Markov Chains

Toolbox

- Search: uninformed/heuristic
- Adversarial search
- Probability
- Bayes nets
  - Naive Bayes classifiers
- Statistical inference

Reasoning over time

- In a Bayes net, each random variable (node) takes on one specific value.
  - Good for modeling static situations.
- What if we need to model a situation that is changing over time?

Example: Comcast

- In 2004 and 2007, Comcast had the worst customer satisfaction rating of any company or gov't agency, including the IRS.
- I have cable internet service from Comcast, and sometimes my router goes down. If the router is online, it will be online the next day with prob=0.8. If it's offline, it will be offline the next day with prob=0.4.
- How do we model the probability that my router will be online/offline tomorrow? In 2 days?

Example: Waiting in line

- You go to the Apple Store to buy the latest iPhone. Every minute, the first person in line is served with prob=0.5.
- Every minute, a new person joins the line with probability

    1 if the line length=0

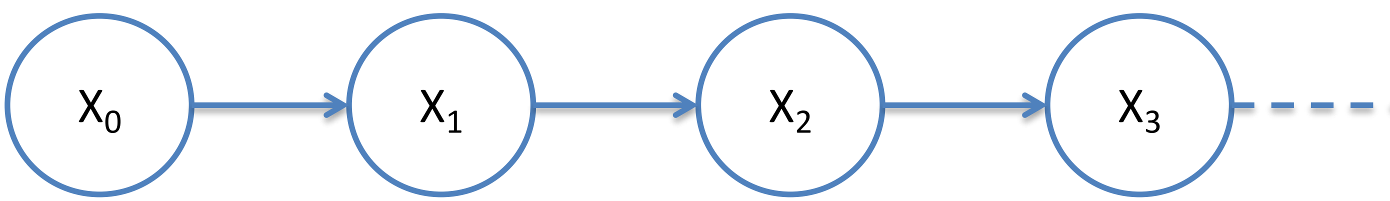
    2/3 if the line length=1

    1/3 if the line length=2

    0 if the line length=3
- How do we model what the line will look like in 1 minute? In 5 minutes?

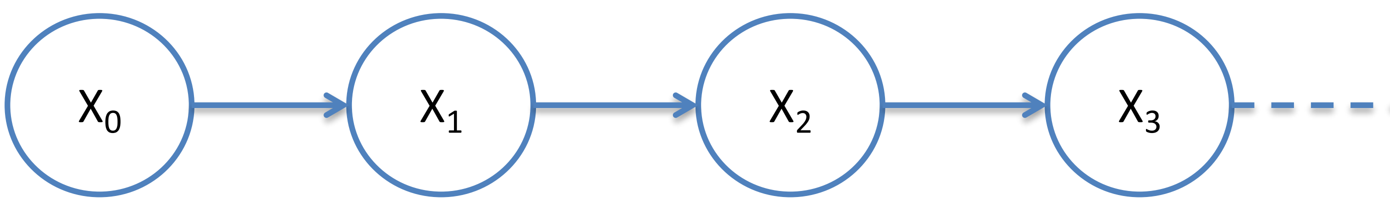
Markov Chains

- A Markov chain is a type of Bayes net with a potentially infinite number of variables (nodes).
- Each variable describes the state of the system at a given point in time (t).



Markov Chains

- Markov property:
  $P(X_t \mid X_{t-1}, X_{t-2}, X_{t-3}, \ldots) = P(X_t \mid X_{t-1})$
- Probabilities for each variable are identical:
  $P(X_t \mid X_{t-1}) = P(X_1 \mid X_0)$



Markov Chains

- Since these are just Bayes nets, we can use standard Bayes net ideas.
  - Shortcut notation: $X_{i:j}$ will refer to all variables X_i through X_j, inclusive.
- Common questions:
  - What is the probability of a specific event happening in the future?
  - What is the probability of a specific sequence of events happening in the future?

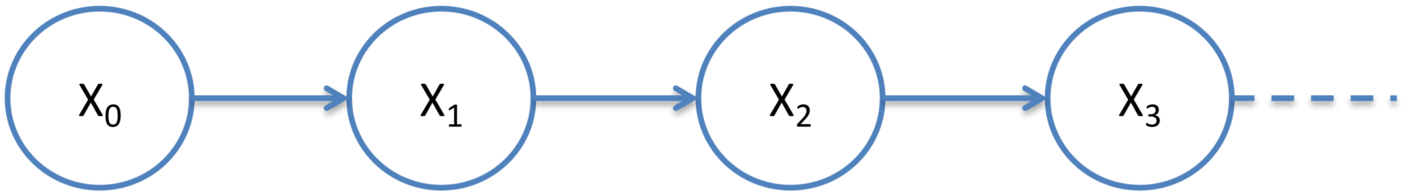
An alternate formulation

- We have a set of states, S.
- The Markov chain is always in *exactly one* state at any given time t.
- The chain transitions to a new state at each time t+1 based only on the current state at time t.

