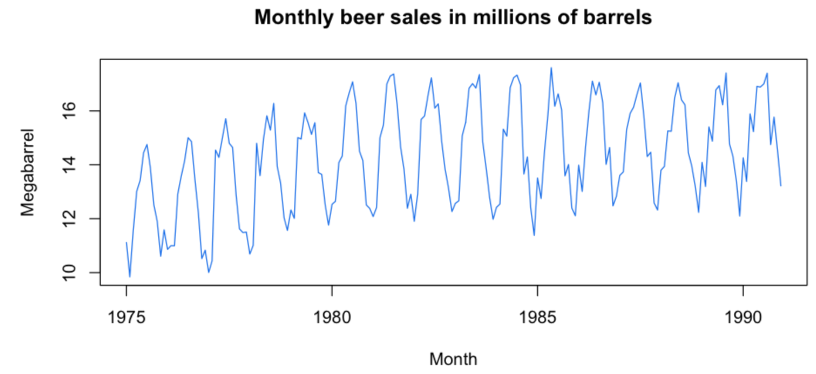
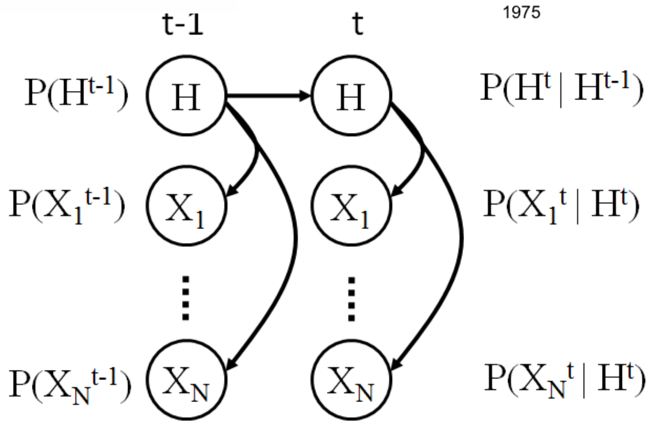
Dynamic Bayesian Networks & Hidden Markov Models – An Introduction





Overview

- In this lecture and lab
  - Introduction to Dynamic Bayesian Networks
  - Markov Chains
  - Hidden Markov Models (HMMs)
  - Models of sequence data / Time-Series
  - Scoring likelihoods of sequences

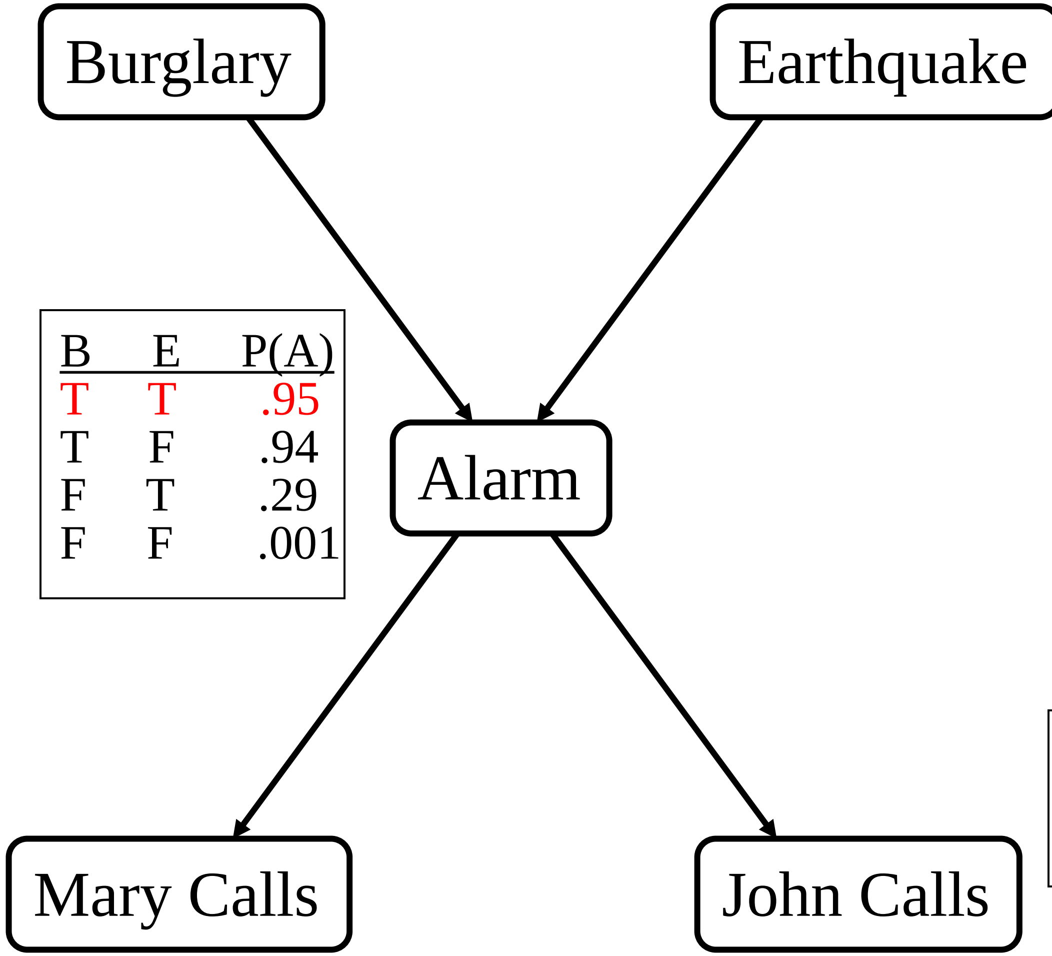
Reading

- Russell & Norvig:
  - Chapter 15, Section 1
- Rabiner’s Paper
- Stamp’s paper

Bayesian Networks Recap

- Method to store joint distribution
- Defined as a directed acyclic graph with local conditional dependencies

Bayesian Network Example



P(B)
.001

P(E)
.002

B	E	P(A)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(M)
T	.70
F	.01

A	P(J)
T	.90
F	.05

Retrieving Probabilities from the Conditional Distributions

$$P(x_1,\ldots,x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(x_i))$$

e.g. Probability of John and Mary Calling, Alarm sounding, no Burglary and no Earthquake:

P(J & M & A & ¬B & ¬E)

= P(J|A)P(M|A)P(A|¬B,¬E)P(¬B)P(¬E)

