Hidden Process Models: Decoding Overlapping Cognitive States with Unknown Timing

Rebecca A. Hutchinson
Department of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
rah@cs.cmu.edu

Tom M. Mitchell
Machine Learning Department
Carnegie Mellon University
Pittsburgh, PA 15213
tom.mitchell@cs.cmu.edu

Abstract

We use Hidden Process Models (HPMs) to evaluate different models of a functional Magnetic Resonance Imaging (fMRI) study in which subjects decide whether stimuli match. We demonstrate the ability of HPMs to simultaneously estimate the hemodynamic response functions *and* the onset times of a set of cognitive processes underlying an fMRI time series, and to compare different models in a principled way.

1 Introduction

Hidden Process Models (HPMs) are a class of probabilistic multivariate time series models developed for the analysis of functional Magnetic Resonance Images (fMRI) [1]. An HPM consists of a set of processes, a set of constraints derived from the experiment design, and a noise model. Using a linearity assumption, HPMs put a Gaussian probability distribution on the observed time series, and the HPM learning algorithm iterates between estimating the timing of unknown processes and the parameters of the processes and noise model.

Consider an experiment in which subjects view pictures and read sentences (in either order) involving symbols and above/below relationships [2]. In each of 40 trials, the subject must push a button to indicate whether the sentence and picture match. To model the fMRI data from this experiment, we could use an HPM with two processes: ViewPicture and ReadSentence. The HPM could model each trial using two process instances, which inherit hemodynamic response function (HRF) parameters from the ViewPicture and ReadSentence processes. Furthermore, the HPM could use the constraint that the process instances in each trial may not match (subjects never saw two pictures or two sentences). Currently, the HPM noise model has an independent parameter for each voxel.

While this 2-process HPM seems reasonable, we may want to study the decision making in this experiment more closely by adding processes to the HPM. While conventional techniques require us to assume either the timing or the HRF of the decision process in order to estimate the other, HPMs can estimate both simultaneously. This allows us to compare different theories of the cognitive processes involved in this experiment in a principled way.

HPM Processes	Log-likelihood
P/S	$-1.0784 * 10^6$
P/S/D	$-1.0759 * 10^6$
P/S+/S-/D	$-1.0742 * 10^6$
P/S/D+/D-	$-1.0742 * 10^6$
P/S/Dy/Dn	$-1.0737 * 10^6$
P/S/D/D?	$-1.0734 * 10^6$

Table 1: Comparison of 6 HPMs. Key: P = ViewPicture, S = ReadSentence, D = Decide, S+ = ReadAffirmativeSentence, S- = ReadNegatedSentence, D+ = DecideOnAffirmativeSentence, D- = DecideOnNegatedSentence, Dy = DecideYes, Dn = DecideNo, D? = DecideNoAnswer.

2 Results and Discussion

In this experiment, we compared the performance of six different HPMs on a single subject. We performed 5-fold cross-validation, using the top 50 most active voxels from each of 22 anatomically-defined Regions of Interest (ROIs). The processes of the HPMs and their associated average log-likelihoods are reported in Table 1. Our results suggest that splitting the Decide process based on whether or not the subject responded fits the data better than splitting on affirmative vs. negated sentences or responding 'Yes' vs. 'No.'

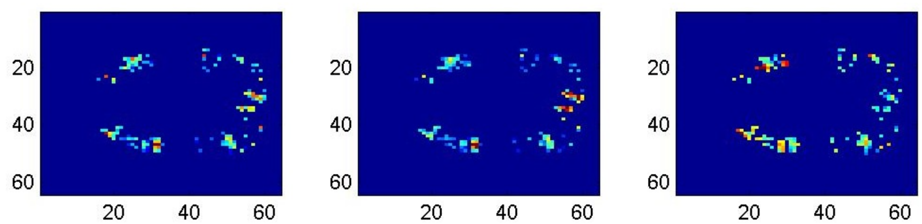

Figure 1: Mean process responses in each voxel for P/S/D HPM, slice 5. From left, each plot is the mean of the learned HRF for the ViewPicture, ReadSentence, and Decide processes in the P/S/D HPM for the voxels in a single slice of the brain.

Each HPM that involved decision processes learned a distribution over onsets from 0 to 4 seconds following the second stimulus presentation. The distributions were all different, but in each case the most probable onset was the latest one. Each of these HPMs also learned a HRF for the decision process(es). For example, the HRF for the processes learned by the P/S/D HPM are summarized in Figure 1.

We can also use HPMs for classification tasks related to the decision processes. For example, the P/S/Dy/Dn HPM identified whether the subject said 'Yes' or 'No' with 88.7% accuracy. (This subject said 'Yes' 11 times, 'No' 15 times, and did not answer 14 times.) In comparison, the leave-one-out cross-validated accuracy for a Gaussian Naive Bayes classifier on this task using the same voxels and the time series of the entire trial is 53.9%.

This experiment is an example of a class of experiments in which we may wish to study cognitive processing with unknown onsets. HPMs allow us to ask new kinds of questions about these experiments that we could not address with, e.g., the classifiers studied in [3]. At the workshop, we hope to discuss these contributions to fMRI data analysis and to get feedback on useful improvements to HPMs.

References

[1] Rebecca A. Hutchinson, Tom M. Mitchell, and Indrayana Rustandi. Hidden process models. In *Proceedings of the International Conference on Machine Learning*, 2006.

[2] T.A. Keller, M.A. Just, and V.A. Stenger. Reading span and the time-course of cortical activation in sentence-picture verification. In *Annual Convention of the Psychonomic Society*, 2001.

[3] Tom M. Mitchell et al. Learning to decode cognitive states from brain images. *Machine Learning*, 57:145–175, 2004.