

SESSION 5: DISCOURSE

Jerry R. Hobbs, Chair

Artificial Intelligence Center
SRI International
Menlo Park, California 94025

The fundamental problem in discourse is “What structure is there in discourse above the level of the sentence?” This question can be asked in terms of text versus dialogue. Is there a kind of structure that is exhibited in text but not in dialogue, or in dialogue but not in text? What are the appropriate structural descriptions of text and of dialogue?

The question can also be asked from the perspective of recognition and from the perspective of generation. There is the structure that the speaker puts there and the structure the hearer discovers there. Are these the same? Does one have primacy over the other? For example, does the speaker impose the structure so that it is the job of the hearer to discover it, no matter what it is, or is it necessary for the speaker to design his discourse in a way that makes it as easy as possible for the hearer to interpret?

The four papers in this session discuss various aspects of this family of questions.

The first paper, by Liddy and her colleagues at Syracuse University, is concerned with text rather than dialogue, and recognition rather than generation. Structure in text can be studied from a general point of view or from a genre-specific point of view. This paper examines the structure specific to the genre of newspaper articles. The structure found can be in the form of a hierarchy, a tree-like structure, much like the syntactic structure of sentences, or it can consist of a division of the text into segments of various kinds, performing various functions, such as the Lead, the Main Body, and so on. Liddy and her colleagues argue that the latter kind of structure is more appropriate for newspaper articles.

Among the questions addressed in this paper are

- What are the structural elements or possibilities?
- How can they be recognized, computationally?
- Once recognized, how can they be used?

The remaining papers are concerned with dialogue. The key idea in investigations of discourse from an artificial intelligence perspective is that the structure of a dialogue is, or is at least derived from, the structure of the participants’ plans. “Plan” here is meant in the AI sense of a hierarchical structure of causal relations, decomposing goals into subgoals, and these subgoals into further subgoals, and so on. An utterance in a dialogue is an action in a larger plan to achieve some goal. Generation is a matter of finding the right such actions. Recognition is a matter of discovering the role that action plays in the overall plan.

Moore’s paper examines a problem that arises in generation. We don’t like to be told too much. If we already know something, we don’t like to hear it again. If the speaker does not take this into account, the result is what my children used to refer to as “talking to me like I’m a retard.” It is also an inefficient way of conveying a plan or some other structured body of information. The problem that Moore addresses is “How should we use knowledge gained from the previous discourse to convey a plan as efficiently as possible?”

The paper by Ferguson and Allen take the perspective of recognition. Part of understanding an utterance is discovering the role it plays in the speaker’s larger plan. But the speaker’s plan, or intentions, are inaccessible to us. All we have to go on is what he or she said. From that we have to hypothesize a plan in which the observable utterances would make the most sense. This is the problem that Ferguson and Allen address: “How can we recognize the speaker’s plan, given only sparse information about it?”

In stretches of dialogue larger than a single utterance, neither participant is exclusively a speaker or exclusively a hearer. They are each executing their own plan, and their plans change as they are impacted by the actions of the other participant. In very focused, collaborative, problem-solving dialogues, they are working together to come up with a single plan for solving the problem. How can a single shared plan arise out of the interaction? This is the problem addressed in the paper by Biermann and

his colleagues at Duke University. Each participant has pieces of a plan; how can they be combined into a single, agreed-upon plan? The particular situation described is a very common one. One participant, in this case, the tutoring system, has control over general knowledge about the domain. The other participant, here, the user, has the knowledge of the specific problematic situation. How can these be combined into a solution of the problem?