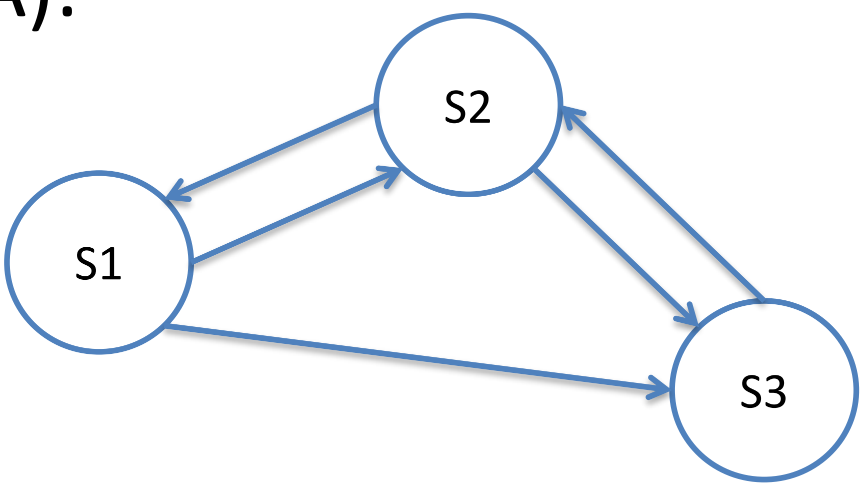
  $p_{ij} = P(X_{t+1} = j \mid X_t = i)$
- Chain must specify p_{ij} for all i and j, and starting probabilities for $P(X_0 = j)$ for all j.

Two different representations

- As a Bayes net:



- As a state transition diagram (similar to a DFA/NFA):



Formulate Comcast in both ways

- I have cable internet service from Comcast, and sometimes my router goes down. If the router is online, it will be online the next day with prob=0.8. If it's offline, it will be offline the next day with prob=0.4.
- Let's draw this situation in both ways.
- Assume on day 0, probability of router being down is 0.5.

Comcast

- What is the probability my router is offline for 3 days in a row (days 0, 1, and 2)?
  - $P(X_2=\text{off}, X_1=\text{off}, X_0=\text{off})$?
  - $P(X_2=\text{off} | X_0=\text{off}, X_1=\text{off}) * P(X_0=\text{off}, X_1=\text{off})$ *[mult rule]*
  - $P(X_2=\text{off} | X_0=\text{off}, X_1=\text{off}) * P(X_1=\text{off} | X_0=\text{off}) * P(X_0=\text{off})$
  - $P(X_2=\text{off} | X_1=\text{off}) * P(X_1=\text{off} | X_0=\text{off}) * P(X_0=\text{off})$
  - $p_{\text{off,off}} * p_{\text{off,off}} * P(X_0=\text{off})$

$$P(x_{0:t}) = P(x_0) \prod_{i=1}^{t} P(x_i \mid x_{i-1})$$

More Comcast

- Suppose I don't know if my router is online right now (day 0). What is the prob it is offline tomorrow?
  - $P(X_1=\text{off})$
  - $P(X_1=\text{off}) = P(X_1=\text{off}, X_0=\text{on}) + P(X_1=\text{off}, X_0=\text{off})$
  - $P(X_1=\text{off}) = P(X_1=\text{off}|X_0=\text{on}) * P(X_0=\text{on})$

    $+ P(X_1=\text{off}|X_0=\text{off}) * P(X_0=\text{off})$

$$P(X_{t+1}) = \sum_{x_t} P(X_{t+1} \mid x_t) P(x_t)$$

More Comcast

- Suppose I don't know if my router is online right now (day 0). What is the prob it is offline **the day after tomorrow**?
  - $P(X_2\text{=off})$
  - $P(X_2\text{=off}) = P(X_2\text{=off}, X_1\text{=on}) + P(X_2\text{=off}, X_1\text{=off})$
  - $P(X_2\text{=off}) = P(X_2\text{=off} | X_1\text{=on}) * P(X_1\text{=on})$

    $+ P(X_2\text{=off} | X_1\text{=off}) * P(X_1\text{=off})$

$$P(X_{t+1}) = \sum_{x_t} P(X_{t+1} \mid x_t) P(x_t)$$

Markov chains with matrices

- Define a transition matrix for the chain:

$$T = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}$$

- Each row of the matrix represents the transition probabilities **leaving** a state.
- Let v_t = a row vector representing the probability that the chain is in each state at time t.
- $v_t = v_{t-1} * T$

Formulate this matrix

- If the stock market is up one day, then it will be up the next day with prob=0.7.
- If it's down one day, it will be down the next day with prob=0.4.

Mini-forward algorithm

- Suppose we are given the value of X_t or a probability distribution over X_t and we want to predict X_{t+1}, X_{t+2}, X_{t+3}…
- Make row vector $v_t = P(X_t)$
  - Note that v_t can be something like [1, 0] if you know the true value of X_t, or it can be a distribution over values.
- $v_{t+1} = v_t * T$
- $v_{t+2} = v_{t+1} * T = v_t * T * T = v_t * T^2$
- $v_{t+3} = v_t * T^3$
- $v_{t+n} = v_t * T^n$

Back to the Apple Store...

- You go to the Apple Store to buy the latest iPhone.
- Every minute, a new person joins the line with probability

  1 if the line length=0

  2/3 if the line length=1

  1/3 if the line length=2

  0 if the line length=3
- Immediately after (in the same minute), the first person is helped with prob = 0.5
- Model this as a Markov chain, assuming the line starts empty. Draw the state transition diagram.
- What is T? What is v_0?