A Real-Time Spoken-Language System for Interactive Problem Solving

Patti Price and Robert C.Moore

SRI International
Menlo Park, CA 94025

PROJECT GOALS

The goal of this project, to develop a spoken-language interface to the Official Airline Guide database, is being developed along two overlapping research and development lines: one focussed on an SLS kernel for database query, and the other on the full interactive system.

RECENT RESULTS

- Evaluation of SRI's NL, SLS, and speech recognition technologies. SRI's February-92 weighted error rate for ATIS class A and D utterances combined was 31.0% for NL and 45.4% for SLS. SRI's February-92 SPREC word error rate for all ATIS utterances was 11.0%. For the speaker-independent, closed 5K vocabulary CSR task, SRI achieved a 16.6% word error rate in the verbalized punctuation test and a 17.1% word error rate in the non-verbalized-punctuation test, using the standard bigram language model.
- Encouraging initial results in a study of verbal repairs in the ATIS MADCOW corpus; the study included analysis and classification of the repairs in the corpus as well as a method using both pattern matching techniques and acoustic evidence to detect and correct the repairs.
- Implementation of RASTA filtering (high-pass filtering in the Log Spectral domain to remove time-invariant or slow moving linear channel effects) to improve channel robustness.
- Experiments showing differences between spontaneous and read speech and the need for including spontaneous speech in training, development, and test sets.
- Development of infrastructure necessary for dealing with very large vocabularies and language models such as those associated with the CSR corpus.
- Improvement in GEMINI natural-language processing system's linguistic coverage of a 2100-utterance ATIS traning corpus to 92% syntactic coverage, 77% semantic coverage.
- Development of algorithms for tracking evolving discourse structure. On the development set of the ATIS corpus, we find these algorithms correctly determine the discourse context with greater than 90% accuracy.
- Improvements to the template matcher: new templates, improved portability, more linguistic knowledge incorporated, reduced overgeneration, improved context handling, improved coverage of database retrieval component, all without significantly increasing response time, which is still less than a second.
- Analysis of data on human-machine problem solving using our SLS ATIS system: we have analyzed the effect on user satisfaction and system performance of system errors, user experience, and instructions to users. We have also explored trade-offs of time vs.speed and user satisfaction.
- Collection of data (speech, transcriptions, and log-files) using our SLS ATIS system. To date this includes 72 speakers, 2301 utterances, and 134 scenarios.
- Improvements in the SLS ATIS demo system, including better paraphrasing of system's understanding, easier to read displays, improved system error messages, simpler control of context mechanism.
- Development with MIT of a method for end-to-end evaluation that takes into account the whole interaction; implementation of the method is in progress.

PLANS FOR THE COMING YEAR

- Fully integrate GEMINI into SRI's ATIS system.
- Explore tighter integration of speech and NL processing for better overall SLS performance.
- Improve speech recognition computational and word-accuracy performance for the ATIS and CSR domains with better models of the channel, speaker, and spontaneous speech effects.