

Call for Papers
IEEE Transactions on Audio, Speech, and Language Processing
IEEE Signal Processing Society

Special Issue on Deep Learning for Speech and Language Processing

Over the past 25 years or so, speech recognition technology has been dominated largely by hidden Markov models (HMMs). Significant technological success has been achieved using complex and carefully engineered variants of HMMs. Next generation technologies require solutions to technical challenges presented by diversified deployment environments. These challenges arise from the many types of variability present in the speech signal itself. Overcoming these challenges is likely to require "deep" architectures with efficient and effective learning algorithms.

There are three main characteristics in the deep learning paradigm: 1) layered architecture; 2) generative modeling at the lower layer(s); and 3) unsupervised learning at the lower layer(s) in general. For speech and language processing and related sequential pattern recognition applications, some attempts have been made in the past to develop layered computational architectures that are "deeper" than conventional HMMs, such as hierarchical HMMs, hierarchical point-process models, hidden dynamic models, layered multilayer perceptron, tandem-architecture neural-net feature extraction, multi-level detection-based architectures, deep belief networks, hierarchical conditional random field, and deep-structured conditional random field. While positive recognition results have been reported, there has been a conspicuous lack of systematic learning techniques and theoretical guidance to facilitate the development of these deep architectures. Recent communication between machine learning researchers and speech and language processing researchers revealed a wealth of research results pertaining to insightful applications of deep learning to some classical speech recognition and language processing problems. These results can potentially further advance the state of the arts in speech and language processing.

In light of the sufficient research activities in this exciting space already taken place and their importance, we invite papers describing various aspects of deep learning and related techniques/architectures as well as their successful applications to speech and language processing. Submissions must not have been previously published, with the exception that substantial extensions of conference or workshop papers will be considered.

The submissions must have specific connection to audio, speech, and/or language processing. The topics of particular interest will include, but are not limited to:

- * Generative models and discriminative statistical or neural models with deep structure
- * Supervised, semi-supervised, and unsupervised learning with deep structure
- * Representing sequential patterns in statistical or neural models
- * Robustness issues in deep learning
- * Scalability issues in deep learning
- * Optimization techniques in deep learning
- * Deep learning of relationships between the linguistic hierarchy and data-driven speech units
- * Deep learning models and techniques in applications such as (but not limited to) isolated or continuous speech recognition, phonetic recognition, music signal processing, language modeling, and language identification.

The authors are required to follow the Author's Guide for manuscript submission to the IEEE Transactions on Audio, Speech, and Language Processing at <http://www.signalprocessingsociety.org/publications/periodicals/taslp/taslp-author-information>

- * Submission deadline: September 22, 2010 (**EXTENDED**)
- * Notification of Acceptance: March 15, 2011
- * Final manuscripts due: May 15, 2011
- * Date of publication: August 2011

For further information, please contact the guest editors:

Dong Yu (dongyu@microsoft.com)
Geoffrey Hinton (hinton@cs.toronto.edu)
Nelson Morgan (morgan@ICSI.Berkeley.edu)
Jen-Tzung Chien (jtchien@mail.ncku.edu.tw)
Shigeki Sagayama (sagayama@hil.t.u-tokyo.ac.jp)