Session 8: Spoken Language Systems II

Charles T. Hemphill, Chair
Texas Instruments
P. O. Box 655474, MS 238
Dallas, TX 75265

This session consisted of four talks from six papers representing the gamut of technology in spoken language systems. Perhaps surprisingly, each of the papers present techniques that are compatible with one another and SLS system builders can benefit from all of these.

Mike Philips of MIT presented recent progress on the VOYAGER system. In particular, MIT has incorporated a "top-N" approach to increase the level of integration between speech and natural language. They use Viterbi search to compute the upper bound estimate for an A* search. Readers should contrast this with papers from BBN and AT&T in the first SLS session. MIT has also developed an interesting way of creating a high-coverage/low-perplexity word-pair grammar based on sentences generated from their natural language grammar.

Debbie Dahl of Unisys presented some observations on training and evaluation of a spoken language system. Using "top-N" Voyager data from MIT, they explored the relationship between training and coverage. They converged on 70% coverage after approximately 1000 sentences. This convergence depends on the domain and perhaps the data collection paradigm, as discussed in Unisys' paper in session five. Unisys has also tabulated empirical data for correct responses versus false alarms to help determine a beneficial value of N for the "top-N" algorithm.

Alex Rudnicky of CMU presented developments in spoken language interaction. While the work discussed above primarily addressed spoken language speed and accuracy, CMU has concentrated on the user interface and system integration issues. They have designed a spoken language system architecture that permits the rapid design of spoken language applications. With this architecture, they have studied computer-human interface design for five different applications. The modularity of the architecture facilitates research and development for each of the individual components.

Finally, Dave Stallard from BBN presented their recent developments in applying unification grammars to spoken language systems. It should be noted that most of the concepts presented can be found in the unification grammar literature, and BBN indicated this by saying that the work contains ideas "so old that they are new." In the first paper, BBN presents various situations that illustrate the advantage of using e-productions in a unification grammar. In the second paper, BBN presents various approaches for reducing the time and space requirements of a unification grammar system. These include: the compilation of rules into rule-groups that share mutually subsumable constituents, a limited form of feature disjunction in unification, and prediction constrained by features.