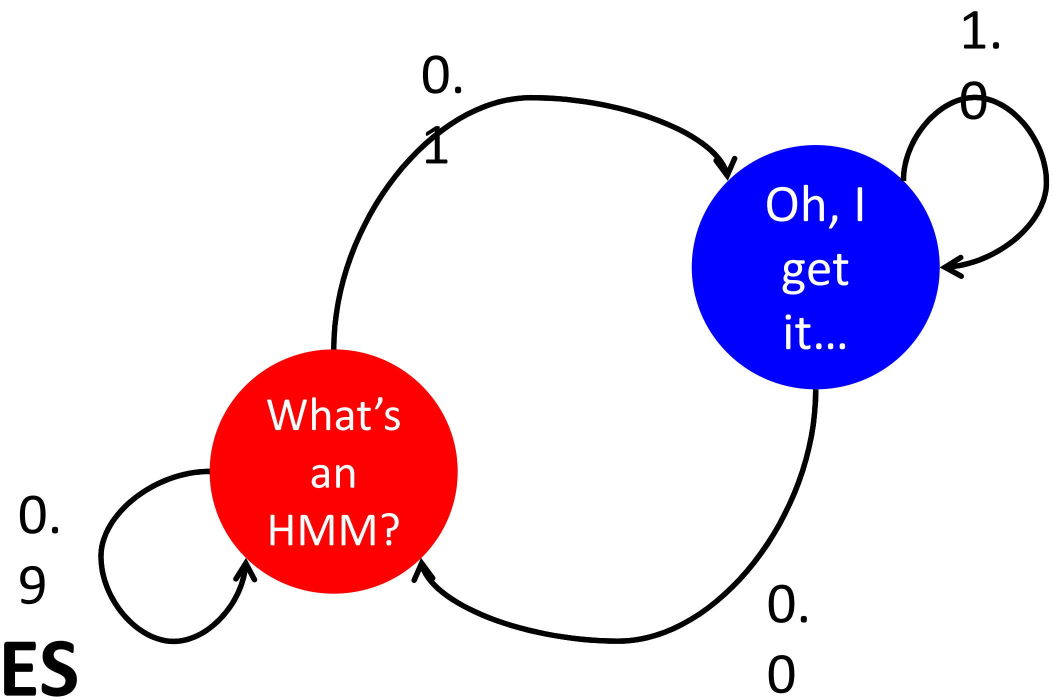
HIDDEN MARKOV MODELS … IN 10 MINUTES



DAVID R. THOMPSON
JET PROPULSION LABORATORY, CALIFORNIA INSTITUTE OF TECHNOLOGY.
DAVID.R.THOMPSON@JPL.NASA.GOV

KECK INSTITUTE STUDY "DIGGING DEEPER: COMPUTATIONAL CHALLENGES IN ASTRONOMY"
10 JUNE 2011

 GOVERNMENT SUPPORT ACKNOWLEDGED.

What is a Hidden Markov Model?

- A probability distribution over time series
- System evolves through discrete states
- Observations can be noisy / irregularly spaced
- Computationally-efficient representation of *some* temporal relationships
- Can formally compute optimal observation times (to best disambiguate the state sequence)

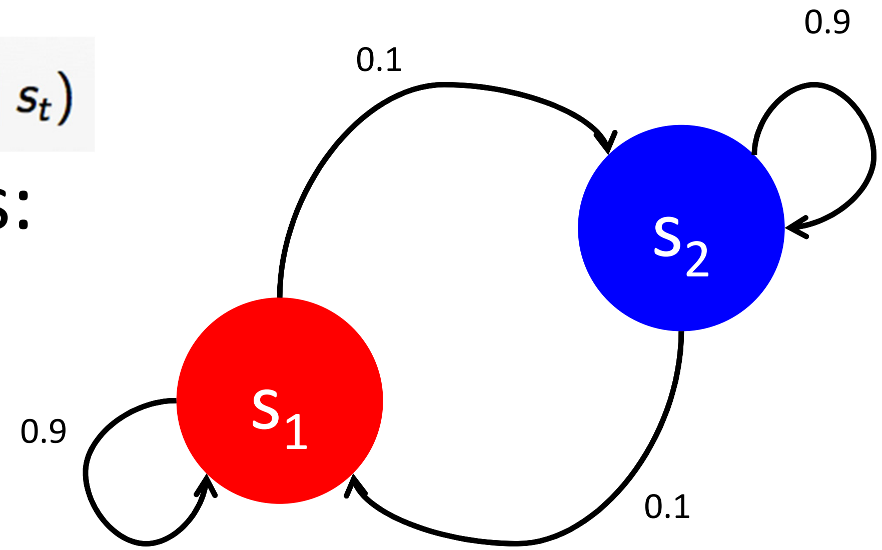
Markov system

- states s=1,2,…n
- time steps t = 1,2,…
- Given any state, the future is independent of the past:

