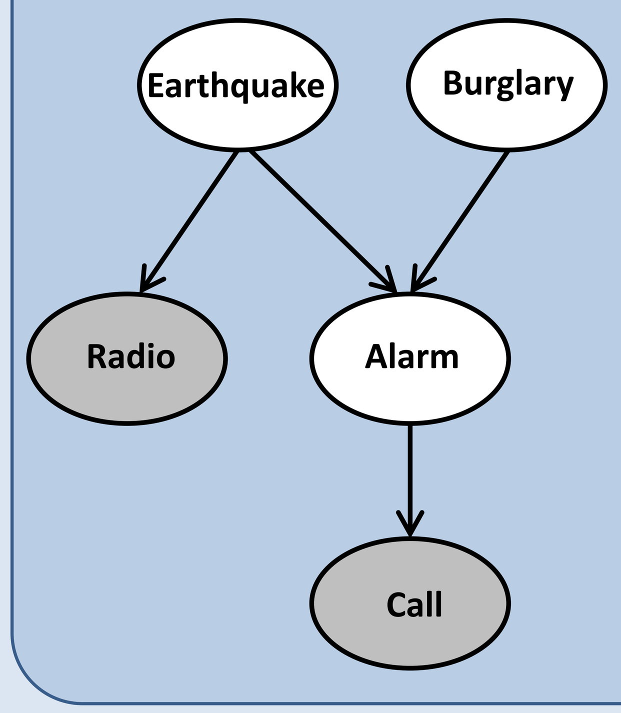
To Infinity and Beyond with Hidden Markov Models

Computational & Biological Learning Lab
University of Cambridge

Jurgen Van Gael† & Zoubin Ghahramani

1. Reasoning with Uncertainty



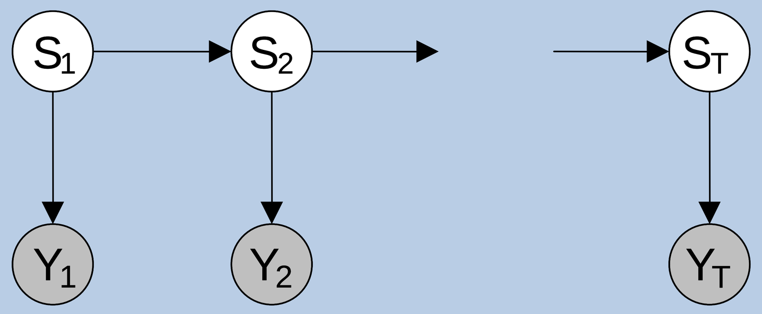
At any point in time, we know:
- there is a small probability of a burglary at our house
- there is a small probability of an earthquake
- if there is a burglary with high probability the alarm goes off
- if there is an earthquake, the alarm might go off
- if there is an earthquake, the radio will report about it
- if the alarm goes off, a security firm will call us

How likely is it that there is a burglary if the security firm calls us?
How likely is it that there is a burglary if the security firm calls us and the radio reports an earthquake?

How do we quantify this reasoning?
using **Graphical Models**

2. Hidden Markov Models

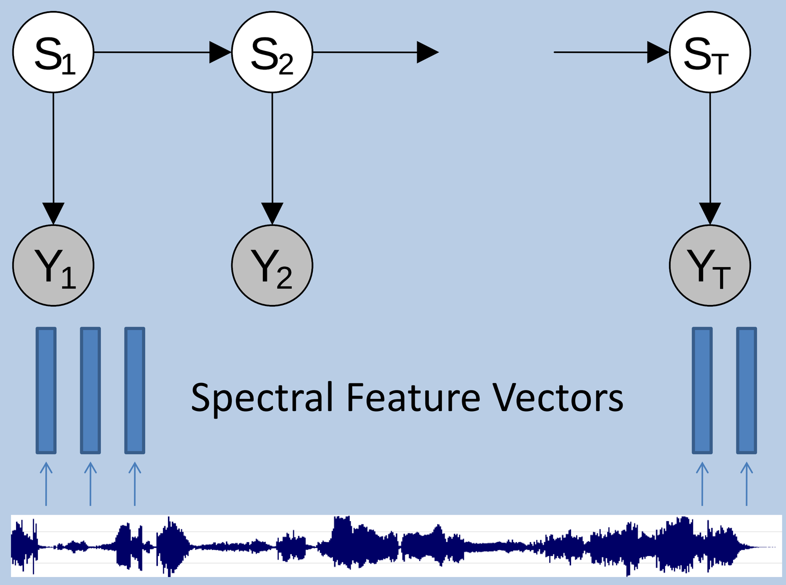
Hidden Markov Models are a form of Graphical Model when there is a sequential structure in the unknowns. We denote this with the picture on the right. The S's are unknown quantities, the Y's are observed quantities.



