



Professor Chin-Hui Lee

Professor
School of Electrical and Computer Engineering
Georgia Institute of Technology



An Information Extraction Approach to Next-Generation Speech Processing

Abstract:

The field of automatic speech recognition (ASR) has enjoyed more than 30 years of technology advancement due to the extensive utilization of the hidden Markov model (HMM) framework and a concentrated effort by the community to make available a vast amount of language resources. However the ASR problem is still far from being solved because not all information available in the speech knowledge hierarchy can be directly and effectively integrated into state-of-the-art systems to improve ASR performance and enhance system robustness. It is believed that some of the current knowledge insufficiency issues can be partially addressed by processing techniques that can take advantage of the full set of acoustic and language information in speech. On the other hand in human speech recognition (HSR) and spectrogram reading we often determine the linguistic identity of a sound based on detected cues and evidences that exist at various levels of the speech knowledge hierarchy, ranging from acoustic phonetics to syntax and semantics. This calls for a bottom-up knowledge integration framework that links speech processing with information extraction, by spotting speech cues with a bank of attribute detectors, weighing and combining acoustic evidences to form cognitive hypotheses, and verifying these theories until a consistent recognition decision can be reached. The recently proposed ASAT (automatic speech attribute transcription) framework is an attempt to mimic some HSR capabilities with asynchronous speech event detection followed by bottom-up speech knowledge integration and verification. In the last few years it has demonstrated potentials and offered insights in detection-based speech processing and information extraction.

This presentation is intended to illustrate new possibilities of speech research via linking analysis and processing of raw speech signals with extracting multiple layers of useful speech information. By organizing these probabilistic evidences from the speech knowledge hierarchy, and integrating them into the already-powerful, top-down HMM framework we can facilitate a knowledge-rich, bottom-up and data-driven framework that will lower the entry barriers to ASR research and further enhance the capabilities and reduce some of the limitations in the state-of-the-art ASR systems. Everyone in and out of the current ASR community will be able to contribute to this worthwhile effort to building a collaborative ASR community of the 21st Century.

About the speaker:

Chin-Hui Lee is a professor at School of Electrical and Computer Engineering, Georgia Institute of Technology. Professor Lee has participated actively in professional societies. He is a member of the IEEE Signal Processing Society (SPS), Communication Society, and the International Speech Communication Association (ISCA). Professor Lee is a Fellow of the IEEE, and has published 400 papers and 30 patents. In 2012 he was invited by ICASSP to give a plenary talk on the future of automatic speech recognition. He was selected as an ISCA Fellow in 2012, and awarded the 2012 ISCA Medal in scientific achievement for "pioneering and seminal contributions to the principles and practice of automatic speech and speaker recognition, including fundamental innovations in adaptive learning, discriminative training and utterance verification". For more information please visit <http://csip.ece.gatech.edu/?q=faculty/chin-hui-lee>

Date: 5 August 2013 (Monday)
Time: 11:00am – 12:30pm
Venue: Room 222, Ho Sin Hang Engineering Building

**** ALL ARE WELCOME ****

