, Professor Ostendorf, has supervised 9 senior projects and several additional undergraduate research students, including several women and two students from under-represented minorities. Her two previous REU awards (at Boston University) supported students Mark Paley, who went on to a very successful career in industry, and Jay Hancock, who is now a PhD student at MIT. Dr. Elizabeth Shriberg is a senior person on the project, and has supervised over a dozen students (many of them women) as indicated in the biographical sketch included with the main proposal.