Speech Example

States correspond to phones: "dh", "a", "iy"

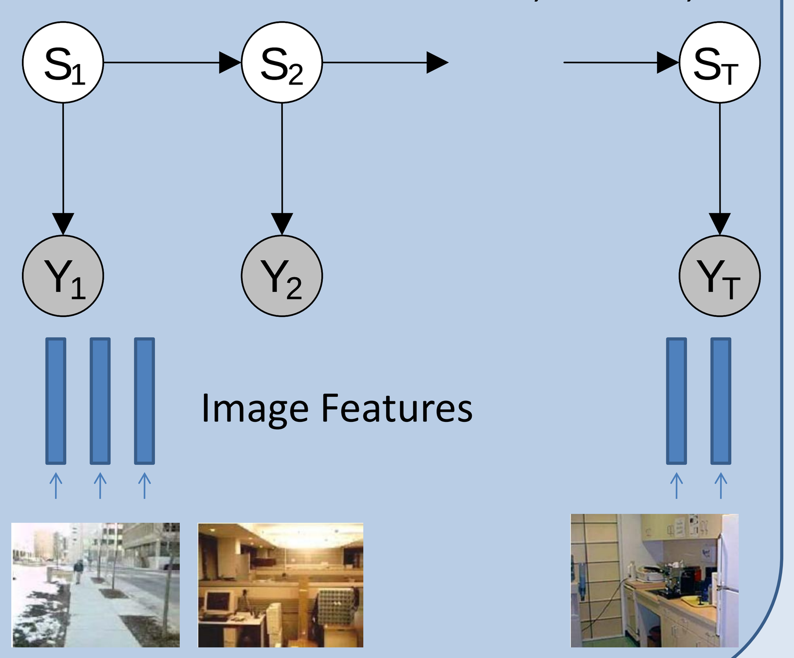
Hidden Markov Models dominate the field of automatic speech recognition. For example, the unknown quantities are the *phones* (short speech sound) and the observed quantities are *spectral feature vectors* which are quantities that are extracted from the speech waveform. In the example on the right, if the unknown quantity corresponds to the phones "dh-a" or "dh-i", this would correspond to the word "the".



Vision Example

States are locations: outdoor, kitchen, ...

Hidden Markov Models have been frequently used in vision and robotics. In the example on the right we sketch a system that was designed to perform context recognition. The unknown quantities in the HMM are the locations; the observations are (wavelet decomposition) features derived from pictures taken using a webcam.

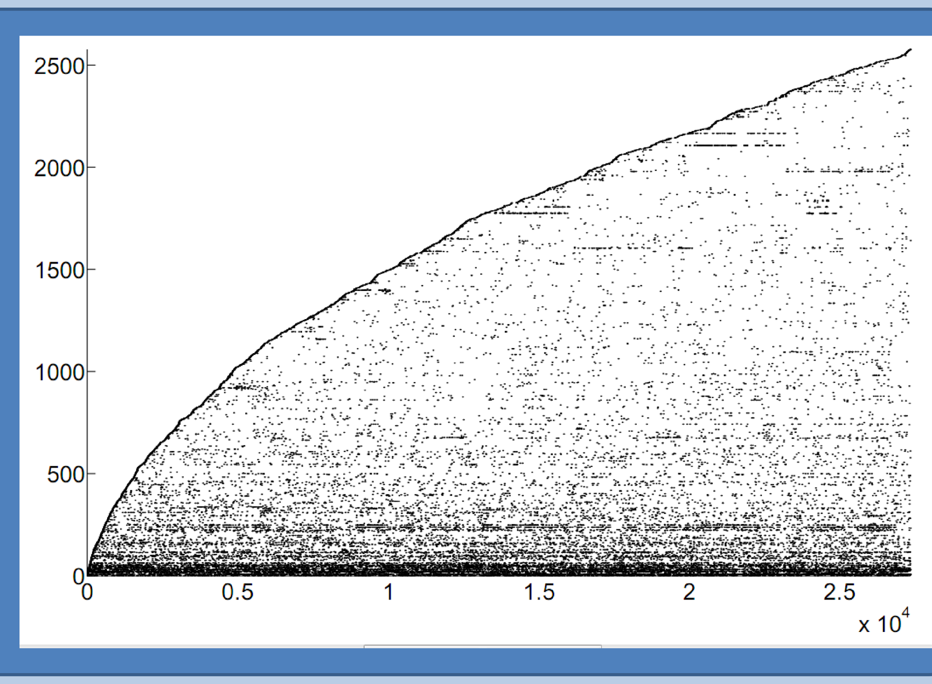


3. The Infinite HMM

What if we don't know how many hidden states there should be? What if we want to treat the number of hidden states as an unknown quantity?