= 0.9 x 0.7 x 0.001 x 0.999 x 0.998

= 0.00062

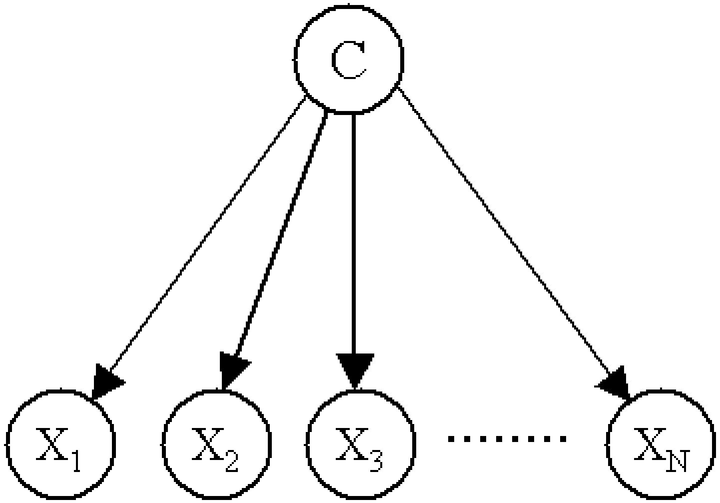
Bayesian Networks Recap

- Method to store joint distribution
- Defined as a directed acyclic graph with local conditional dependencies
- Evidence can be entered in to the model using any of the nodes
- Inference can be used to determine the probability distribution over remaining nodes given the observed evidence

Classification / Clustering

Classification:

- Naïve Bayes Classifier assumes independence between variables given the class: p(X|C)
- Can calculate p(C|X) using Bayes Rule or inference



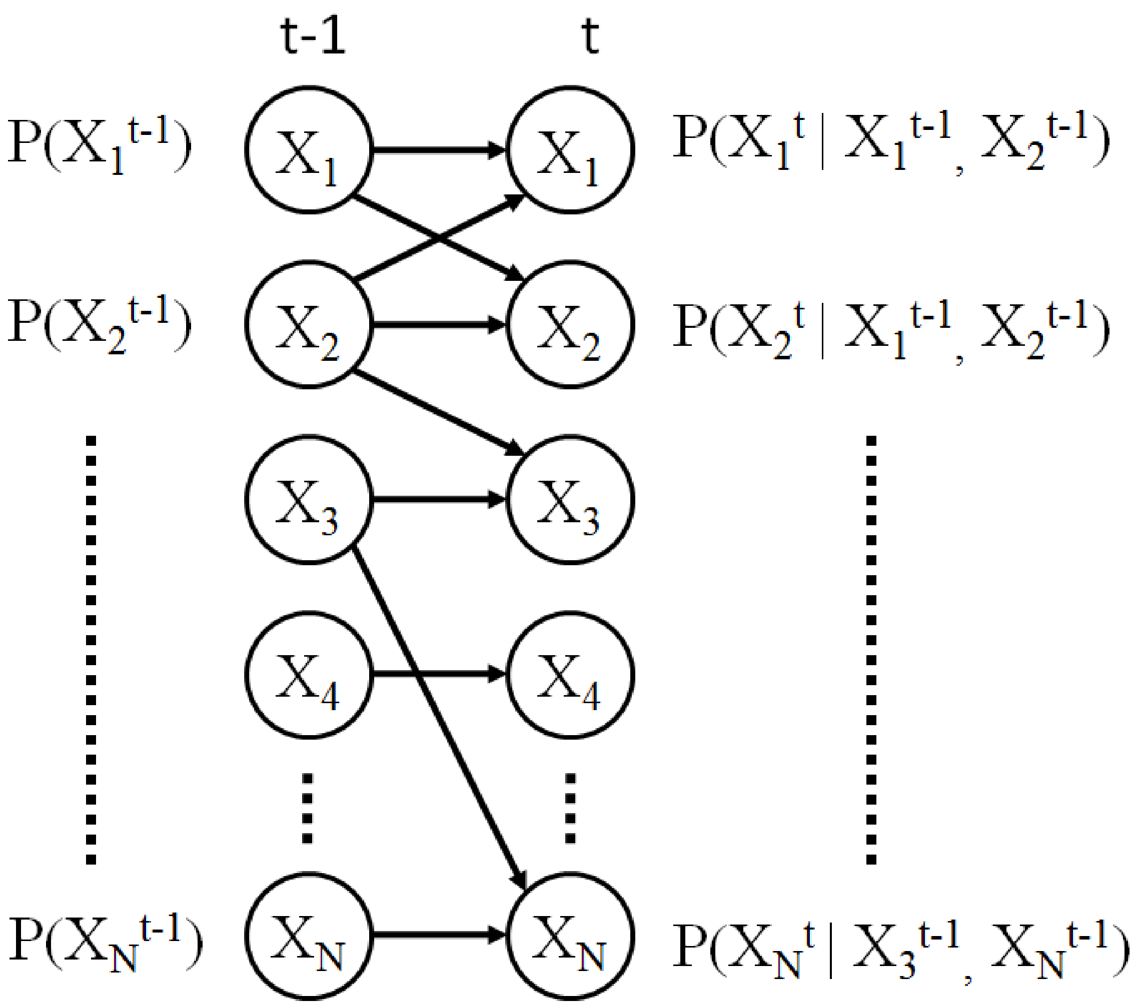
Clustering

- "Hidden" variables to model unmeasured features
- Make *C* a hidden variable

Dynamic Bayesian Networks

- Nodes represent variables at distinct time slices
- Directed links between nodes
  - Within the same "time slice"
  - Between "time slices"
- Often assume 1^{st} Order Markov Assumption:
  - "Links only need to span *adjacent* time slices in order to make predictions over time"

Dynamic Bayesian Networks



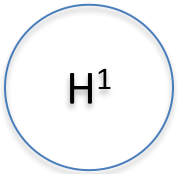
Markov Chains

- Essentially links a number of possible state transitions over a sequence, *H*, or time window
- Requires:
  - An initial probability distribution over the states
  - A transition probability distribution determining state changes
- Assumes 1st order Markov Assumption:
  - only info at time t is needed to calculate the info at time: $P(H^t|H^{t-1})$

Brunel University London

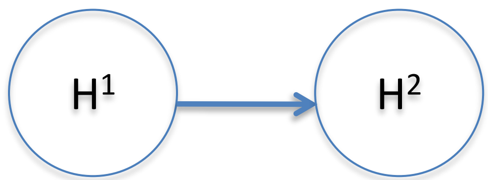
Markov Chains

- Initial distribution: $\pi = [0.5 \quad 0.5]$



Markov Chains

- Initial distribution: $\pi = [0.5 \quad 0.5]$

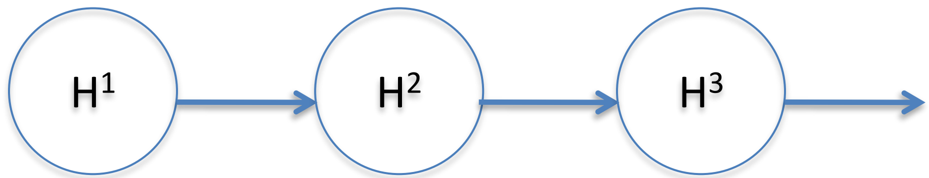


- Transition Distribution Matrix:

$$\begin{array}{cc} & \text{To} \\ A = & \begin{bmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{bmatrix} \text{From} \end{array}$$

