
This document is also available in PDF¹ format.

Purpose: Discuss types of HMMs and other related issues.

Material: Paper by Rabiner and Juang; book by Deller et al; paper by SE Levinson, “Structural methods in automatic speech recognition”, *Proceedings of the IEEE*, vol 73 pp 1625–1650, Nov 1985.

General: The key concepts of HMMs should now be in place. In this lecture we are going to look at some of the many variations and uses of HMMs. Primarily HMMs are distinguished by three factors: the type of pdf used, the state transition topology and the order of the model. In this lecture we will be primarily discussing densities, with some mention of topology issues. In the next lecture we will discuss the rôle of Markov order in more detail.

Topics:

- **Variants on the state PDFs:** In the article *discrete* and *continuous* HMMs are mentioned. Quite often the descriptive qualifier for an HMM really describes the state PDFs. Here are some examples:
 - **Discrete HMMs:** As a first step a preprocessing stage classifies the feature vectors according to a VQ codebook or some other description. The resultant one-dimensional features are then represented with a discrete PDF. This has the advantages of being fairly fast and also very flexible (i.e. close to a non-parametric PDF). The price paid for this lies in the quantisation error being made when substituting a continuous vector for a single discrete feature.
 - **Continuous HMMs** are normally more computationally intensive, but can yield more accurate results. Care must however be taken that the shape of the PDFs are appropriate to the task at hand. The state PDFs could also be mixtures of PDFs, normally a mixture of Gaussian PDFs. This allows for very general PDF shapes.
 - **Semi-continuous HMMs:** These really are continuous HMMs using mixtures that shares a common (global) set of underlying PDFs. The mixture weights remind of discrete HMMs while the underlying continuous PDFs remind of continuous HMMs – hence the name.
 - **ANN HMM hybrids:** It has been shown that Artificial Neural Networks (ANNs) can be used to approximate the state PDFs under certain circumstances. This is a mechanism for using continuous HMMs with very flexible densities that are also discriminatively trained. It is, however, much slower and more prone to bad local optima. More on this when we cover ANNs later in the course.
- **Topology:** The structure in which the states are joined together has a large effect on the modelling capacity of the HMM. Two popular versions are the *left-to-right* structures used for phoneme and word models and *ergodic* structures which are more fully connected, allowing recurring visits to previous states and groups of states. There are also special duration modelling structures which aim to improve the somewhat limited duration modelling capacity of the HMM. Interestingly, these structures build a link between the HMM and the so-called semi-HMM (another beast altogether – when you enter a state you remain there for a time dictated by a PDF explicitly modelling duration, after which you can make a transition).
- **HMM order:** In a later lecture we will see that higher-order HMMs really are a powerful way to describe (first-order) HMM topologies.
- **Combining HMMs:** Smaller HMMs can be joined together to form bigger networks which are once again HMMs. This is the mechanism utilised in word-spotters and phonetically-based speech recognition systems.
- **Sharing (tying) parameters:** States, as well as subparts of the HMM, can be *tied together*, thereby sharing PDFs and possibly also link probabilities. This is very useful in building bigger nets from smaller ones.

¹http://www.dsp.sun.ac.za/pr813/lectures/8_hmm_b/8_hmm_b.pdf