Example:
We analyze the text "Alice in Wonderland" and assign each word type to a unique integer. Then, we plot each word's position (horizontal) against its word type (vertical).

Any training sequence is not expected to contain all possible observations

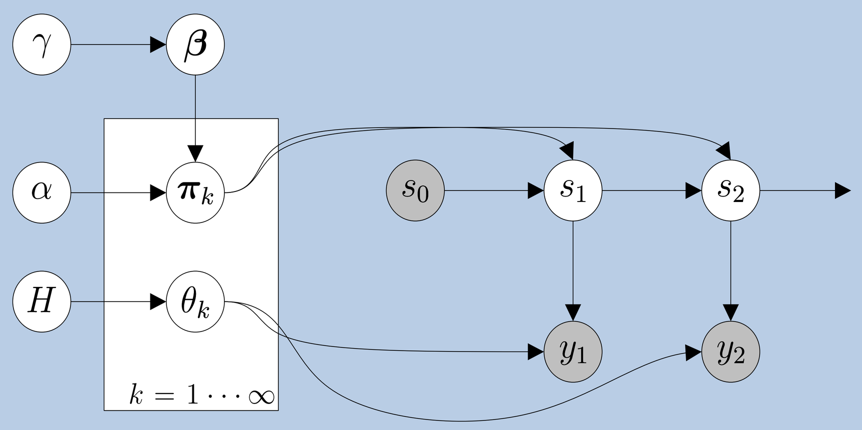


The *Infinite Hidden Markov Model* has a potentially unbounded number of states but we encode a preference for a small number of states.

The Graphical Model

$$\forall j \in \{1 \cdots \infty\}$$
$$\forall t \in \{1 \cdots T\}$$

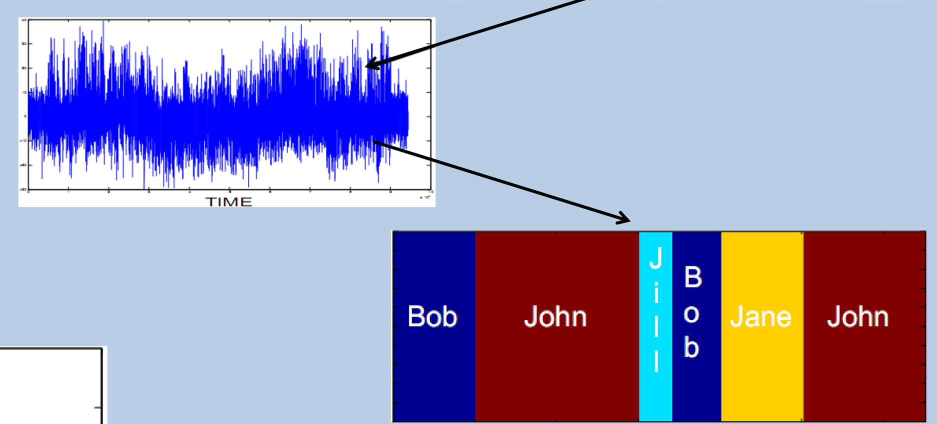
Encoding Small # States	$\boldsymbol{\beta}$	$\sim$	$\text{Stick}(\gamma)$
Encoding Transition Matrix	$\boldsymbol{\pi}_{j\cdot}$	$\sim$	$\text{DP}(\alpha, \boldsymbol{\beta})$
Observation Parameters	θ_j	$\sim$	H
Initial State	s_0	$=$	1
State Sequence	s_t	$\sim$	$\boldsymbol{\pi}_{s_{t-1},\cdot}$
Observation Sequence	y_t	$\sim$	$F(y_t \mid \theta_{s_t})$

