Project Summary: Linguistic Knowledge Sources for Spoken Language Understanding

Principal Investigators
Lynette Hirschman and Deborah Dahl

Unisys Defense Systems
Center for Advanced Information Technology
Paoli, PA 19301

The objective of the Unisys Spoken Language Systems effort is to develop and demonstrate technology for the understanding of goal-directed spontaneous speech input. The Unisys spoken language architecture couples a speech recognition system (the MIT Summit system) with the Unisys discourse understanding system Pundit. Pundit is a broad-coverage language understanding system used in a variety of message understanding applications and extended to handle spoken language input. Its power comes from the integration of syntax, semantics and pragmatics (context), the ability to port rapidly to new task domains, and from an open, modular architecture. Pundit is unique in its ability to handle connected discourse; it includes a reference resolution module that tracks "discourse entities" and distinguishes references to previously mentioned entities from the introduction of new entities. The Pundit front-end supports turn-taking dialogue and permits the system to include both questions and answers in building an integrated discourse context, required for the handling of interactive communication. Pundit has been interfaced to speech recognition systems (both Summit and the ITT continuous speech recognizer) to perform applications on direction-finding assistance (Voyager), air travel planning (ATIS), and air traffic control (flight strip updating).

Progress

During the first year of our contract, we have ported the MIT Summit system and provided a loosely coupled interface with Pundit, via an N-best interface. Pundit processes the N top-scoring acoustic hypotheses until it finds an "understandable" hypothesis. We have been able to evaluate performance of this system on a demonstration task for direction finding assistance (based on the MIT Voyager system). We have also participated in the DARPA Common Evaluation Task in the Air Travel Information System domain. In support of the evaluation efforts, we have proposed an evaluation technique for automated dialogue evaluation and have developed techniques for "outside evaluators" to evaluate correctness of query/response pairs. Our major results to date include:

Demonstration of 70% correct test results on unseen data for a typed direction-assistant task, based on 1000-sentence training corpus, using the Pundit language understanding system plus a dialogue management front-end which manages the user/computer conversation.

Demonstration of a loosely coupled Spoken Language System (50-100x real-time); in initial experiments, application accuracy is approximately 50% for Spoken Language test data, with a false alarm rate of less than 5%.

Demonstration of a typed query front-end for the Air Travel Information System application, interfaced to an INGRES database. Metrics will be reported at the June DARPA Workshop.

Development of a proposal for evaluation of discourse, extending the notion of the database comparator for use in evaluation of sentences in context.

Plans

Our plans for the coming year focus on increasing the performance of our spoken language system, with particular focus on measuring and improving system usability, that is, the degree to which the system helps (or hinders) a user to accomplish a particular task. This will require extensions to the vocabulary (1000-2000 words), increase in perplexity (60-100), increased coverage (at least 70%), and improved error diagnosis and feedback. It will also require the development of metrics for user satisfaction and for task completion. This goal will also drive development of interactive dialogue capability, providing co-operative responses to users. We plan to develop a testbed, where we can interface multiple speech recognition front-ends to Pundit, via the N-best interface and the stack-decoder interface developed at Lincoln Labs. This will provide insight into some of the tradeoffs between recognizer improvements and improvements in language understanding.