

Neuroinformatics Journal Club

Dynamic Brain Networks (HMM)

Horacio Gómez-Acevedo
Department of Biomedical Informatics
University of Arkansas for Medical Sciences

July 8, 2022



Today's paper

NeuroImage 180 (2018) 646–656



ELSEVIER

Contents lists available at [ScienceDirect](#)

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Discovering dynamic brain networks from big data in rest and task

Diego Vidaurre^{a,*}, Romesh Abeysuriya^a, Robert Becker^a, Andrew J. Quinn^a, Fidel Alfaro-Almagro^b, Stephen M. Smith^a, Mark W. Woolrich^{a,b}

^a Oxford Centre for Human Brain Activity (OHBA), University of Oxford, UK

^b Oxford University Centre for Functional MRI of the Brain (FMRIB), University of Oxford, UK

Journal Club Format

This is a new format for Journal Club

- ▶ We will explore more carefully the methodology.
- ▶ Most of the papers are going to require "heavy" tools from computer sciences, statistics and mathematics.
- ▶ It will be a transition time to understand each other (be patient)
- ▶ The dynamical nature of brain activity will make it very challenging to hold on to a single model.
- ▶ Complexity vs. Explainability is an implicit trade-off.
- ▶ The ultimate goal is to use and expand "acceptable" methodologies into new and more "appropriate" ones. Yes, this means **coding** (R, Python, MatLab).

Hidden Markov Models

The implementation of Hidden Markov models dates back in the late 60's by Leonard Baum and collaborators. It is still an active area of research.

What is a Markov chain?

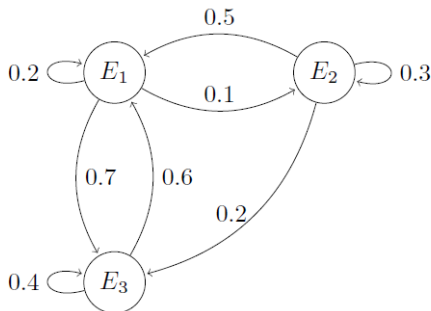
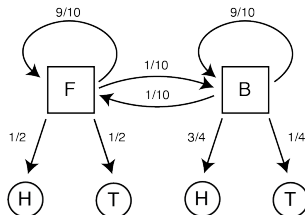


Figure 1: Markov chain H_2

Hidden Markov Chains

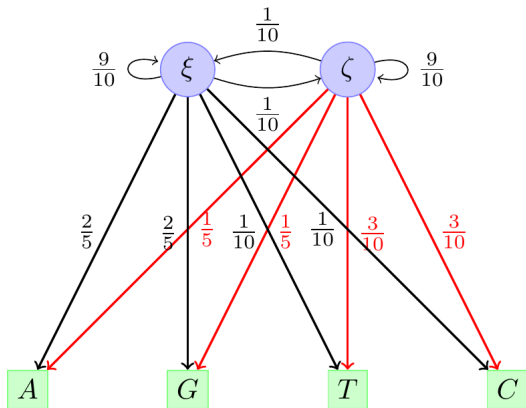
The fair casino problem

- ▶ You don't know if the dealer has a fair or unfair coin (Hidden states)
- ▶ You only observe the output (Emissions)

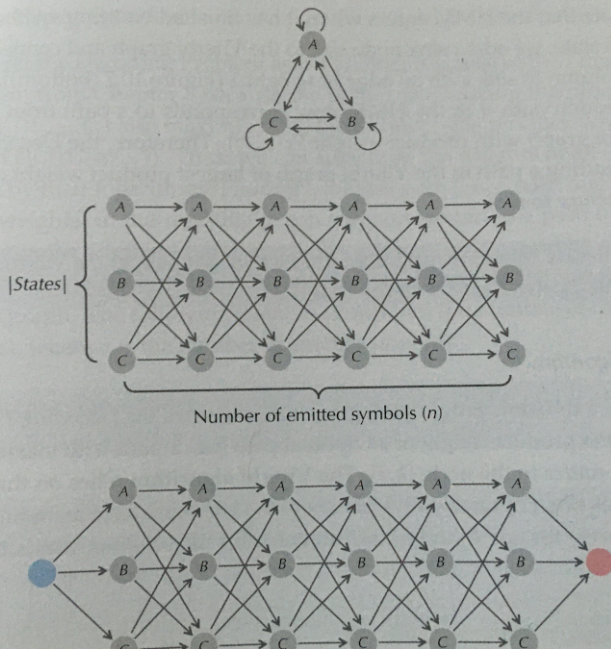


Hidden Markov Chains (Genomics)

Another (unrealistic) example is to build a HMM to describe a (genetic) sequence



Viterbi's representation HMMs



Building a HMM

From Baker et al. 2014

Here, we present a study that identifies transient networks of brain activity, with no prior assumptions on the brain areas or time scales involved. This uses a distinct methodology based on a hidden Markov model (HMM), which infers a number of discrete brain states that recur at different points in time. Each inferred state corresponds to a unique pattern of whole-brain spontaneous activity, which is modeled by a multivariate normal distribution and a state time course indicating the points in time at which that state is active. These two outputs are shown schematically in Figure 1, and allow us to describe both the spatial and temporal characteristics of each inferred state.