Markov Chains

- Initial distribution: $\pi = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}$



- Transition Distribution Matrix:

$$\begin{array}{cc} & \text{To} \\ A = & \begin{bmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{bmatrix} \text{From} \end{array}$$

Markov Chains

- Initial distribution: $\pi = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}$

P(State 1 at t=1) P(State 2 at t=1)

- Transition Distribution Matrix:

P(State 2 at t | State 1 at t-1)

$$\begin{array}{ccc} & \text{To} & \\ A = & \begin{bmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{bmatrix} & \text{From} \end{array}$$

P(State 2 at t | State 2 at t-1)

Markov Chains

- To calculate the probability of the sequence:

  H = 1,2,1,2

  Given the Markov Chain, M: $P(H|M) = \pi_1\, a_{12}\, a_{21}\, a_{12}$

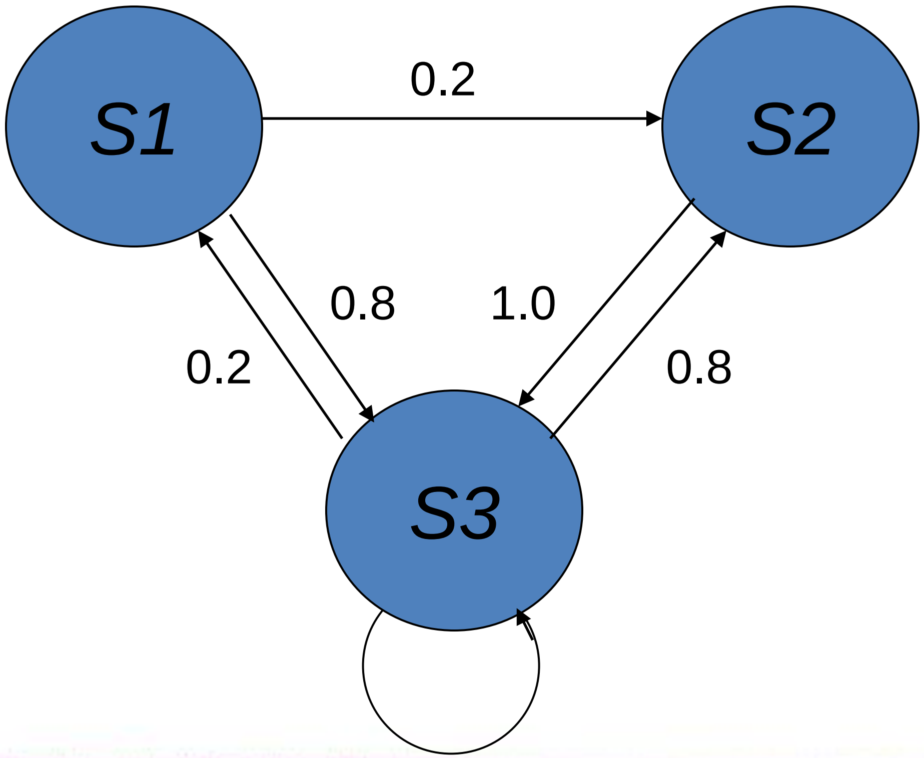
- Start with the initial probability $\pi_1 = 0.5$
- And multiply through with each transition based upon the transition matrix: $\pi_1\, a_{12}\, a_{21}\, a_{12}$

  = 0.5 x 0.1 x 0.3 x 0.1

  = 0.0015 (often these are very small numbers)

State Transition Diagram

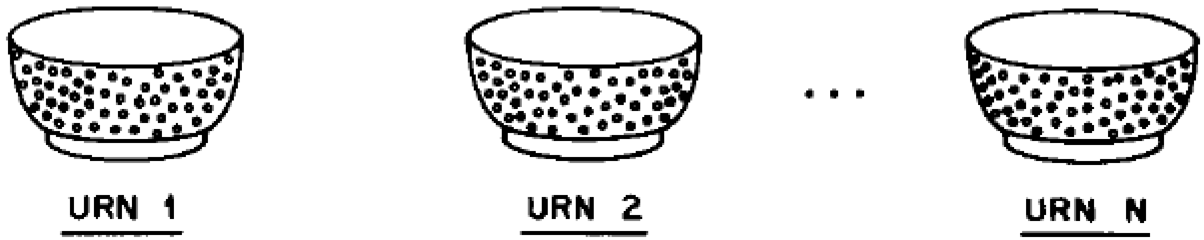
- Visualising the state transition probabilities (here for a Markov chain with three states)



Demo

http://setosa.io/blog/2014/07/26/markov-chains/

Hidden States



- Imagine there is an urn of coloured balls with 5 red balls and 3 green balls: P(G) = 3/8
- Now imagine there is another urn with 2 red balls and 6 green, P(G) = 6/8
- Someone hands you an urn *from behind a curtain* and asks you the P(G)
- After selecting a ball and returning the bag, the process is repeated

