Markov Chains

Prof. S. Shakya

Markov Chains

- If the future states of a process are independent of the past and depend only on the present , the process is called a Markov process
- A discrete state Markov process is called a Markov chain.
- A Markov Chain is a random process with the property that the next state depends only on the current state.

Markov Chains

- Since the system changes randomly , it is generally impossible to predict the exact state of the system in the future.
- However, the statistical properties of the system's future can be predicted.
- In many applications it is these statistical properties that are important current state depends only the current state.
- M/M/m queues can be modeled using Markov processes.
- The time spent by the job in such a queue is Markov process and the number of jobs in the queue is a Markov chain.

Markov Chain

- A simple example is the nonreturning random walk, where the walkers are restricted to not go back to the location just previously visited.

Markov Chains

- Markov chains is a mathematical tools for statistical modeling in modern applied mathematics, information science

Why Study Markov Chains?

- Markov chains are used to analyze trends and predict the future. (Weather, stock market, genetics, product success, etc.

Markov Chains

As we have discussed, we can view a stochastic process as sequence of random variables

$$\{X_1, X_2, X_3, X_4, X_5, X_6, X_7, \dots\}$$

Suppose that X_7 depends only on X_6, X_6 depends only on X_5, X_5 on X_4, and so forth. In general, if for all i,j,n,

$$P(X_{n+1} = j | x_n = i_n, x_{n-1} = i_{n-1}, \dots, x_0 = i_0) = P(X_{n+1} = j | X_n = i_n),$$

then this process is what we call a Markov chain.

Markov Chains

- The conditional probability above gives us the probability that a process in state i_n at time n moves to i_{n+1} at time n + 1.
- We call this the transition probability for the Markov chain.
- If the transition probability does not depend on the time n, we have a stationary Markov chain, with transition probabilities

$$P_{ij} = P(X_{n+1} = j | X_n = i)$$

Now we can write down the whole Markov chain as a matrix P:

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \vdots & \vdots & \ddots & \\ P_{n1} & P_{n2} & & P_{nn} \end{bmatrix}$$

Markov Chains

- The probability of going from state *i* to state *j* in *n* time steps is

$$p_{ij}^{(n)} = \Pr(X_n = j \mid X_0 = i)$$

and the single-step transition is

$$p_{ij} = \Pr(X_1 = j \mid X_0 = i).$$

- For a time-homogeneous Markov chain:

$$p_{ij}^{(n)} = \Pr(X_{n+k} = j \mid X_k = i)$$

and

$$p_{ij} = \Pr(X_{k+1} = j \mid X_k = i).$$

Key Features of Markov Chains

- A sequence of trials of an experiment is a **Markov chain** if

1) the outcome of each experiment is one of a set of discrete states;
2) the outcome of an experiment depends only on the present state, and not on any past states;
3) the transition probabilities remain constant from one transition to the next.

Markov Chains

- The Markov chain has network structure much like that of website, where each node in the network is called a state and to each link in the network a transition probability is attached, which denotes the probability of moving from the source state of the link to its destination state.

Markov Chains

- The process attached to a Markov chain moves through the states of the networks in steps, where if a any time the system is in state i, then with probability equal to the transition probability from state i, to state j, it moves to state j.
- We will model the transitions from one page to another in a web site as a Markov chain.
- The assumption we will make , called Markov property, is that the probability of moving from source page to a destination page doesn't depend on the route taken to reach the source.

