SCoT: A Model of Conversational and Tutorial Intelligence

Brady Clark, Elizabeth Owen Bratt, Karl Schultz ({bzack,ebratt,schultzk}@csli.stanford.edu)
Center for the Study of Language and Information; 220 Panama Street; Stanford, CA 94305 USA

Martha W. Evens (evens@iit.edu)
Department of Computer Science (Illinois Institute of Technology); 10 W. 31st Street; Chicago, IL 60616 USA

Spoken Conversational Tutor (SCoT)

Human tutoring is known to be more effective than classroom instruction (Bloom 1984). Sophisticated intelligent tutoring systems (ITS) are only about half as effective as human tutors, however (Graesser et al. 2001). To approach the effectiveness of human tutors, ITS may need to not only use sophisticated tutoring strategies, but also use natural language dialogue (Graesser et al. 2001).

How should dialogue be modeled as part of an ITS? One approach has been to catalogue the dialogue moves found in human tutorial dialogues, and realize them as states in finite state automata that govern the dialog move option space for any input. This approach intertwines the mechanisms of dialogue (e.g., discourse markers, turn management) with the mechanisms of tutoring (e.g., hints, explanations). The joint activity theory (Clark 1996) separates conversational intelligence, i.e. how to use dialogue mechanisms in conversation, from the activity that a dialogue accomplishes. This separation provides for a clearer representation of how and why the structure of a task changes the structure of a dialogue.

Joint activities are activities in which participants have to coordinate their individual actions to succeed (e.g., a tutor identifying and addressing a student's misconceptions), using language and other signals. The functions of many of these signals are shared across domains (e.g., discourse markers), whereas dialogue structure varies as a consequence of the activity the dialogue serves. This suggests that linguistic knowledge should be kept separate from domain knowledge in an ITS.

We have adopted the joint activity approach in the development of a spoken conversational tutor (SCoT). SCoT is composed of three separate components: a knowledge base, a dialogue manager (the component that handles conversational intelligence), and a tutor. SCoT is developed within the Architecture for Conversational Intelligence (ACI; Lemon et al. 2001), a general purpose architecture which supports multi-modal, mixed-initiative dialogues with devices.

Conversational Intelligence

Conversational intelligence includes turn management, a structured representation of the dialogue, and appropriate use of discourse markers.

The ACI dialogue manager creates and updates an *Information State* (a dialogue context). Dialogue moves update information states. The dialogue manager includes multiple dynamically updated components; e.g., the dialogue move tree (a structured history of dialogue moves) and the activity tree (a hierarchical representation of joint activities). These structured histories provide a representation for past topics, should a student or tutor want to refer back to previous parts of the discussion.

Tutorial Intelligence

Tutorial intelligence includes strategies for determining if a student possesses a piece of knowledge and tactics for reacting to a student's needs during a session.

Strategies are methods for constructing a plan for post-practice reflective dialogue. SCoT uses a record of the student's performance to construct an initial tutorial plan.

Tactics (e.g., hinting) are triggered by student input (e.g., an answer to a question). Each tactic has a goal, preconditions, and a multi-step recipe. Preconditions include classifications of the student's answer and action in the problem-solving session. Recipes are composed of a sequence of actions. Actions can consist of a single primitive like providing feedback ("Yes, exactly") or an embedded tactic (giving the student a hint and then reasking a question).

The activity tree serves to represent the particular tutoring strategies and tactics chosen by the ITS, to allow the dialogue manager to interpret student actions and convey the tutor's queries, statements, etc. as linguistic utterances. Thus conversational and tutorial intelligence are distinct, but interact.

Acknowledgments

This work is supported by the Department of the Navy under research grant N000140010660. Further information is available at `http://www-csli.stanford.edu/semlab/muri`.