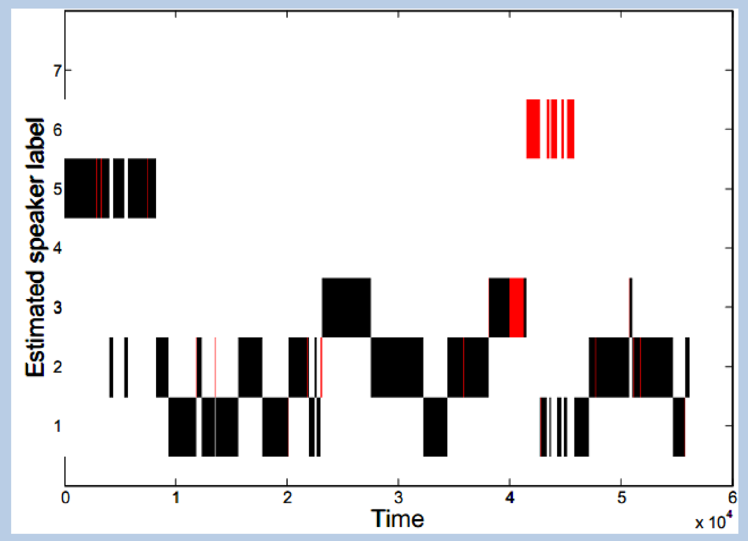


Speech Example

Consider a system that performs *speech diarization*: a recording device records a meeting and we want to automatically find out who is speaking when.

The unknown quantities now correspond to which speaker is talking. The iHMM does not assume the number of speakers; it will automatically adjust this to the observed speech signal.



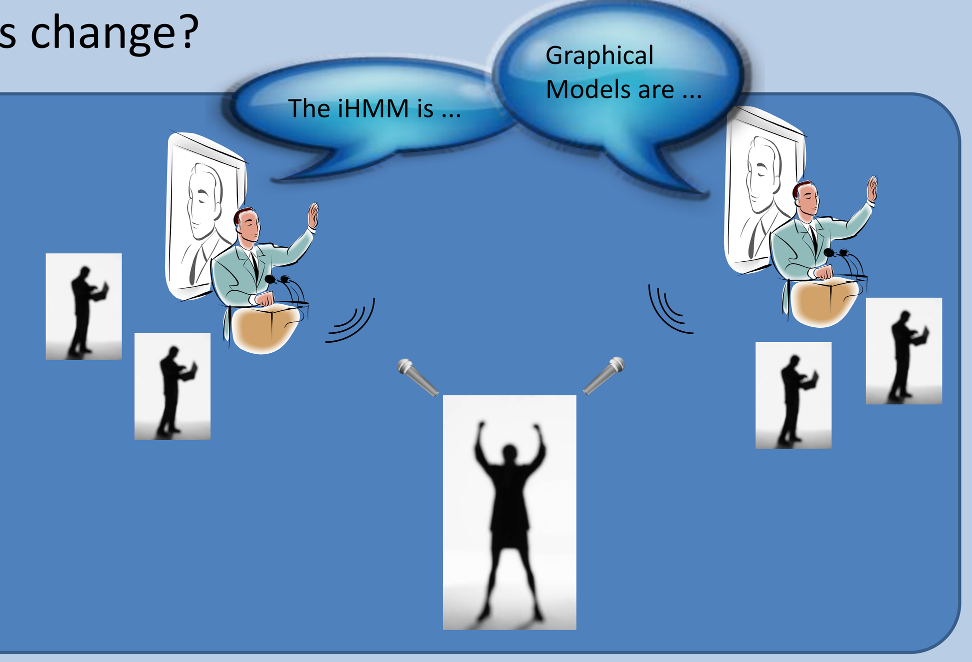


The plot on the right displays the estimated speaker label (vertical). Black indicates a correct labelling, red indicates a mistake with respect to a human labelled diarization sequence.

4. The Infinite Factorial HMM

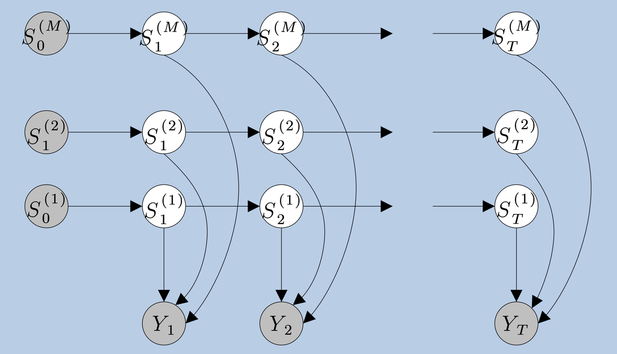
What if not one unknown quantity changes state at any one time but multiple independent quantities change?

The Cocktail Party Problem:
A group of academics is attending a poster session at a conference. Different speakers (close to each other) are discussing their work. How do you filter out one speaker's discourse?



