

Integration of Speech and Natural Language Understanding for Spoken Language Systems (SLS)

Patti Price
SRI International

OBJECTIVES

SRI is developing a multi-modal interface to an air travel database that will permit cooperative planning via interactive human-machine problem solving. The project goals are real-time, large vocabulary (3000 words), high semantic accuracy (90%), speaker-independent understanding of spontaneously spoken natural language in a restricted domain. The grammar should be shown to be habitable and the system should be robust to individual differences in dialect and speaking style. The design should be easily portable to a variety of applications.

RECENT ACCOMPLISHMENTS

- Demonstrated the feasibility of a unique frame-level integration of speech and natural language which permits independent development and, computationally efficient, integrated application of constraints from phonetics through semantics,
- Developed a phonological rule set with high coverage and low overcoverage that when combined with algorithms developed under the Real Time Speech Recognition contract improved performance for both speaker-independent and speaker-dependent recognition,
- Incorporated techniques (developed under an NSF grant) that use statistical information about the speech signal and the noise to improve recognition accuracy in noisy environments,
- Shown the feasibility of integrating prosodic information (e.g., duration models developed under a DARPA-NSF contract) in the grammar-parser,
- Collected, transcribed, and analyzed over 8 hours of human-human air travel dialogues for use in vocabulary estimates and initial system design,
- Acquired the Official Airline Guide Database; implemented it in a relational non-proprietary form delivered to TI; created an initial SLS simulation; used it to collect 88 air travel planning dialogues which have been transcribed, analyzed and used to design: (a) data collection methods for evaluation of an SLS kernel, and (b) a more complex multi-modal interactive interface in the air travel planning domain.

PLANS

- Port the syntax from the resource management domain to the air travel planning domain; incorporate domain-specific semantics for the air travel planning domain,
- Begin integration of phonology, syntax, semantics, and prosody for accurate modeling of spontaneous speech; statistical modeling of linguistic relationships that can control perplexity and maintain habitability,
- Continue coordination with related contracts where appropriate,
- Collect and analyze more human-machine data under various conditions in order to guide the design and evaluation of the SRI SLS.