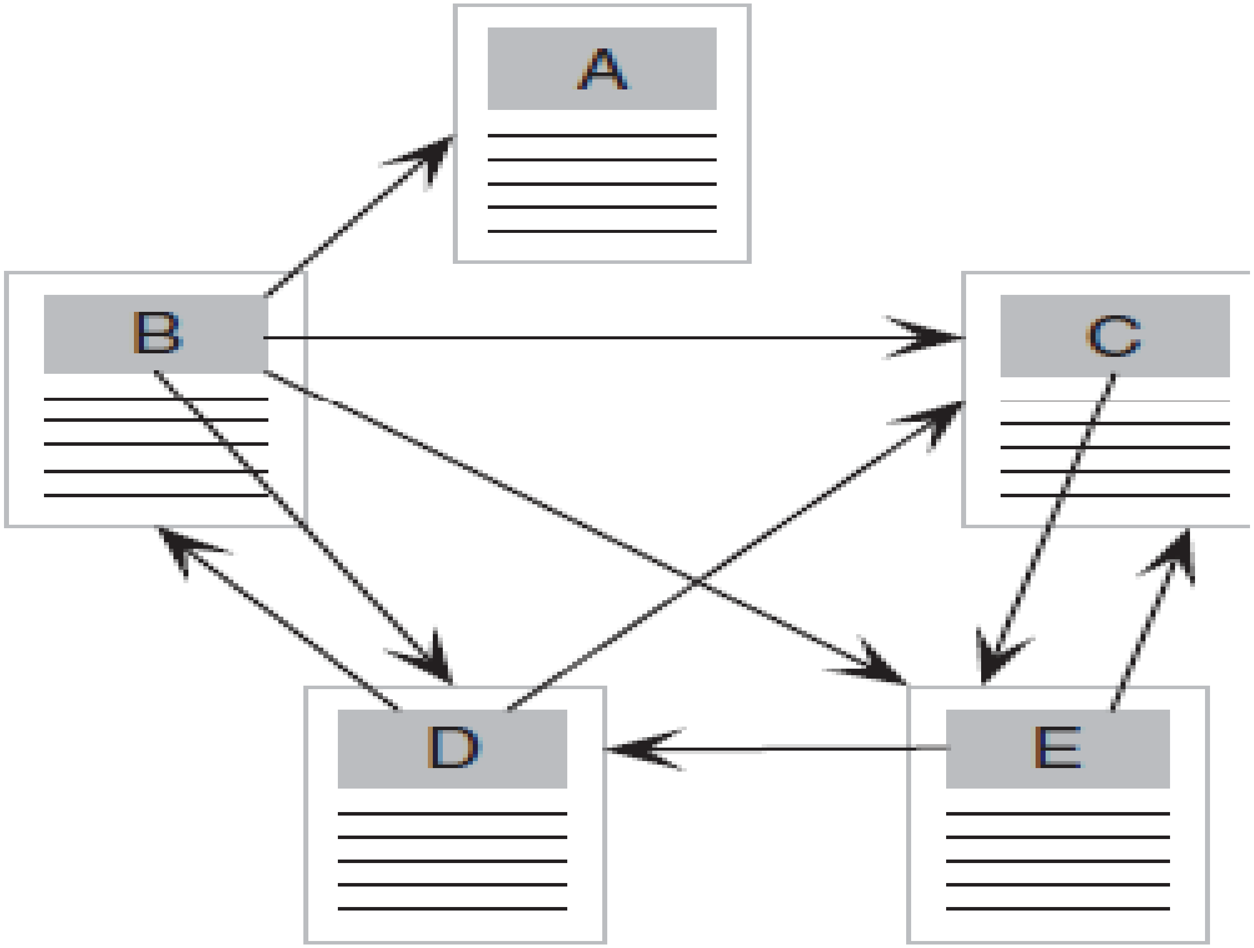
Internet application

- The PageRank of a webpage as used by Google is defined by a Markov chain.
- It is the probability to be at page *i* in the stationary distribution on the following Markov chain on all (known) webpages. If *N* is the number of known webpages, and a page *i* has k_i links then it has transition probability $\frac{\alpha}{k_i} + \frac{1-\alpha}{N}$ for all pages that are linked to and $\frac{1-\alpha}{N}$ for all pages that are not linked to.
- The parameter α is taken to be about 0.85