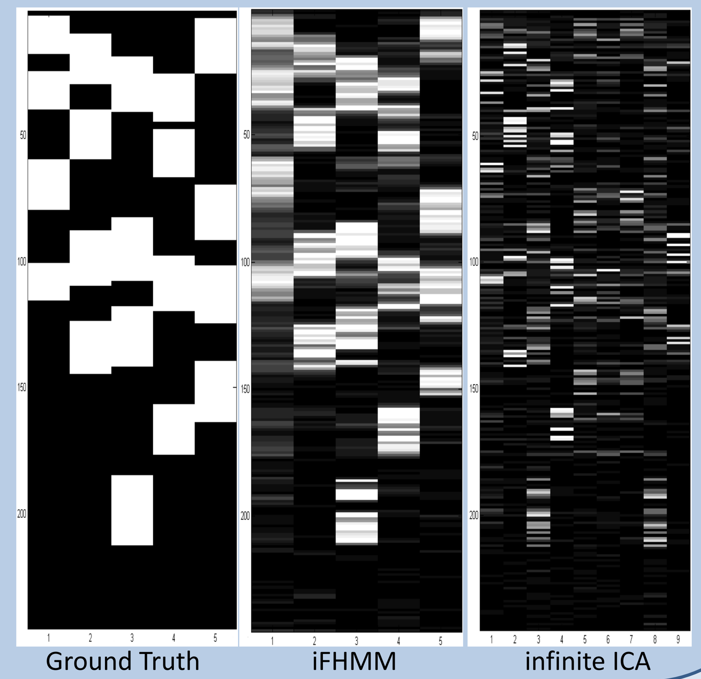
The *Infinite Factorial Hidden Markov Model* has multiple independent chains of unknown states.

(Partial) Graphical Model on the right



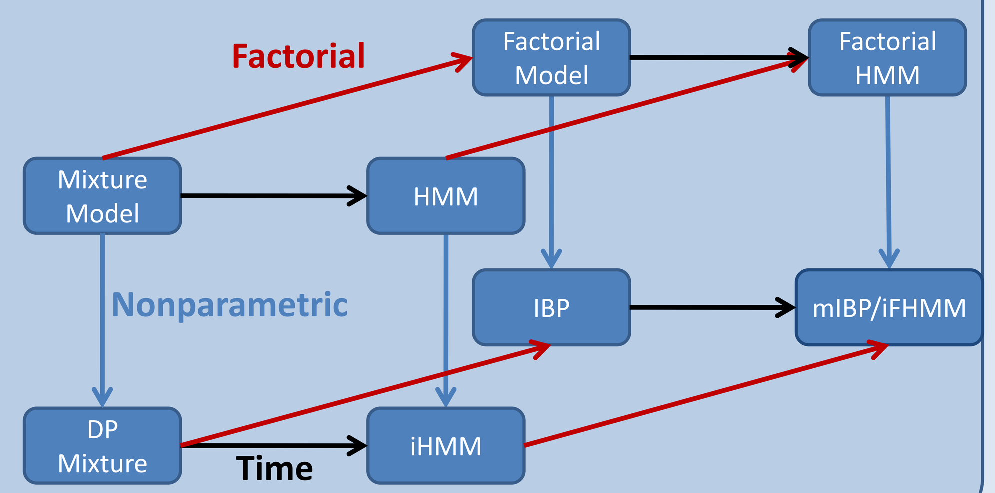
Speech Example

We recorded speech from 5 people. The left plot shows time on the vertical axis, speakers on the horizontal axis. White denotes when a particular speaker is talking, black when he is silent. The middle plot shows a sample from the iFHMM solution: it has discovered both the number of speakers and a reasonably accurate segmentation. The right plot shows the solution found by infinite Independent Component Analysis.



The Landscape

Graphical Models have unified many known models (Mixture, HMM, ...). This unifying picture has also been a guide to extending known models in different directions. Our work covers the *nonparametric time series models.*



References

- *Context-based vision system for place and object recognition*
  A Torralba, KP Murphy, WT Freeman, MA Rubin, Computer Vision, 2003
- *Speech and Language Processing*
  Daniel Jurafsky and James H. Martin, Prentice Hall, 2008
- *The Infinite Hidden Markov Model*
  M.J. Beal, Z. Ghahramani, C.E. Rasmussen, NIPS 2002
- *An HDP-HMM for Systems with State Persistence*
  E.B. Fox, E.B. Sudderth, M.I. Jordan, A.S. Willsky, ICML 2008
- *Beam Sampling for the Infinite Hidden Markov Model*
  J. Van Gael, Y. Saatci, Y.W. Teh, Z. Ghahramani, ICML 2008
- *The Infinite Factorial Hidden Markov Model*
  J. Van Gael, Y.W. Teh, Z. Ghahramani, NIPS 2008

†Jurgen Van Gael is supported by a Microsoft Research Scholarship, joint work with Yee Whye Teh (Gatsby unit, University College London), Yunus Saatçi