Brunel University London

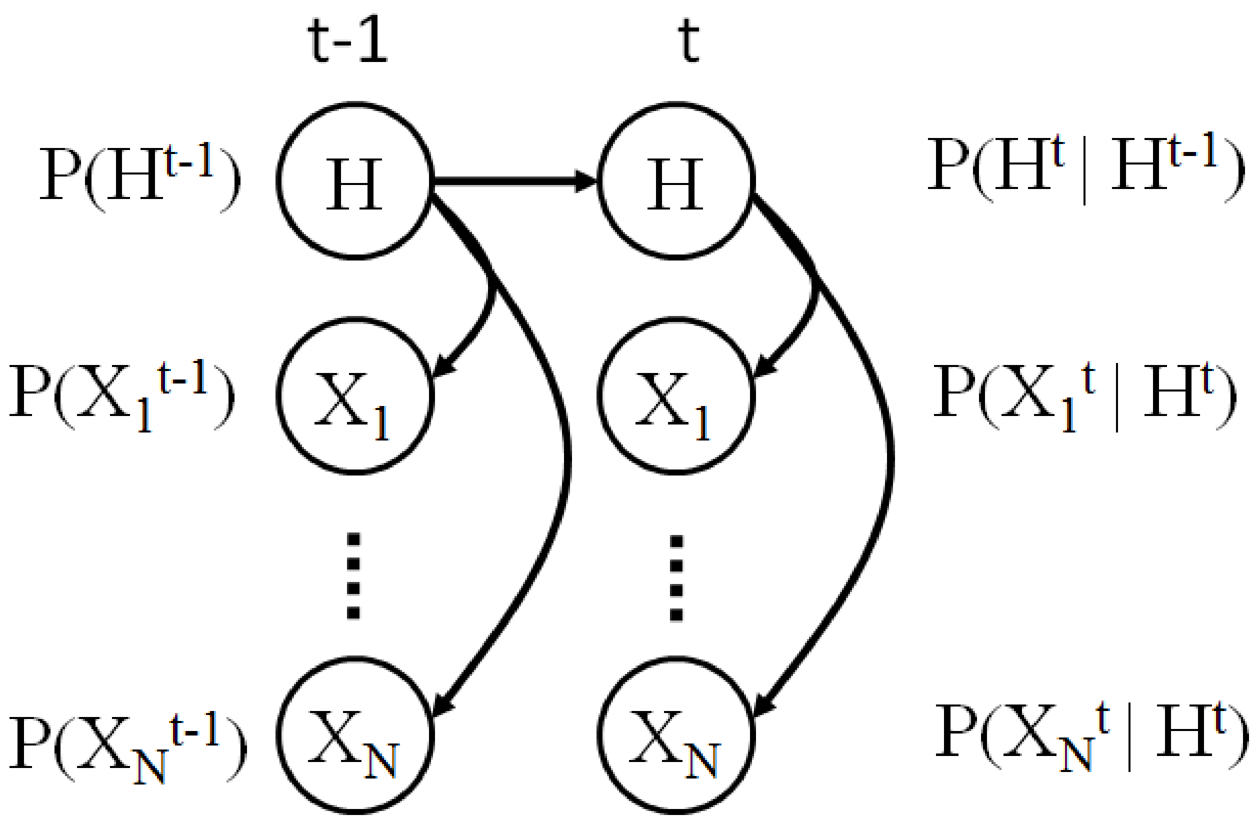
Hidden Markov Models

- We can model this using a HMM:
  - Consists of a Markov Chain, *H*, representing *hidden* states
  - Linked by possible transitions between time points
  - The current state determines the probability distribution of the next state
  - Markov assumption assumed: $P(H_t | H_{t-1})$

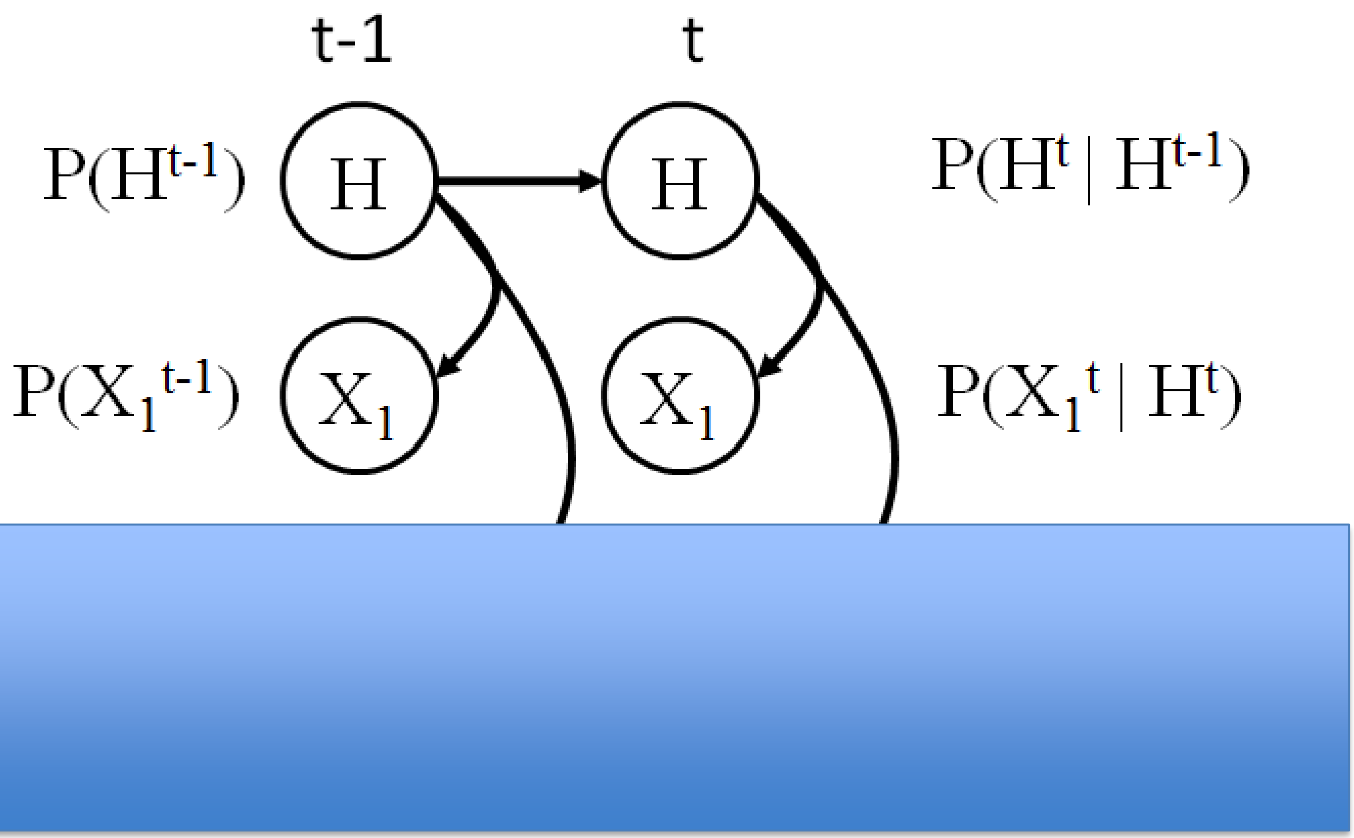
Hidden Markov Models – Emission Probabilities

- As well as hidden nodes, there are observed nodes, *X*
- These are linked directly to the hidden node in their own "time-slice"
- The probability distribution of these observed nodes is conditioned upon the hidden state
- $P(X_t | H_t)$

Hidden Markov Models



Hidden Markov Models



Example:

- Given the following HMM, *M* :
  - Initial State Distribution - P($H^1 = i$): $\pi = \begin{bmatrix} 1 & 0 \end{bmatrix}$
  - Transition Probability Distribution – P($H^t = j \mid H^{t-1} = i$):

$$\begin{array}{rcl} & t & \\ A = & \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} & t\text{-}1 \end{array}$$

- Probability of hidden state seq: {1, 2, 2}:

$$\mathrm{P(H \mid M)} = \pi_1 \, \mathrm{a}_{12} \, \mathrm{a}_{22}$$

= 1 x 0.2 x 0.6 = 0.12

Example:

