This project developed technology to enable multi-modal dialogue-based interfaces for interactive analysis and problem solving. We developed a system that acts as a planning assistant, interacting with the user in conversational English to help with situation assessment and logistics planning. This enabled users without prior training to interact successfully with a set of sophisticated AI reasoning tools. The novel system architecture we developed will serve as the starting point for a new generation of mixed-initiative spoken dialogue-based systems.
Final Report for
A Dialogue-based Architecture for a Tactical Picture Agent
ONR grant N00014-95-1-1008
ending Sept. 30, 2001

James F. Allen
Department of Computer Science
University of Rochester
Rochester, NY, USA
http://www.cs.rochester.edu/research/trains/

1. Executive Summary
This project developed technology to enable multi-modal dialogue-based interfaces for interactive analysis and problem solving. We developed a system that acts as a planning assistant, interacting with the user in conversational English to help with situation assessment and logistics planning. This approach enabled users without any prior training to interact successfully with a set of sophisticated AI reasoning tools including planners, schedulers, and knowledge based reasoning systems. The novel system architecture we developed will serve as the starting point for a new generation of mixed-initiative spoken dialogue-based systems.

2. Approach
One of the main premises in the project was that effective interactive situation assessment and collaborative plan development should have many of the characteristics of human dialogue, especially in the way topics are developed incrementally, then refined, clarified, and corrected iteratively until an acceptable solution is found. We developed a model where the human planner is always in the loop, and where both the human and the planning system could make contributions to the plans under development. The system recognizes the human planners intentions, determines how these affect the plan(s) under development, reasoned about the results of changes to the plans, and finally presented this information as a natural part of the dialogue. The state of the dialogue was captured by an abstract plan that serves as the interface between the humans perception of the goals and solutions so far, and the specialized domain reasoners (such as schedulers, routes planners, feasibility assessment, etc).

The research was driven by experimentation. We built robust end-to-end dialogue systems that interact with a user in order to accomplish a quantifiable task. By robust, we mean that everyday people could sit down at the machine and with less than a minute of instruction, engage in a dialogue to solve problems. By end-to-end, we mean that the system is complete: from speech input, through planning and reasoning and back to speech output. By a quantifiable task we mean that these were real problems for which there are right and wrong answers, or measurable degrees of effectiveness of solutions.

DISTRIBUTION STATEMENT A:
Approved for Public Release
Distribution Unlimited
3. The Architecture

One of the key results of this project was the new system architecture. Figure 1 shows the key new parts of the architecture. The earlier version of the system had a centralized Discourse Manager that controlled the interpretation of user input and the planning of system responses. In the new design these functions are split apart. The Interpretation Manager (IM) is responsible for interpreting the user input and computing any discourse obligations (which constrain allowable system responses). The Behavioral Agent (BA) is responsible for determining the system's overall behavior, using problem solving obligations that it computes from users problem solving actions (computed by the IM) plus using information from other information sources as well. The Generation Manager (GM) is responsible for determining the system's communication with the user, and it uses the discourse obligations from the IM as well as the actions to perform from the BA. This new architecture has one key new technological advantage and two key theoretical advantages:

1. It improves the portability of the system to new domains;
2. It allows the system to take more initiative in dialogues rather than mainly responding to the user input;
3. It allows the system to be more responsive and natural (e.g., the GM can provide acknowledgement even when the BA is deciding on a response).

In addition, we have made many advances in a range of topics relating to reference resolution and discourse context, generation, and robust parsing. These results are described in the papers listed below.
References


Stent, A.J., “Content planning and generation in continuous-speech spoken dialog systems,” *Proc., “May I Speak Freely?” Workshop* (to be held at the *German Annual Conf. on AI* (KI-99), Bonn, September 1999.

