

# UW CENTER FOR PATTERN ANALYSIS AND MACHINE INTELLIGENCE

## INVITED SEMINAR SERIES

### Suspending Deep Belief

**Speaker:** Gerald Penn, University of Toronto

**Date:** May 8, 2013

**Time:** 2:00 pm – 3:00 pm

**Place:** DC-1304 (Refreshments will be served)

#### **Abstract :**

Deep neural networks (DNNs), often accompanied by generative pre-training with deep belief networks (DBNs), have started to supplant Gaussian mixture models (GMMs) as the default acoustic models for automatic speech recognition (ASR). When the output nodes of DNNs are expanded from a small number of phonemes into a large number of tied-states of triphone HMMs, it has been reported that the resulting so-called context-dependent DNN/HMM hybrid model achieves an unprecedented performance gain in many challenging ASR tasks, including the well-known Switchboard task. Why, specifically, is this happening? What is it about the DNN approach that grants them this ability? In this talk, I'll describe some experiments that reveal some clues to help us answer this question. The results of these experiments suggest that DNNs do not necessarily yield better modelling capability than conventional GMMs for standard speech features. DNNs, however, are very powerful in terms of leveraging highly correlated features. The unprecedented gain of the context-dependent DNN/HMM model can be almost entirely attributed to the DNN's input feature vectors, having been concatenated from several consecutive speech frames within a relatively long context window. Then we'll turn our attention to DBN pre-training. Again, where is the benefit coming from? Our recent attempts at answering this have revealed a simple but novel use of convolutional neural networks that can beat a DBN pre-trained network with a similar number of trainable weights.

#### **Biography:**

Gerald Penn is the Associate Chair, Research and Industrial Relations in the Department of Computer Science at the University of Toronto, and a Fellow of Computer Science at the University of Trinity College. He received his Ph.D. in 2000 from Carnegie Mellon University. From 1998 to 2001, he was a Member of Technical Staff in the Multimedia Communications Research Laboratory at Bell Labs in the United States. His research interests include speech and natural language processing, mathematical linguistics and programming language theory.

**Co-Sponsored by:** KW IEEE CIS, KW IEEE CSS, KW IEEE SMC, KW IEEE SP

**WATERLOO  
ENGINEERING**

**CPAMI** 