HMM Analysis of Musical Structure:
Identification of Latent Variables Through Topology-Sensitive Model Selection

Panayotis Mavromatis

Music and Audio Research Lab (MARL),
New York University,
35 West 4th St., Suite 777,
New York, NY 10012, USA
panos.mavromatis@nyu.edu
http://theory.smusic.nyu.edu/pm/

Abstract. Hidden Markov Models (HMMs) have been successfully employed in the exploration and modeling of musical structure, with applications in Music Information Retrieval. This paper focuses on an aspect of HMM training that remains relatively unexplored in musical applications, namely the determination of HMM topology. We demonstrate that this complex problem can be effectively addressed through search over model topology space, conducted by HMM state merging and/or splitting. Once successfully identified, the HMM topology that is optimal with respect to a given data set can help identify hidden (latent) variables that are important in shaping the data set’s visible structure. These variables are identified by suitable interpretation of the HMM states for the selected topology. As an illustration, we present two case studies that successfully tackle two classic problems in music computation, namely (i) algorithmic statistical segmentation and (ii) meter induction from a sequence of durational patterns.

1 Introduction

Hidden Markov Models have been successfully employed in the exploration and modeling of musical structure [1,2], with applications in Music Information Retrieval [3].

Simply put, a Hidden Markov Model is a probabilistic version of a Finite State Machine (FSM), or formal specification of a finite state grammar. A FSM is formally defined by states and transitions, graphically represented by circles and arrows respectively. A FSM generates a symbolic sequence by traversing a path of states connected by transitions, following the direction of the arrows. The generated sequence is the string of output symbols encountered in the path. A FSM is a simple and flexible way to specify finite-memory constraints on the symbolic values of variables that characterize musical structure (e.g., pitch, duration, etc.) and as such offers useful formal characterizations of the structure of musical sequences.
A Hidden Markov Model (HMM) is a FSM with probabilities attached to its transitions and output symbols. The generation of a sequence through a specific HMM path has probability equal to the product of all transition and output probabilities encountered in traversing the generating path.

What gives the HMM technique its strength and flexibility is the fact that selecting the best HMM for a given data set can be generally accomplished through efficient algorithms. For instance, given a data set of symbolic sequences whose structure we wish to explore, it is customary to assume a HMM of fixed topology (i.e., number of states, and how they are connected by transitions) and identify the model parameters (i.e., transition and output probabilities) that best fit the data set, in the sense of Maximum Likelihood Estimation, using the so-called Baum-Welch algorithm.

This paper focuses on an aspect of HMM training that remains relatively unexplored in musical applications, namely the determination of HMM topology. Our aim is to algorithmically construct models whose topologies consist of states interpretable as values of latent ("hidden") variables that may play important role in the determination of musical structure. In a given application, one may wish to focus on a particular ("visible") musical variable, aiming to model syntactical constraints on its successive values (e.g., stylistically acceptable patterns of note durations). The states of a HMM obtained through topology-sensitive search should indicate which additional variables must be taken into consideration (e.g., metric position) in order to understand the syntax of the original "visible" variable that one set out to model. This can be accomplished by showing a close correspondence between HMM states and particular values of the candidate "hidden" variables.

For an HMM topology to be interpretable in the manner suggested in the preceding paragraph, special effort must be put in the topology selection algorithm. If one simply relies on Baum-Welch optimization of the HMM parameters, one will in most cases obtain HMMs whose states are not readily interpretable, however well these models may fit the data. Previous studies that attempted to address this complex problem have generally employed some form of search over model topology space, which was conducted by HMM state merging or splitting. In this paper, we use the same basic search procedure, except that we allow state merging and splitting to be combined in the same search. In addition, we evaluate each candidate model using a Bayesian approach, in which a HMM’s prior probability is determined through the Minimum Description Length principle. This prior is optimal in that it leads to models that are neither too large nor too small, and has been found to provide a reliable termination criterion for the state merging/splitting search.

We will illustrate our method with the help of two case studies that successfully tackle two classic problems in music computation, namely (i) algorithmic statistical segmentation and (ii) meter induction from a sequence of durational patterns.