- Given the following HMM, *M* :
  - Initial State Distribution - P($H^1 = i$): $\pi = \begin{bmatrix} 1 & 0 \end{bmatrix}$
  - Transition Probability Distribution – P($H^t = j \mid H^{t-1} = i$):

$$\begin{array}{cc} & t \\ A = & \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} & t\text{-}1 \end{array}$$

  - We also need an Emission (or Sensor) Distribution:

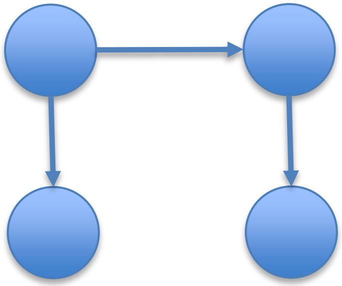
$$\begin{array}{cc} & X \\ B = & \begin{bmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{bmatrix} & H \end{array}$$

Brunel University London

Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{cc} t & \\ \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} & t\text{-}1 \end{array} \qquad B = \begin{array}{cc} X & \\ \begin{bmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{bmatrix} & H \end{array}$$

- Use this to generate a sequence



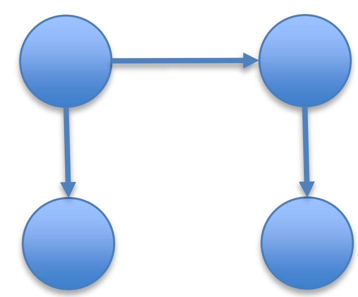
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{cc} & t \\ & \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} & t\text{-}1 \end{array} \qquad B = \begin{array}{cc} & X \\ & \begin{bmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{bmatrix} & H \end{array}$$

- Use this to generate a sequence

H

X



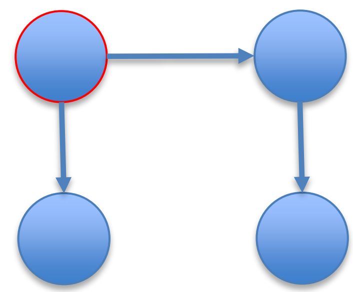
Example:

$$\pi = \begin{bmatrix} \boxed{1} & 0 \end{bmatrix} \qquad A = \begin{array}{cc} & t \\ & \\ \end{array} \!\!\!\!\!\! \begin{array}{c} \\ \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} \end{array} \ t\text{-}1 \qquad B = \begin{array}{c} X \\ \begin{bmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{bmatrix} \end{array} \ H$$

- Use this to generate a sequence

H 1

X



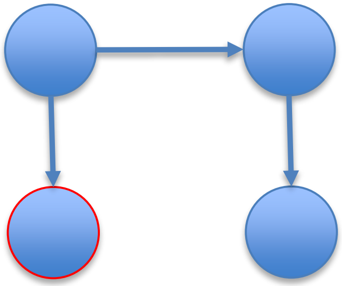
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{cc} t & \\ \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} & t\text{-}1 \end{array} \qquad B = \begin{array}{cc} X & \\ \begin{bmatrix} \boxed{0.5} & 0.5 \\ 0.1 & 0.9 \end{bmatrix} & H \end{array}$$

- Use this to generate a sequence

H 1

X 1



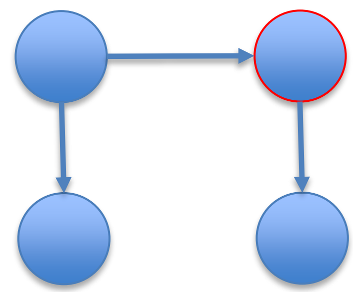
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{cc} t & \\ \begin{bmatrix} \boxed{0.8} & 0.2 \\ 0.4 & 0.6 \end{bmatrix} & t\text{-}1 \end{array} \qquad B = \begin{array}{cc} X & \\ \begin{bmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{bmatrix} & H \end{array}$$

- Use this to generate a sequence

H 1 1

X 1



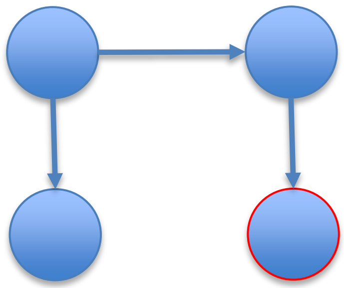
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{c} t \\ \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} \end{array} \; t\text{-}1 \qquad B = \begin{array}{c} X \\ \begin{bmatrix} 0.5 & \boxed{0.5} \\ 0.1 & 0.9 \end{bmatrix} \end{array} \; H$$

- Use this to generate a sequence

H 1 1

X 1 2



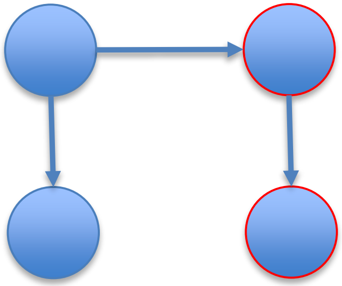
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{cc} t & \\ \begin{bmatrix} \boxed{0.8} & 0.2 \\ 0.4 & 0.6 \end{bmatrix} & t\text{-}1 \end{array} \qquad B = \begin{array}{cc} X & \\ \begin{bmatrix} 0.5 & \boxed{0.5} \\ 0.1 & 0.9 \end{bmatrix} & H \end{array}$$

- Use this to generate a sequence

H 1 1 1

X 1 2 2



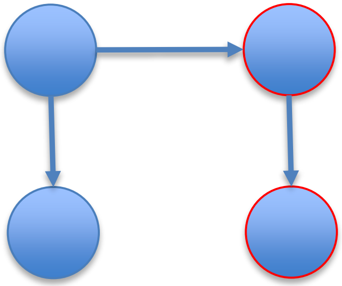
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{cc} \begin{array}{cc} t & \end{array} & \\ \begin{bmatrix} \boxed{0.8} & 0.2 \\ 0.4 & 0.6 \end{bmatrix} & t\text{-}1 \end{array} \qquad B = \begin{array}{cc} \begin{array}{cc} X & \end{array} & \\ \begin{bmatrix} \boxed{0.5} & 0.5 \\ 0.1 & 0.9 \end{bmatrix} & H \end{array}$$

- Use this to generate a sequence

H 1 1 1 1

X 1 2 2 1



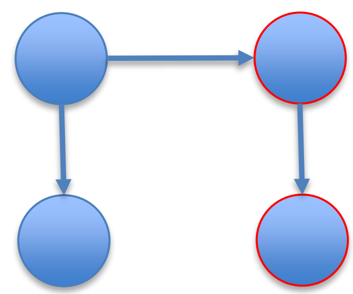
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{c} t \\ \begin{bmatrix} 0.8 & \boxed{0.2} \\ 0.4 & 0.6 \end{bmatrix} \end{array} \; t\text{-}1 \qquad B = \begin{array}{c} X \\ \begin{bmatrix} 0.5 & 0.5 \\ 0.1 & \boxed{0.9} \end{bmatrix} \end{array} \; H$$

