A preliminary version of this work has been published here:

Harati Nejad Torbati, A. H., Picone, J., & Sobel, M. (2014). A Left-to-Right HDP-HMM with HDPM Emissions. In *Proceedings of the Conference on Information Sciences and Systems* (pp. 1–6). Princeton, New Jersey, USA.

This paper presented the initial idea, but did not include many important mathematical details or a comprehensive set of experiments.

Major Theoretical Differences:

# The model introduced in the conference paper was restricted to only a special case of the more general model introduced in this paper. Specifically, in this paper we have introduced a general model named DHDPHMM and its inference algorithm while in the conference paper we only introduced a special case of the model with no details about the inference algorithm.

# In this paper, we provide important theoretical derivations and implementation details regarding DHDPHMM. We discuss its differences relative to other models (e.g. HDPHMM). These details are not in the conference paper.

# In this paper, we introduce a general framework for non-Ergodic structures while in the conference paper we only derived a left to right structure.

# In this paper, we have included an inference algorithm for a Bayesian approach of adding non-emitting states while in conference paper these details were missing.

Major Experimental Differences:

# In this paper we have a very extensive experimental section while in the conference paper the experimental section was very brief. The only commonality is section of simulated data.

# Phoneme classification experiments are much more comprehensive in this paper. In the conference paper they were limited to left to right HDPHMMs and HMMs. In this paper we have added results for DHDPHMM (both ergodic and several non-ergodic structures) and also for many other state of the art models. These new models provide state of the art results, while in the conference paper the results were not as good since they used an older, less sophisticated model.

# We have more experiments that demonstrate how learning complexity and scalability are handled by DHDPHMM versus HDPHMM and HMM.

# The phoneme recognition experiments in this paper are completely new.