

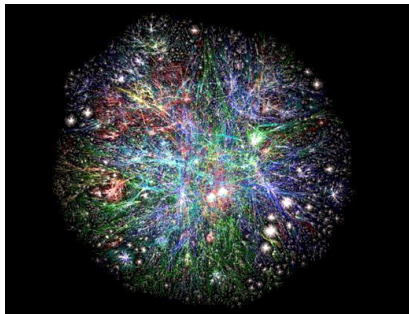
CMPSCI 240: Reasoning about Uncertainty

Lecture 14: Markov Chains

Andrew McGregor

University of Massachusetts

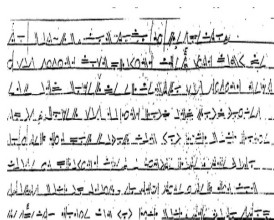
The Web Graph and PageRank



- **Web Graph:** Nodes are pages. Edges are hyperlinks between pages.
- How does Google pick which webpages to display when you search?
 - **First:** Find all webpages that contain your search terms. But there could be thousands, which should be displayed?
 - **Second:** Consider a user who randomly click links. Assume the “importance” of a page is proportional to the probability of ending up at that page. Amongst the pages that matched the search, return those with the highest probability.

Cryptography

Stanford's Statistics Dept. has a drop-in consulting service. A prison psychologist showed up with a collection of coded messages...



The approach they took to crack it:

- 1 Assume symbol corresponds to a letter. But which new symbol corresponds to each letter? There are $26!$ possible mappings.
- 2 They collected statistics from novels on how often each letter followed each other letter and then used this information to narrow down the options.

Markov Chain

In many examples, the future depends on the past only through present!

Example

The web page that visited at time t only depends on the web page visited at time $t - 1$.

Example

Whether the next decoded symbol is “B” only depends on the decoded symbol immediately preceding it.

Example

Whether you understand the content of the next class only depends on whether you understand the concept in today's class.

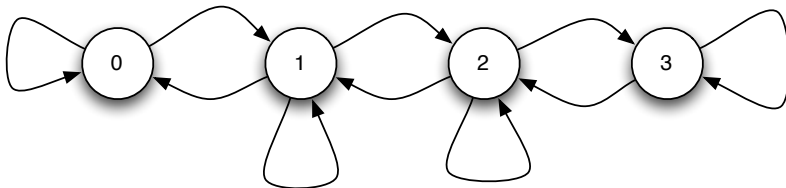
Analyzing the Queue at Earth Foods Cafe

- Consider a queue at Earth Foods Cafe
- Every minute, someone joins the queue. . .
 - With probability 1 if the queue has length 0
 - With probability $2/3$ if the queue has length 1
 - With probability $1/3$ if the queue has length 2
 - With probability 0 if the queue has length 3.
- Every minute, the server serves a customer with probability $1/2$.

Suppose 1 person in line at noon. How many people in line at 12:10pm?

States

- At any given time, the queue is in one of four *states*: either there are 0,1,2, or 3 people in the queue.
- Arrows indicate that it is possible to move from one state to the next at each step.



Transition Probabilities

- Let X_t be the number of people in the queue at time t
- Suppose the queue has 1 person at time t ,

$$P(X_{t+1} = 0 | X_t = 1) = P(\text{no-one arrived \& someone served} | X_t = 1) = 1/6$$

$$P(X_{t+1} = 2 | X_t = 1) = P(\text{someone arrived \& no-one served} | X_t = 1) = 1/3$$

$$P(X_{t+1} = 3 | X_t = 1) = 0$$

$$\begin{aligned} P(X_{t+1} = 1 | X_t = 1) &= P(\text{no-one arrived and no-one served or} \\ &\quad \text{someone arrived and someone served} | X_t = 1) \\ &= 1/3 \cdot 1/2 + 2/3 \cdot 1/2 = 1/2 \end{aligned}$$

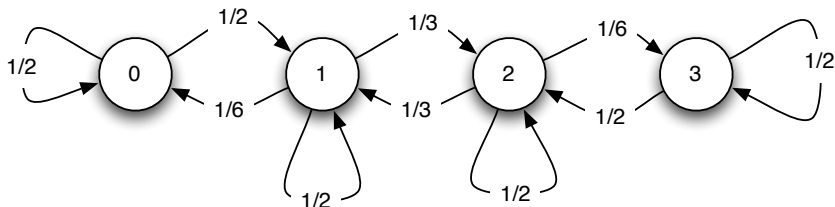
- Suppose the queue has 0 people at time t ,

$$P(X_{t+1} = 0 | X_t = 0) = 1/2, \quad P(X_{t+1} = 1 | X_t = 0) = 1/2,$$

$$P(X_{t+1} = 2 | X_t = 0) = 0, \quad P(X_{t+1} = 3 | X_t = 0) = 0$$

States with Transition Probabilities

- Weight p_{ij} on arrow from state i to state j indicates the probability of transitioning to state j given we're in state i .