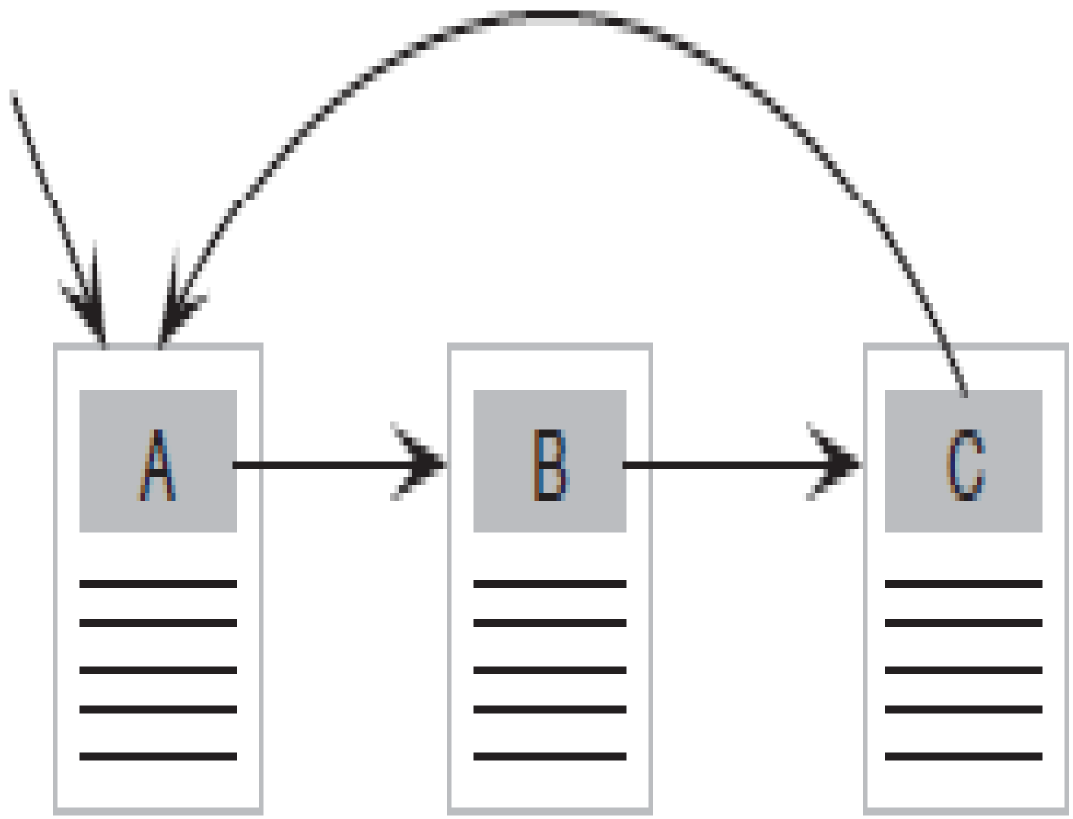
Internet application

- Markov models have also been used to analyze web navigation behavior of users.
- A user's web link transition on a particular website can be modeled using first- or second-order
- Markov models and can be used to make predictions regarding future navigation and to personalize the web page for an individual user.

Example website to illustrate PageRank



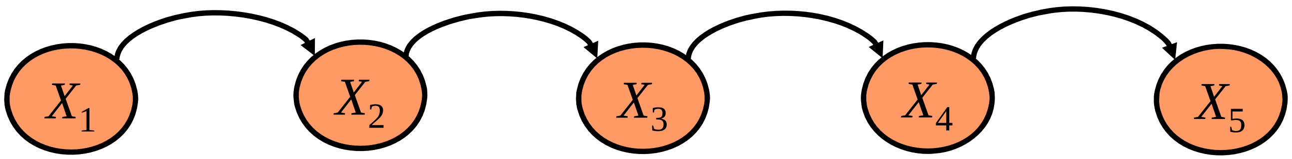
Example of a rank sink



Markov Process

- **Markov Property:** The state of the system at time $t+1$ depends only on the state of the system at time t

$$\boxed{\Pr\left[X_{t+1} = x_{t+1} \mid X_1 \cdots X_t = x_1 \cdots x_t\right] = \Pr\left[X_{t+1} = x_{t+1} \mid X_t = x_t\right]}$$



