

## **CAREER: Modeling and optimizing user-centric mixed-initiative spoken dialog systems**

Shrikanth Narayanan, University of Southern California

### **Project Summary**

Conversational interfaces hold the promise of providing natural, easy and universal access to information. The research effort of this project targets two specific problems in conversational engineering under a unifying stochastic modeling framework: (1) dialog interaction modeling (2) user modeling in spoken dialog. These problems are related: user models are essential in designing optimal dialog strategies, while discourse state information is essential for user behavior modeling. Conversational participants -- humans or machines -- are modeled as stochastic dynamical systems, interacting with one another over a noisy communication channel. Noise in the channel may be due to errors in machine speech recognition, language understanding or other machine/user communication errors. Algorithms based on mathematical game theory and robust design are formulated for derivation of optimal mixed-initiative dialog control strategies. The approach to modeling user behavior is data-driven with an emphasis on behavior under error conditions and on the inclusion of automatic emotion tracking. Again, a unified statistical framework provides a way for integrating multiple sources of information (e.g., acoustic, lexical, nonverbal and discourse) based on information-theoretic principles. The research development will be based on annotated spoken language corpora from real human-machine dialogs from a variety of real application domains including air travel and call center applications and an immersive training system for military/security applications being developed at the University of Southern California (USC).

The educational component creates a new curriculum in spoken language processing: a two-semester graduate sequence, developed for simultaneous offering both on campus and on USC "Distance Education Network" will target basic speech processing through the latest advances in speech recognition, speech synthesis and spoken dialog. The courses will be relevant to a wide range of programs in addition to core electrical engineering, including computational linguistics, intelligent systems in computer science, integrated media and creative technologies, phonetics and psychology. In addition, a unique interdisciplinary undergraduate course with a theme on human versus machine processing of spoken language will be developed and implemented. The research and education components are tightly integrated, with courses featuring the latest research results through demonstrations and hands-on projects, with an emphasis on "learning by doing". Likewise, research serves as an effective vehicle for education in terms of enabling training both undergraduate and doctoral students.

This project in spoken language engineering is unique in terms of its scope and societal relevance, spanning both fundamental and applied aspects of spoken language research, including an industry perspective, and creating synergies with other activities within the university in multimedia, natural language processing and agent technology. It will serve as a vehicle to promote cross-disciplinary graduate and undergraduate research and education.