- **N.B.** Weights on arrows out of each state i sum to one: $\sum_j p_{ij} = 1$
- **Question:** If we're in state 2, what's the probability we're in state 3 after one step: A) 1, B) 1/2, C) 1/6, D) 0, E) 1/3. Ans: 1/6.
- **Question:** If we're in state 2, what's the probability we're in state 2 after two steps: A) 1/3, B) 4/9, C) 1/4, D) 1/12, E) 1/9. Ans: 4/9.

What if the current state is uncertain?

- What if don't know X_{t-1} , but know $P(X_{t-1} = i)$ for each i ?
- Then, by the Law of Total Probability:

$$P(X_t = j) = \sum_i P(X_t = j | X_{t-1} = i) P(X_{t-1} = i)$$

- **Question:** If there's a $1/3$ probability we're in state 1 and a $2/3$ probability we're in state 3, what's the probability we're in state 2 after one step. A) $1/3$, B) $1/4$, C) $4/9$, D) $7/9$, E) $1/9$. Ans: $4/9$.

Analyzing Markov Chains via Matrices

- Define **transition probability matrix**:

$$A = \begin{pmatrix} p_{0,0} & p_{0,1} & p_{0,2} & p_{0,3} \\ p_{1,0} & p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,0} & p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,0} & p_{3,1} & p_{3,2} & p_{3,3} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/6 & 1/2 & 1/3 & 0 \\ 0 & 1/3 & 1/2 & 1/6 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}$$

where $p_{i,j} = P(X_t = j | X_{t-1} = i)$

- Define $v_t = \langle P(X_t = 0), P(X_t = 1), P(X_t = 2), P(X_t = 3) \rangle$
- By the Law of Total Probability,

$$P(X_1 = j) = \sum_i P(X_1 = j | X_0 = i) P(X_0 = i) = v_0 \begin{pmatrix} p_{0,j} \\ p_{1,j} \\ p_{2,j} \\ p_{3,j} \end{pmatrix}$$

- Markov Chain Thm:** Given v_0 , we can compute $v_1 = v_0 A$, and

$$v_t = v_{t-1} A = v_{t-2} A A = v_{t-3} A A A = \dots = v_0 A^t$$

Simulation of the queue if there is initially one person

$$\begin{aligned}v_0 &= \langle 0.000, 1.000, 0.000, 0.000 \rangle \\v_1 &= \langle 0.167, 0.500, 0.333, 0.000 \rangle \\v_2 &= \langle 0.167, 0.444, 0.333, 0.056 \rangle \\v_3 &= \langle 0.158, 0.416, 0.342, 0.084 \rangle \\v_4 &= \langle 0.148, 0.401, 0.352, 0.099 \rangle \\v_5 &= \langle 0.142, 0.391, 0.359, 0.109 \rangle \\v_6 &= \langle 0.136, 0.386, 0.364, 0.114 \rangle \\v_7 &= \langle 0.133, 0.382, 0.368, 0.118 \rangle \\v_8 &= \langle 0.130, 0.380, 0.370, 0.120 \rangle \\&\vdots \quad \vdots \quad \vdots \\v_\infty &= \langle 0.125, 0.375, 0.375, 0.125 \rangle\end{aligned}$$

Simulation of the queue if there is initially three people

$$\begin{aligned}v_0 &= \langle 0.000, 0.000, 0.000, 1.000 \rangle \\v_1 &= \langle 0.000, 0.000, 0.500, 0.500 \rangle \\v_2 &= \langle 0.000, 0.167, 0.500, 0.333 \rangle \\v_3 &= \langle 0.028, 0.250, 0.472, 0.251 \rangle \\v_4 &= \langle 0.056, 0.296, 0.404, 0.204 \rangle \\v_5 &= \langle 0.078, 0.324, 0.423, 0.177 \rangle \\v_6 &= \langle 0.093, 0.341, 0.407, 0.159 \rangle \\v_7 &= \langle 0.104, 0.353, 0.397, 0.148 \rangle \\v_8 &= \langle 0.111, 0.360, 0.389, 0.140 \rangle \\&\vdots \quad \vdots \quad \vdots \\v_\infty &= \langle 0.125, 0.375, 0.375, 0.125 \rangle\end{aligned}$$