$$P(s_{t+1} \mid s_t, s_{t-1}, \dots, s_0) = P(s_{t+1} \mid s_t)$$

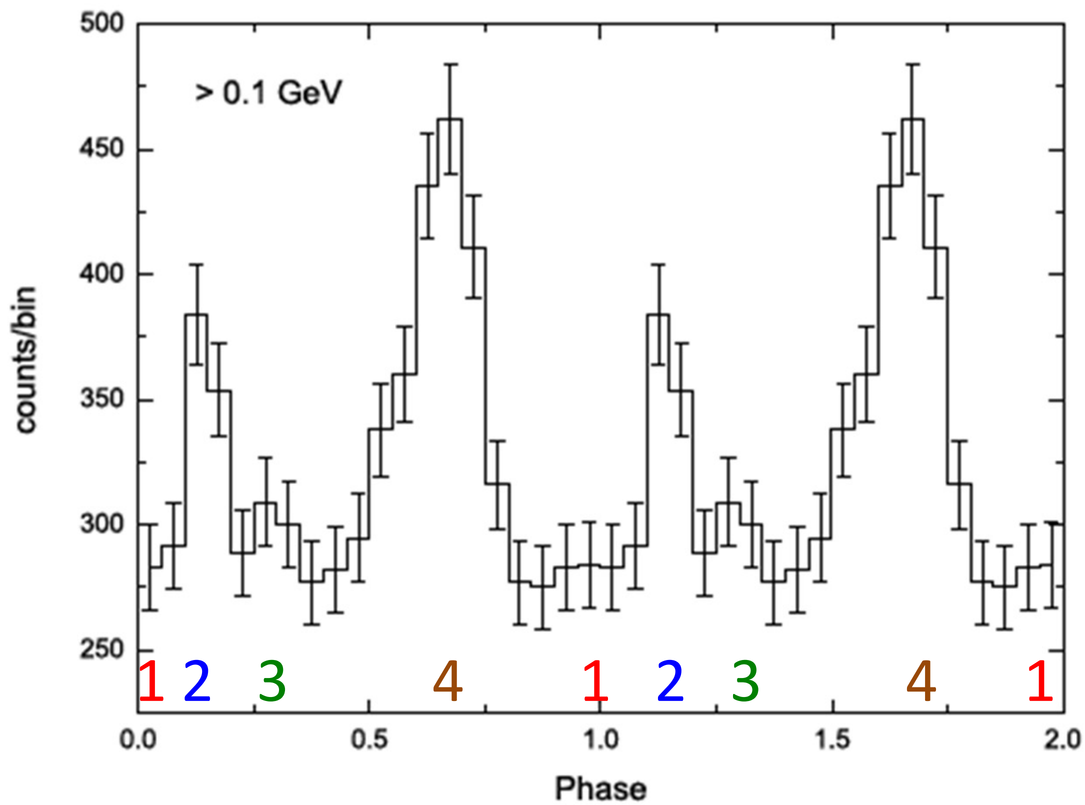
- Transition probabilities:

$$a_{k,l} = P(s_{t+1} = l \mid s_t = k)$$



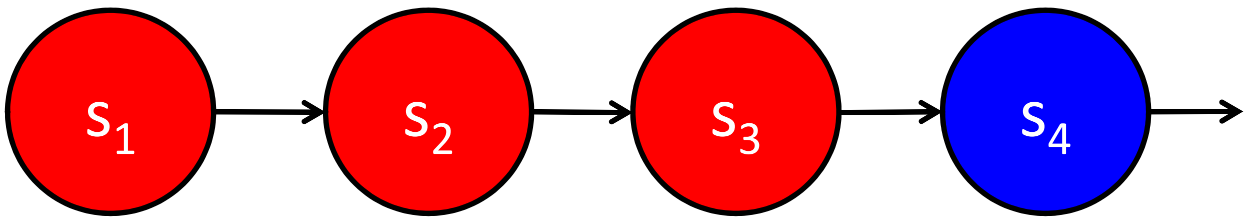
Equations from Jia Li (http://www.stat.psu.edu/~jiali)

A four-state Markov system?



Gamma-Ray Light Curve for PSR J2021+4026 as observed by Fermi/LAT in the range of 0.1 GeV -300 GeV (from Trepl et al. 2010)