- Use this to generate a sequence

H 1 1 1 1 2

X 1 2 2 1 2



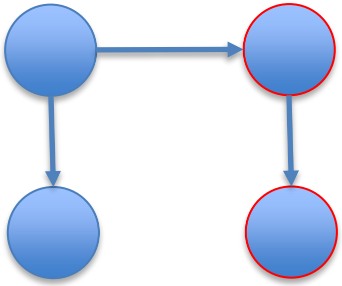
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{cc} t & \\ \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & \boxed{0.6} \end{bmatrix} & t\text{-}1 \end{array} \qquad B = \begin{array}{cc} X & \\ \begin{bmatrix} 0.5 & 0.5 \\ 0.1 & \boxed{0.9} \end{bmatrix} & H \end{array}$$

- Use this to generate a sequence

H 1 1 1 1 2 2

X 1 2 2 1 2 2



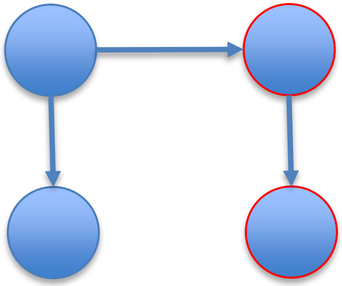
Example:

$$\pi = \begin{bmatrix}1 & 0\end{bmatrix} \qquad A = \begin{array}{cc} t & \\ \begin{bmatrix}0.8 & 0.2\\ 0.4 & \boxed{0.6}\end{bmatrix} & t\text{-}1 \end{array} \qquad B = \begin{array}{cc} X & \\ \begin{bmatrix}0.5 & 0.5\\ 0.1 & \boxed{0.9}\end{bmatrix} & H \end{array}$$

- Use this to generate a sequence

H 1 1 1 1 2 2 2

X 1 2 2 1 2 2 2



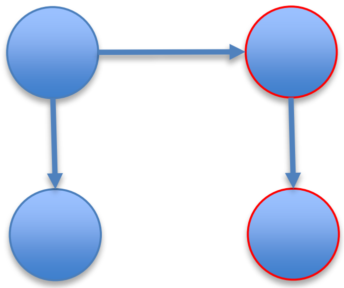
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{cc} t & \\ \begin{bmatrix} 0.8 & 0.2 \\ \boxed{0.4} & 0.6 \end{bmatrix} & t\text{-}1 \end{array} \qquad B = \begin{array}{cc} X & \\ \begin{bmatrix} \boxed{0.5} & 0.5 \\ 0.1 & 0.9 \end{bmatrix} & H \end{array}$$

- Use this to generate a sequence

H 1 1 1 1 2 2 2 1

X 1 2 2 1 2 2 2 1



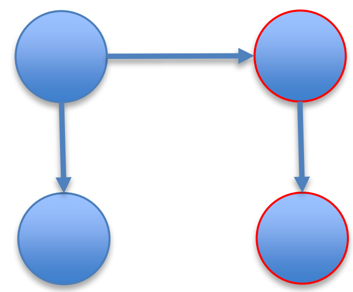
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{cc} \begin{array}{cc} & t \end{array} & \\ \begin{bmatrix} \boxed{0.8} & 0.2 \\ 0.4 & 0.6 \end{bmatrix} & t\text{-}1 \end{array} \qquad B = \begin{array}{cc} \begin{array}{cc} & X \end{array} & \\ \begin{bmatrix} 0.5 & \boxed{0.5} \\ 0.1 & 0.9 \end{bmatrix} & H \end{array}$$

- Use this to generate a sequence

H 1 1 1 1 2 2 2 1 1

X 1 2 2 1 2 2 2 1 2



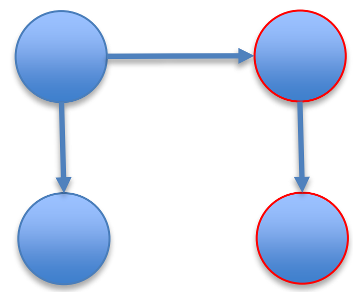
Example:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad A = \begin{array}{c} t \\ \begin{bmatrix} \boxed{0.8} & 0.2 \\ 0.4 & 0.6 \end{bmatrix} \end{array} \; t\text{-}1 \qquad B = \begin{array}{c} X \\ \begin{bmatrix} \boxed{0.5} & 0.5 \\ 0.1 & 0.9 \end{bmatrix} \end{array} \; H$$

- Use this to generate a sequence

H 1 1 1 1 2 2 2 1 1 1

X 1 2 2 1 2 2 2 1 2 1



Examples

- Given the following HMM, *M* to model tossing a coin using 3 hidden states:

$$\pi = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.2 & 0.3 & 0.5 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$B = \begin{array}{c} \text{Heads} \quad \text{Tails} \\ \begin{bmatrix} 0.5 & 0.5 \\ 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \end{array}$$

- You observe the sequence: X=H H H H T H T T T T

1) What state sequence is most likely?

Different ways to interpret “most likely”!

Brunel University London

Examples

- Given the following HMM, *M* to model tossing a coin using 3 hidden states:

$$\pi = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.2 & 0.3 & 0.5 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$\begin{array}{cc} & \begin{array}{cc} \text{Heads} & \text{Tails} \end{array} \\ B = & \begin{bmatrix} 0.5 & 0.5 \\ 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \end{array}$$

- You observe the sequence: X=H H H H T

2) What is the probability of the observation sequence and the “most likely” state sequence?

1) What state sequence is "most likely"?

$$\pi = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.2 & 0.3 & 0.5 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$\begin{array}{cc} & \begin{array}{cc} \text{Heads} & \text{Tails} \end{array} \\ B = & \begin{bmatrix} 0.5 & 0.5 \\ 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \end{array}$$

X=HHHHT

For now let's assume "most likely" state sequence is the one for which the probability of each individual observation is maximum: for each Heads, the most likely state is S2 and for each Tails, the most likely state is S3:

H = 2, 2, 2, 2, 3

2) Observing Seq & Most Likely Hidden Seq?

$$\pi = \begin{bmatrix} 0.2 & \boxed{0.3} & 0.5 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.2 & 0.3 & 0.5 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$\begin{array}{cc} & \text{Heads} \quad \text{Tails} \\ B = & \begin{bmatrix} 0.5 & 0.5 \\ 0.75 & 0.25 \\ 0.2 & 0.8 \end{bmatrix} \end{array}$$

X	H	H	H	H	T
H	2	2	2	2	3

$P(X,H \mid M) = P(X^1|H^1) \mathrm{x} P(H^1) \text{ x } P(X^2|H^2) \text{ x } P(H^2|H^1) \ldots$

$= (\boxed{0.3} \text{ x } 0.75) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (0.5 \text{ x } 0.8)$

Brunel University London

2) Prob of the Observation Seq & Most Likely Hidden Seq?

$$\pi = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.2 & 0.3 & 0.5 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$\begin{array}{cc} & \begin{array}{cc} \text{Heads} & \text{Tails} \end{array} \\ B = & \begin{bmatrix} 0.5 & 0.5 \\ \boxed{0.75} & 0.25 \\ 0.2 & 0.8 \end{bmatrix} \end{array}$$

X	H	H	H	H	T
H	2	2	2	2	3

P(*X*,*H* | *M*) = P(X^1|H^1)xP(H^1) x P(X^2|H^2) x P(H^2|H^1) ...