Building a HMM

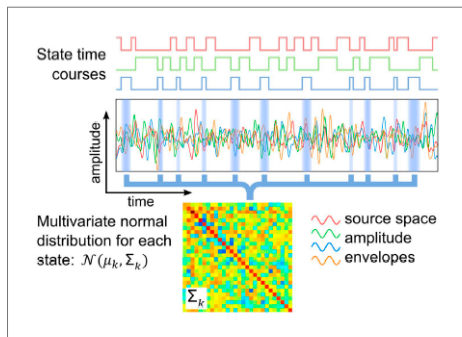


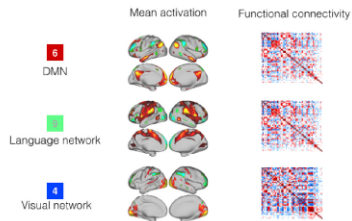
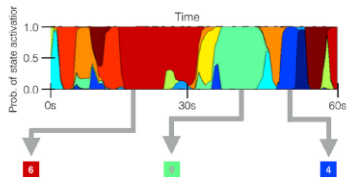
Figure 1. Schematic of the HMM outputs. An HMM with K states is inferred from band-limited amplitude envelopes of source reconstructed MEG data. Each state is characterized by a multivariate normal distribution (defined by means μ_k and covariance matrix Σ_k) and a state time course, which is a binary sequence that indicates the points in time at which the state is active.

DOI: [10.7554/eLife.01867.003](https://doi.org/10.7554/eLife.01867.003)

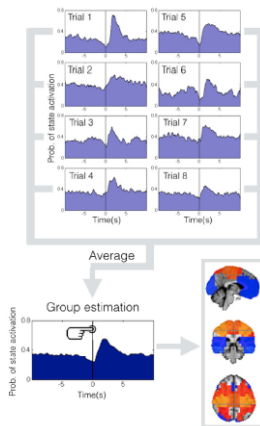
In the paper

The HMM is a family of models that can describe time series of data using a discrete number of states, all having the same probabilistic distributions but each having different distribution parameters. Thus, the states correspond to unique patterns of brain activity that recur in different parts of the time series. For each time point t , a state variable dictates the probability of each state being active at that moment

General HMM at rest



General HMM in task



General HMM cont

an HMM generally comprises the description of the states, the state time courses (which determines the probability of each state to be active at each time point in the time series) and the transition probabilities between the states (i.e. the probability to transition from each state to each other state). Because here we run the HMM on all concatenated subjects' datasets, the states and the transition probabilities are defined at the group level; the state time courses are however particular to each subject - that is, states can come active at different moments for each subject. Since the probability distribution of each part of the model depends on all others, there is no closed-form solution available

Autoregressive Model

If you consider time series data (meaning measuring the same variable over time), say X_t . An **autoregressive model** defines the current value based on previous values

$$X_t = \beta_0 + \beta_1 X_{t-1}$$

The **order** of the AR model refers to the number of past variables that you will consider. For instance, an AR(2) model would be

$$X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2}$$

Training a HMM

The normal way to train the parameters from a HMM is based on the Baum-Welch algorithm (backward-forward algorithm).

This algorithm requires the whole series of data to determine the parameters.

To avoid this issue the authors suggest a method called Stochastic Variational Inference.

Stochastic Variational Inference

The methodology is to split the HMM into two parts: local and global variables

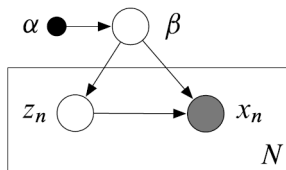
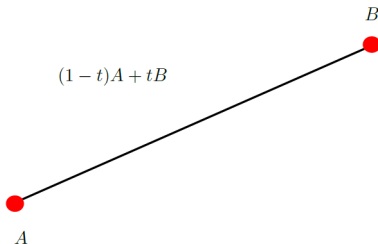


Figure 2: A graphical model with observations $x_{1:N}$, local hidden variables $z_{1:N}$ and global hidden variables β . The distribution of each observation x_n only depends on its corresponding local variable z_n and the global variables β . (Though not pictured, each hidden variable z_n , observation x_n , and global variable β may be a collection of multiple random variables.)

What about those equations?

$$\varphi_{new} = (1 - \rho)\varphi_{old} + \rho\varphi_{interum} \quad (1)$$



$$\rho(c) = (c + \alpha)^\beta \quad (2)$$

Final stochastic step

It runs a random sampling of subjects with a twist:
Promote those subjects that have been historically sampled fewer times.

1. Initialise the observation models.
2. Repeat:
 - 2.1 Choose M subjects at random using probabilities, w_i , computed as in (Eq. (3)).
 - 2.2 Compute the state time courses for the M subjects.
 - 2.3 Compute the interim state probability distribution using the M subjects as though we had N subjects.
 - 2.4 Perform an approximate update of the state probability distributions using (Eq. (1)).
 - 2.5 Perform an exact update of the state transition probabilities.
 - 2.6 Update ρ using (Eq. (2)).
 - 2.7 Compute the free energy.

Until the free energy converges.