Can efficiently compute probabilities of any state sequence



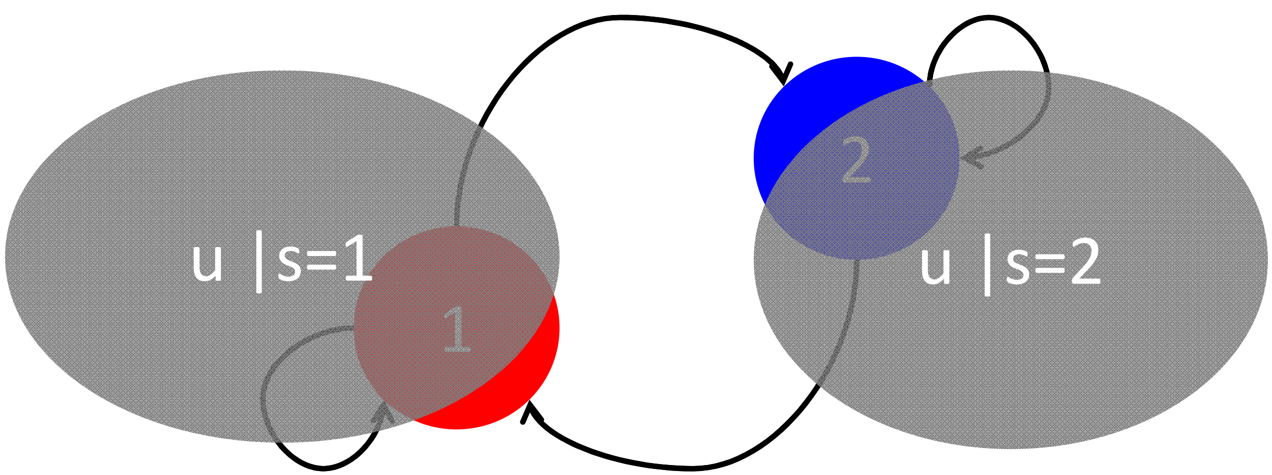
$$P(s_1, s_2, \ldots, s_T) = P(s_1)P(s_2|s_1)P(s_3|s_2)\cdots P(s_T|s_{T-1})$$
$$= \pi_{s_1} a_{s_1,s_2} a_{s_2,s_3} \cdots a_{s_{T-1},s_T} \,.$$

Initial probability of state s_1

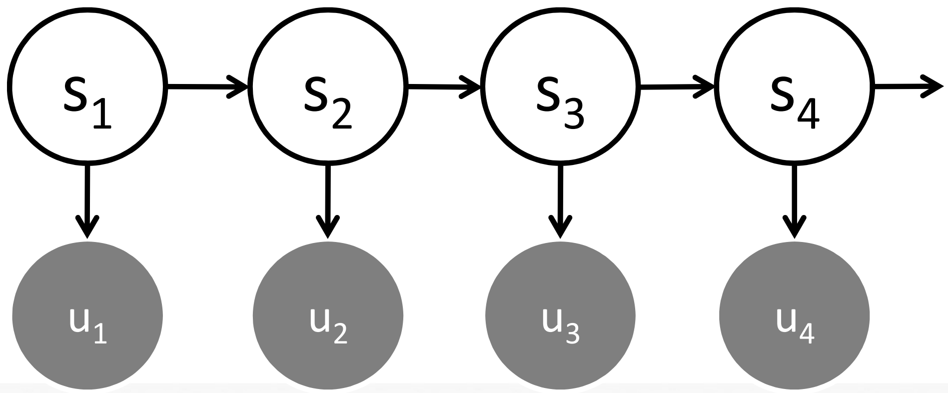
Equations from Jia Li (http://www.stat.psu.edu/~jiali)

Hidden markov models

- We can't observe the state directly, but instead get noisy observations u_1, u_2, …
- These depend only only on the current state
- They can be continuous, discrete, uni- or multivariate.



Can efficiently compute joint probabilities of any state/observation sequence



$$
\begin{aligned}
P(\mathbf{u}, \mathbf{s}) &= P(\mathbf{s})P(\mathbf{u} \mid \mathbf{s}) \\
&= \pi_{s_1} b_{s_1}(u_1) a_{s_1,s_2} b_{s_2}(u_2) \cdots a_{s_{T-1},s_T} b_{s_T}(u_T)\ .
\end{aligned}
$$

$$
\begin{aligned}
P(\mathbf{u}) &= \sum_{\mathbf{s}} P(\mathbf{s})P(\mathbf{u} \mid \mathbf{s}) \qquad \textit{total prob. formula} \\
&= \sum_{\mathbf{s}} \pi_{s_1} b_{s_1}(u_1) a_{s_1,s_2} b_{s_2}(u_2) \cdots a_{s_{T-1},s_T} b_{s_T}(u_T)
\end{aligned}
$$

Equations from Jia Li (http://www.stat.psu.edu/~jiali)

What questions can you ask an HMM?

1. **State estimation:** What is $P(s_i=x \mid u_1, u_2, \dots u_T)$
   – Efficient dynamic programming solution – "Forward-Backward algorithm," **scales $O(TN^2)$**

2. **Most Probable Sequence:** Given u_1, u_2, u_3…, what is the most probable state sequence and what is that probability?
   – Uses Dynamic Programming – "Viterbi algorithm," **scales $O(TN^2)$**

3. **Learning HMMs:** Given u_1, u_2, u_3…, what is the maximum likelihood HMM that could have produced this string of observations?
   – Uses the Expectation Maximization – "Baum-Welsh algorithm," **scales $O(TN^2)$**

Other sources of information

- L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proc. of the IEEE, Vol.77, No.2, pp.257--286, 1989.
  - http://ieeexplore.ieee.org/iel5/5/698/00018626.pdf?arnumber=18626
- Andrew Moore's tutorial
  - www.autonlab.org/tutorials/hmm.html
- For continuous domains, **linear dynamical systems**