= (0.3 x $\boxed{0.75}$) x (0.3 x 0.75) x (0.3 x 0.75) x (0.3 x 0.75) x (0.5 x 0.8)

Brunel University London

2) Prob of the Observation Seq & Most Likely Hidden Seq?

$$\pi = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.2 & \boxed{0.3} & 0.5 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$\begin{array}{cc} & \begin{array}{cc} \text{Heads} & \text{Tails} \end{array} \\ B = & \begin{bmatrix} 0.5 & 0.5 \\ \boxed{0.75} & 0.25 \\ 0.2 & 0.8 \end{bmatrix} \end{array}$$

X	H	H	H	H	T
H	2	2	2	2	3

$$P(X,H \mid M) = P(X^1|H^1) \text{x} P(H^1) \text{ x } P(X^2|H^2) \text{ x } P(H^2|H^1) \ldots$$

$$= (0.3 \text{ x } 0.75) \text{ x } (\boxed{0.3} \text{ x } \boxed{0.75}) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (0.5 \text{ x } 0.8)$$

2) Prob of the Observation Seq & Most Likely Hidden Seq?

$$\pi = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.2 & \boxed{0.3} & 0.5 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$B = \begin{array}{cc} \text{Heads} & \text{Tails} \\ \end{array}$$
$$B = \begin{bmatrix} 0.5 & 0.5 \\ \boxed{0.75} & 0.25 \\ 0.2 & 0.8 \end{bmatrix}$$

X	H	H	H	H	T
H	2	2	2	2	3

$P(X,H \mid M) = P(X^1|H^1) \times P(H^1) \times P(X^2|H^2) \times P(H^2|H^1) \ldots$

$= (0.3 \times 0.75) \times (0.3 \times 0.75) \times (\boxed{0.3} \times \boxed{0.75}) \times (0.3 \times 0.75) \times (0.5 \times 0.8)$

Brunel University London

2) Prob of the Observation Seq & Most Likely Hidden Seq?

$$\pi = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.2 & \boxed{0.3} & 0.5 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$B = \begin{array}{c} \begin{array}{cc} \text{Heads} & \text{Tails} \end{array} \\ \begin{bmatrix} 0.5 & 0.5 \\ \boxed{0.75} & 0.25 \\ 0.2 & 0.8 \end{bmatrix} \end{array}$$

X	H	H	H	H	T
H	2	2	2	2	3

$P(X,H \mid M) = P(X^1|H^1) x P(H^1)$ x $P(X^2|H^2)$ x $P(H^2|H^1)$ ...

= (0.3 x 0.75) x (0.3 x 0.75) x (0.3 x 0.75) x ($\boxed{0.3}$ x $\boxed{0.75}$) x (0.5 x 0.8)

2) Prob of the Observation Seq & Most Likely Hidden Seq?

$$\pi = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.2 & 0.3 & \boxed{0.5} \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$\begin{array}{cc} & \text{Heads} \quad \text{Tails} \\ B = & \begin{bmatrix} 0.5 & 0.5 \\ 0.75 & 0.25 \\ 0.2 & 0.8 \end{bmatrix} \end{array}$$

X	H	H	H	H	T
H	2	2	2	2	3

$P(X,H \mid M) = P(X^1|H^1) \text{x} P(H^1) \text{ x } P(X^2|H^2) \text{ x } P(H^2|H^1) \ldots$

$= (0.3 \text{ x } 0.75) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (\boxed{0.5} \text{ x } 0.8)$

Brunel University London

2) Prob of the Observation Seq & Most Likely Hidden Seq?

$$\pi = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.2 & 0.3 & 0.5 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$\begin{array}{cc} & \text{Heads} \quad \text{Tails} \\ B = & \begin{bmatrix} 0.5 & 0.5 \\ 0.75 & 0.25 \\ 0.2 & \boxed{0.8} \end{bmatrix} \end{array}$$

X	H	H	H	H	T
H	2	2	2	2	3

$P(X,H \mid M) = P(X^1|H^1) \text{x} P(H^1) \text{ x } P(X^2|H^2) \text{ x } P(H^2|H^1) \ldots$

$= (0.3 \text{ x } 0.75) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (0.3 \text{ x } 0.75) \text{ x } (0.5 \text{ x } \boxed{0.8})$

Brunel University London

Typical Algorithms for HMMs

1. Given an observed sequence, *X* and a HMM, *M,* how do we compute its probability given the model?: p(*X* | *M*) - *Forward Algorithm*

2. Given the observed sequence and the model, how do we choose an *optimal* hidden state sequence?: *Most Probable Explanation using the Viterbi algorithm*

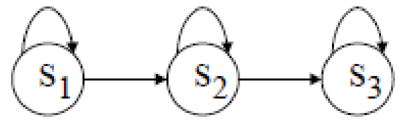
Typical Algorithms for HMMs

3. How do we adjust the model parameters to maximise the probability of the observed sequence given the model?: *Learning using the Expectation Maximisation algorithm*

Sequences for Characters

Exercise: character recognition with HMM(1)

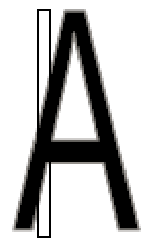
- The structure of hidden states:



- Observation = number of islands in the vertical slice.
- HMM for character 'A' :

Transition probabilities: $\{a_{ij}\} = \begin{pmatrix} .8 & .2 & 0 \\ 0 & .8 & .2 \\ 0 & 0 & 1 \end{pmatrix}$

Observation probabilities: $\{b_{jk}\} = \begin{pmatrix} .9 & .1 & 0 \\ .1 & .8 & .1 \\ .9 & .1 & 0 \end{pmatrix}$

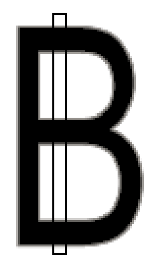


1 1 2 2 2 2 2 1 1

- HMM for character 'B' :

Transition probabilities: $\{a_{ij}\} = \begin{pmatrix} .8 & .2 & 0 \\ 0 & .8 & .2 \\ 0 & 0 & 1 \end{pmatrix}$

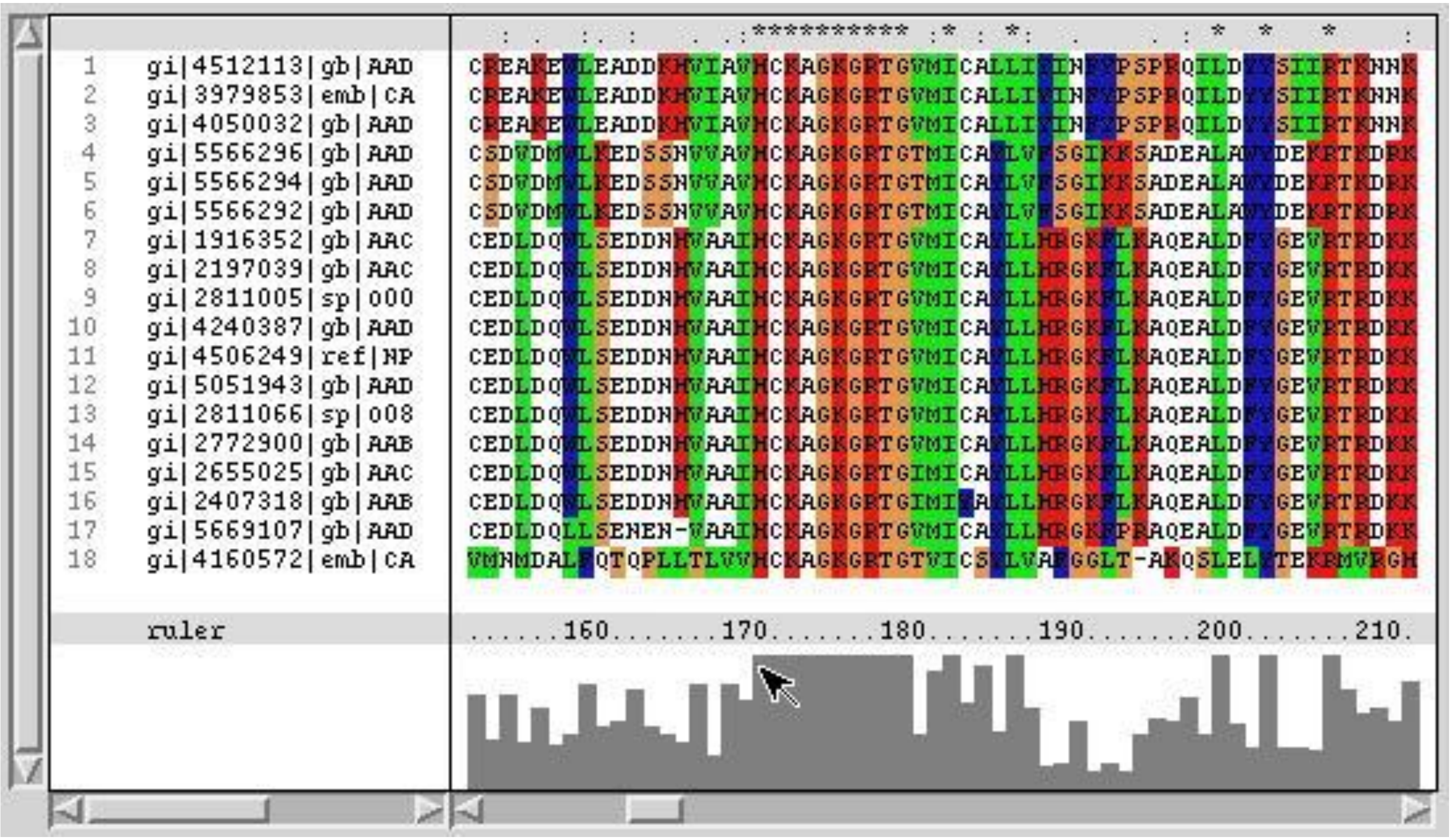
Observation probabilities: $\{b_{jk}\} = \begin{pmatrix} .9 & .1 & 0 \\ 0 & .2 & .8 \\ .6 & .4 & 0 \end{pmatrix}$



1 1 3 3 3 3 3 3 3

Brunel University London

Aligning Sequences of Nucleotides



- From the UBC Bioinformatics Centre (UBiC)

Brunel University London

Applications

- Dynamic mouse gesture recognition

http://www.youtube.com/watch?v=0CNJ2fCj4xQ

- Text Modelling – sequences of words

https://projects.haykranen.nl/markov/demo/

- Speech Recognition (Rabiner's paper)
- Tracking Speech to create avatar:

https://www.youtube.com/watch?v=J95TiRLoci8

Applications

- Dynamic Bayesian Network for monitoring / tracking behaviour:

https://www.youtube.com/watch?v=nEsS9dukgvM

- Dynamic Bayesian Network Automatic Driving

https://www.youtube.com/watch?v=aWtd8ZCcuZg

Summary

- Sequences and how to model them
- How to build HMMs
  - How to generate sequences
  - Calculate useful probabilities
- Three key useful algorithms:
  - Viterbi, Forward, EM
- Examples

Next Week: Philosophy & Ethics of Artificial Intelligence

- AI in the news
- Impacts on Society


Support The Guardian
Available for everyone, funded by readers
Contribute → Subscribe →
Search jobs | Sign in | Search | The Guardian | UK edition
News | Opinion | Sport | Culture | Lifestyle | More
UK World Business Coronavirus Football Environment UK politics Education Society Science Tech Global development Obituaries

Uber

Uber ditches effort to develop own self-driving car

Ride-hailing company sells division to startup Aurora but says they will collaborate


BBC | Sign in | News | Sport | Weather
NEWS TECHNOLOGY
Home | World | UK | England | N. Ireland | Scotland | Wales | Business | Politics

2 December 2014 Last updated at 13:02

Stephen Hawking warns artificial intelligence could end mankind

COMMENTS (1027)

By Rory Cellan-Jones
Technology correspondent

Stephen Hawking: "Humans, who are limited by slow biological evolution, couldn't compete and would be superseded"

Prof Stephen Hawking, one of Britain's pre-eminent scientists, has said that efforts to create thinking machines pose a threat to our very existence.

Vivas Today and Tomorrow

- Deadline was TODAY
- If you attempt today (having submitted already) but do not get vivad in time we will hold an emergency session next week
- ONLY VIVAS FOR PEOPLE WHO HAVE COMPLETED ALL 3 LABS
- Preference: LAB3 > LAB2&3 > LAB1&2&3