# Variational Inference for Hierarchical Dirichlet Process Based Nonparametric Models

**Will Stephenson**

**Advisor: Erik Sudderth , Reader: Ben Raphael**

## Abstract

We examine two popular statistical models, the hidden Markov model and mixed membership stochastic blockmodel. Using the hierarchical Dirichlet process, we define nonparametric variants of these models. We develop a memoized online variational inference algorithm that uses a new objective function to properly penalize the addition of unneeded states in either model. Finally, we demonstrate that our models outperform competing methods in a wide array of tasks, including speaker diarization, academic co-authorship network analysis, and motion capture comprehension.

## 1 Introduction

Many real-world processes follow complex or ill-understood procedures and rules; networks of human relationships develop from a series of often byzantine social rules, and human exercise generates motion capture data that depends on the specifics of human physiology. We seek mathematical models for such processes that are able to explain observed data well while remaining concise. Many popular statistical models do so by assuming the existence of a number of unobserved, or *hidden*, states. The diverse nature of data is then explained by the diversity among the hidden states. The hidden state of human exercise might be the specific exercise being performed at any point in time; an entire sequence of motion capture data can then be understood as a series of particular exercises along with the specific dynamics of each exercise.

We focus on two particular models: the hidden Markov model (HMM) [16] and mixed membership stochastic blockmodel (MMSB) [1] which model time series and relational data, respectively. Both seek to understand the structure of data by defining some number of hidden states or communities. Given these states, they specify a data generation process. For example, audio data of human conversation may be understood as a series of speakers along with their individual speech patterns; a social network may be understood in terms of cliques of friends with high probability of intra-clique interaction and low probability of inter-clique interactions.

Crucially, training either of these models requires us to specify the number of states or communities. In many cases this number is difficult to estimate or explicitly unknown; for example, the speaker diarization task [19] requires us to segment audio data of meetings with an unknown number of speakers. To address this issue, we turn to *nonparametric* models, which assume an unbounded number of states, some finite subset of which are represented in the data. Specifically, we use the *hierarchical Dirichlet process* [18] to define the HDP-HMM [9] and HDP-MMSB [14].

Given a dataset and generic description of a model, we wish to infer the specific model parameters that best match the data. To perform this inference, we use *memoized variational inference* [12]. We derive a new variational objective function for HDP-based models and exploit the structure of the HDP to show it generalizes to both our models. These choices follow the work of [11] on HDP topic models and address issues with previous inference methods for the HDP-HMM and HDP-MMSB.