- **Stationary Assumption:** Transition probabilities are independent of time (t)

$$\boxed{\Pr\left[X_{t+1} = b \mid X_t = a\right] = p_{ab}}$$

Markov Process

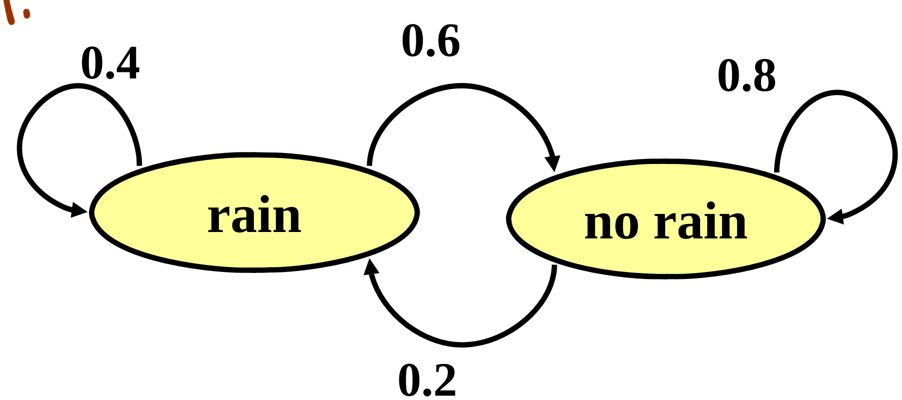
Simple Example

Weather:

- raining today ⇒ 40% rain tomorrow
  ⇒ 60% no rain tomorrow

- not raining today ⇒ 20% rain tomorrow
  ⇒ 80% no rain tomorrow

Stochastic FSM:



Markov Process

Simple Example

Weather:

- **raining today** ⇒ **40% rain tomorrow**
  ⇒ **60% no rain tomorrow**

- **not raining today** ⇒ **20% rain tomorrow**
  ⇒ **80% no rain tomorrow**

The transition matrix:

$$P = \begin{pmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{pmatrix}$$

- Stochastic matrix:
  Rows sum up to 1
- Double stochastic matrix:
  Rows and columns sum up to 1

Markov Process

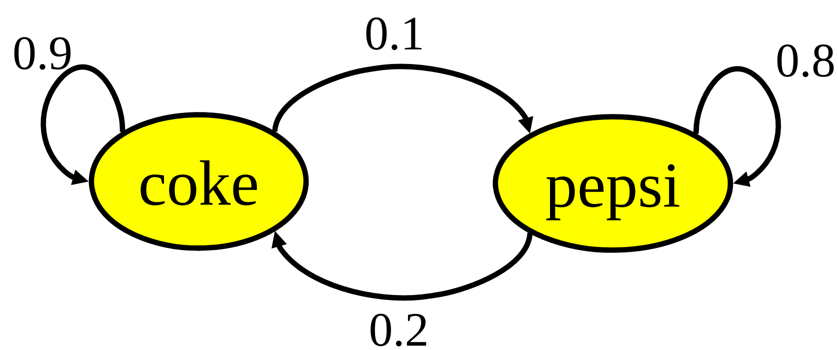
Coke vs. Pepsi Example

- Given that a person's last cola purchase was Coke, there is a 90% chance that his next cola purchase will also be Coke.

- If a person's last cola purchase was Pepsi, there is an 80% chance that his next cola purchase will also be Pepsi.

transition matrix:

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$



Markov Process

Coke vs. Pepsi Example (cont)

Given that a person is currently a Pepsi purchaser, what is the probability that he will purchase Coke two purchases from now?

Pr[Pepsi→?→Coke] =

Pr[Pepsi→Coke→Coke] + Pr[Pepsi→ Pepsi →Coke] =

0.2 * 0.9 + 0.8 * 0.2 = 0.34

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

Pepsi → ?   ? → Coke

Markov Process

Coke vs. Pepsi Example (cont)

Given that a person is currently a Coke purchaser, what is the probability that he will purchase Pepsi **three** purchases from now?

$$P^3 = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix} = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

Markov Process

Coke vs. Pepsi Example (cont)

- Assume each person makes one cola purchase per week
- Suppose 60% of all people now drink Coke, and 40% drink Pepsi
- What fraction of people will be drinking Coke three weeks from now?

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \qquad P^3 = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

$$\Pr[X_3 = \text{Coke}] = 0.6 * 0.781 + 0.4 * 0.438 = 0.6438$$

Q_i - the distribution in week i

$Q_0=(0.6,0.4)$ - initial distribution

$Q_3 = Q_0 * P^3 = (0.6438, 0.3562)$

Markov Process

Coke vs. Pepsi Example (cont)

Simulation: