Chapter 5
Markov Chains

Markov Process

- Markov process is a process whose future probabilities are determined by its most recent values.
- 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.
- Markov process is a simple **stochastic process** in which the distribution of future states depends only on the present state and not on how it arrived in the present state.
- Markov process models are useful in studying the evolution of systems over repeated trails or sequence time periods or stages.
- Examples of Markov Process:

1. Airplane at Airport
2. Rainfall
3. Behavior of Business or Economy
4. Flow of traffic

Markov Chain

- A Markov chain is a stochastic model describing a sequence of possible events in which the probability of next state depends only on the previous event.
- A discrete state Markov process is called a Markov chain.
- A Markov chain is a probabilistic model describing a system that changes from state to state, and in which the probability of the system being in a certain state at a certain time step depends only on the state of the preceding time step
- 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..
- Markov chains is a mathematical tools for statistical modeling in modern applied

mathematics, information science.

- 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 chains are used to compute the probabilities of events occurring by viewing them as states transitioning into other states, or transitioning into the same state as before.
- Markov chains are used to analyze trends and predict the future such as weather forecasting, stock market prediction, genetics, product success, etc.
- A Markov chain consists of states and transition probabilities. Each transition probability is the probability of moving from one state to another in one step.
- The probability that j is the next event of the chain given that the current state is i is called the transition probability from i to j.
- These transition probabilities are independent of the past and depend only on the two states involved.

- 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.
- If at 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 make an assumption, called Markov property, according to which the probability of moving from source page to a destination page doesn’t depend on the route taken to reach.

- If the transition probability does not depend on the time n, we have a stationary Markov chain, with transition probabilities

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

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

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

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

Transition Matrix

- The matrix of transition probabilities is called the transition matrix.
- It is a square matrix and is represented by P.
- The transition probability matrix is the matrix that shows probabilities of moving from one state to another state.
- So the transition matrix for whole Markov chain can be represented as:

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

where P_{11}, P_{12},….P_{nn} are the transition probabilities.

- All the entries of the matrix lie between 0 and 1. The sum of entries of any row is equal to 1.

$$
\begin{array}{c} \\ \text{Nice} \\ \text{Rainy} \\ \text{Snowy} \end{array}
\begin{array}{c}
\begin{array}{ccc} \text{Nice} & \text{Rainy} & \text{Snowy} \end{array} \\
\begin{bmatrix} 0.0 & 0.75 & 0.25 \\ 0.25 & 0.25 & 0.5 \\ 0.25 & 0.5 & 0.25 \end{bmatrix}
\end{array}
$$

Key Features of Markov Chain

- 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 probability remain constant from one transition to the next.

Markov Process

- Markov process is a process whose future probabilities are determined by its most recent values.
- According to **Markov Property**, the state of the system at time t+1 depends only on the state of the system at time t.
- Also a stationary assumption is made according to which transition probabilities are independent of time(t). So,

$$Pr[X_{t+1} = b \mid X_t = a] = P_{ab}$$

Current Status Distribution Matrix

- Current status distribution matrix is a row matrix that provides the status of current state of all the discrete states of the Markov chain.
- It is denoted by Q_0.
- Each entry must be between 0 and 1 inclusive.
- The sum of entries of each row must be 1.

Scenario

Given that chance of a Honda bike user to buy Honda bike at next purchase is 70% and that his next purchase will be Yamaha is 30%. The chance of Yamaha bike user to buy Yamaha bike at next purchase is 80% and that his next purchase will be Honda is 20%. What is the probability to buy Yamaha bike after three purchase of a current Honda Bike user?

$$\text{Current Status Matrix } Q_0 = \begin{array}{c} \begin{array}{cc} \text{Honda User} & \text{Yamaha User} \end{array} \\ \begin{bmatrix} 1 & \qquad\qquad 0 \end{bmatrix} \end{array}$$

Steps on making predictions

1. Create current status distribution matrix Q_o.
2. Create probability distribution matrix or transition matrix P.
3. Calculate $Q_n = Q_o * P^n$, which represents probability vector after n repetitions of the experiment.

Application of Markov Chain

1. Internet Application

Link Analysis

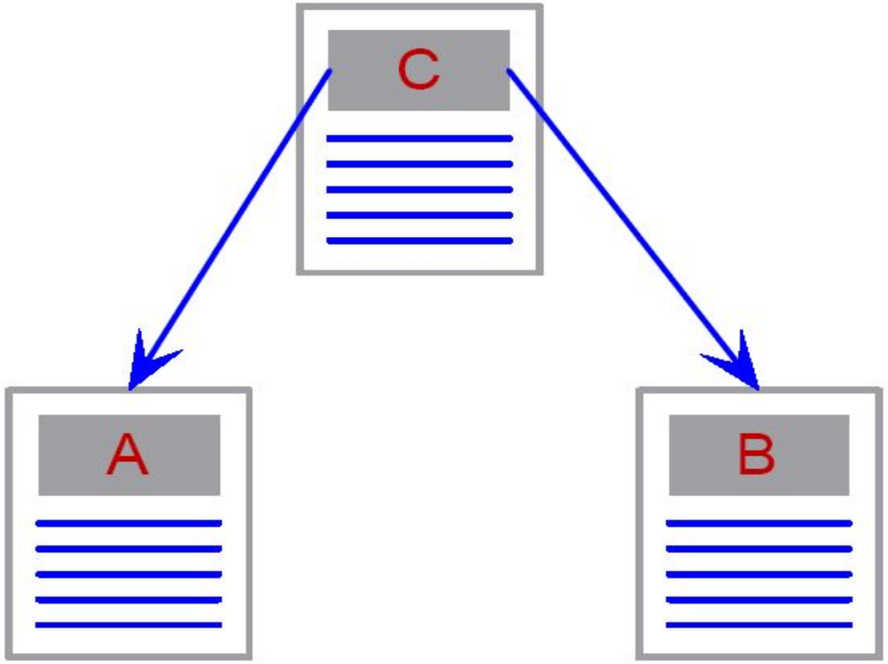
- Link analysis is a data-analysis technique used to evaluate relationships (connections) between nodes.
- Relationships may be identified among various types of nodes (objects), including organizations, people and transactions.
- A link from page A to page B is a vote of the author of A for B, or a recommendation of the page.
- The number of incoming links to a page is a measure of importance and authority of the page.
- A page is more important if the sources of its incoming links are important.

Why link analysis?

- The web is not just a collection of documents – its hyperlinks are important!
- A link from page A to page B may indicate:
  - A is related to B, or
  - A is recommending, citing, voting for or endorsing B
- Links are either
  - referential – click here and get back home, or
  - Informational – click here to get more detail
- Links effect the ranking of web pages and thus have commercial value.
- Link analysis has been used for:
  - investigation of criminal activity (fraud detection, counterterrorism, and intelligence)
  - computer security analysis
  - search engine optimization
  - market research
  - medical research

Citation Analysis – Impact Factor

- The **impact factor** of a journal = $\frac{A}{B}$
  - A is the number of current year citations to articles appearing in the journal during previous two years.
  - B is the number of articles published in the journal during previous two years.
- Co-citation: A and B are co-cited by C, implying that they are related or associated.
- The strength of co-citation between A and B is the number of times they are co-cited.



Page Rank

- The PageRank of a webpage as used by Google is defined by a Markov chain.
- Google's PageRank (PR) is method of ranking web pages for placement on a Search Engine Results Page (SERP).
- PageRank is a mathematical formula (algorithm) that Google uses to calculate the importance of a particular web page/URL based on incoming links.
- PageRank algorithm assigns each web page a relevancy score.
- It is used to measure the <u>relative importance of a website within it's set of hyperlinked pages.</u>
- If we rank better in organic search, then we should get more website traffic from search engines.

- Markov models can be used to make predictions regarding future navigation and to personalize the web page for an individual user.
- Markov models have also been used to analyze web navigation behavior of users.
- **Rank Sink:** A page or group of pages is a rank sink if they can receive rank propagation from their parent but can’t propagate rank to other pages. A rank sink occurs when a page does not link out.
- **Dangling pages**: Dangling pages are pages which do not have any out link or the page which not provide reference to other pages

The Page Rank Algorithm

 The original Page Rank algorithm was described by Lawrence Page and Sergey Brin in several publications. It is given by

PR(A) = (1-d) + d (PR(T1)/C(T1) + ... + PR(Tn)/C(Tn))

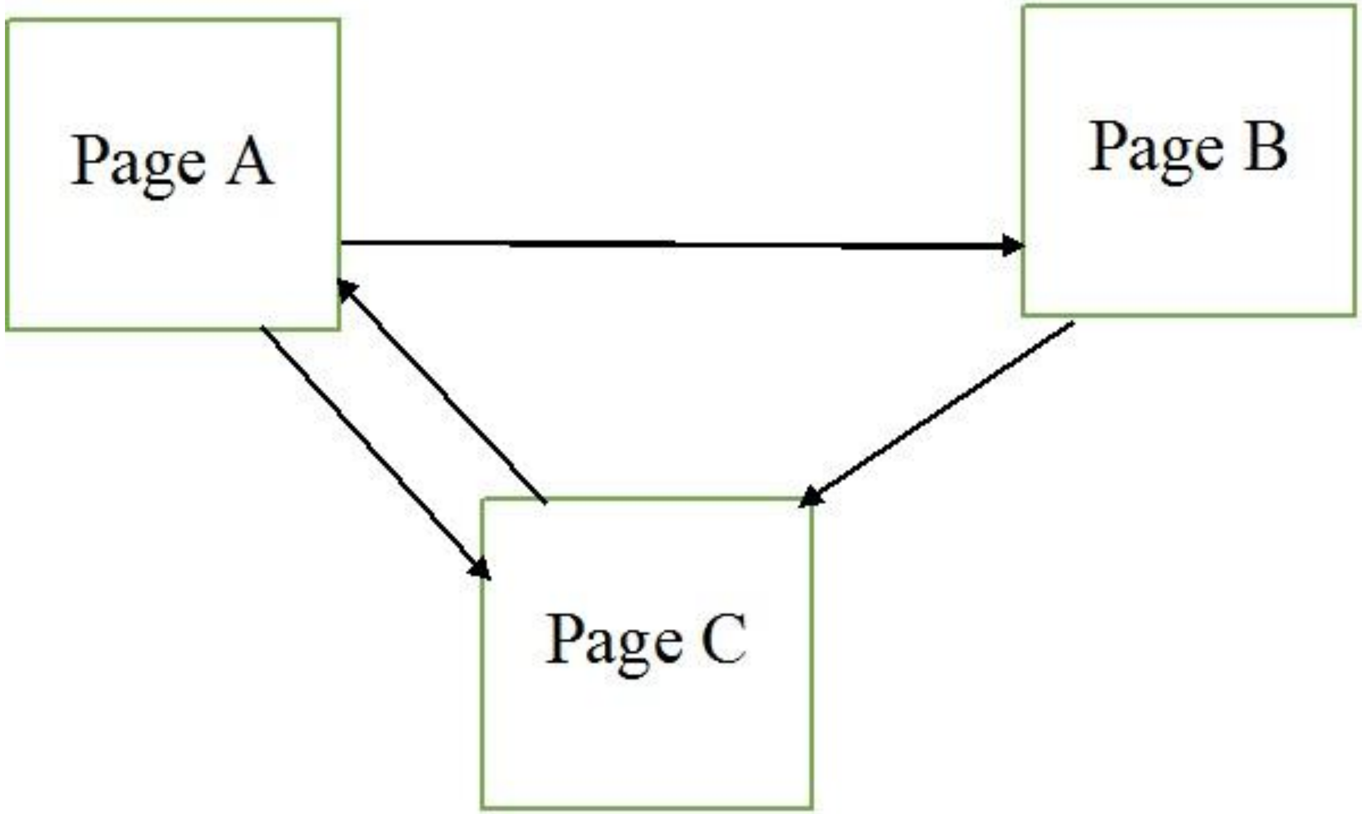
Where

- PR(A) is the Page Rank of page A,
- PR(Ti) is the Page Rank of pages Ti which link to page A,
- C(Ti) is the number of outbound links on page Ti and
- d is a damping factor which can be set between 0 and 1.

Page Rank Computation

We regard a small web consisting of three pages A, B and C, whereby page A links to the pages B and C, page B links to page C and page C links to page A. According to Page and Brin, the damping factor d is usually set to 0.85, but to keep the calculation simple we set it to 0.5. The exact value of the damping factor d admittedly has effects on Page Rank, but it does not influence the fundamental principles of Page Rank.

Initially let the page rank of each page be 1. Calculate iteratively and conclude which page has the highest score.


Page A
Page B
Page C

- Using formula for page rank

  $$PR(A) = 0.5 + 0.5\ PR(C)/1$$
  $$PR(B) = 0.5 + 0.5\ (PR(A) / 2)$$
  $$PR(C) = 0.5 + 0.5\ (PR(A) / 2 + PR(B))$$

- These equations can easily be solved. We get the following Page Rank values for the single pages:

PR(A) = 14/13 = 1.07692308
PR(B) = 10/13 = 0.76923077
PR(C) = 15/13 = 1.15384615

- It is obvious that the sum of all page's Page Ranks is 3 and thus equals the total number of web pages.

2. **Market Research and Market Trend Prediction:** Markov chains and their respective diagrams can be used to model the probabilities of certain financial market climates and thus predicting the likelihood of future market conditions.

3. **Asset pricing and other financial predictions:** Markov chain and Markov process can be used to predict the price and financial factors of certain assets.

4. **Markov text generator:** Can be used in automatic text generation. A Markov chain algorithm basically determines the next most probable suffix word for a given prefix.

5. **Population Genetics:** Markov chain models have been the most widely used ones in the study of random fluctuations in the genetic compositions of populations over generations.

Numericals

Spring time has 3 possible conditions nice, rainy and snowy. If its nice today then tomorrow it will be:

a. rainy 75% of the time
b. snowy 25% of the time

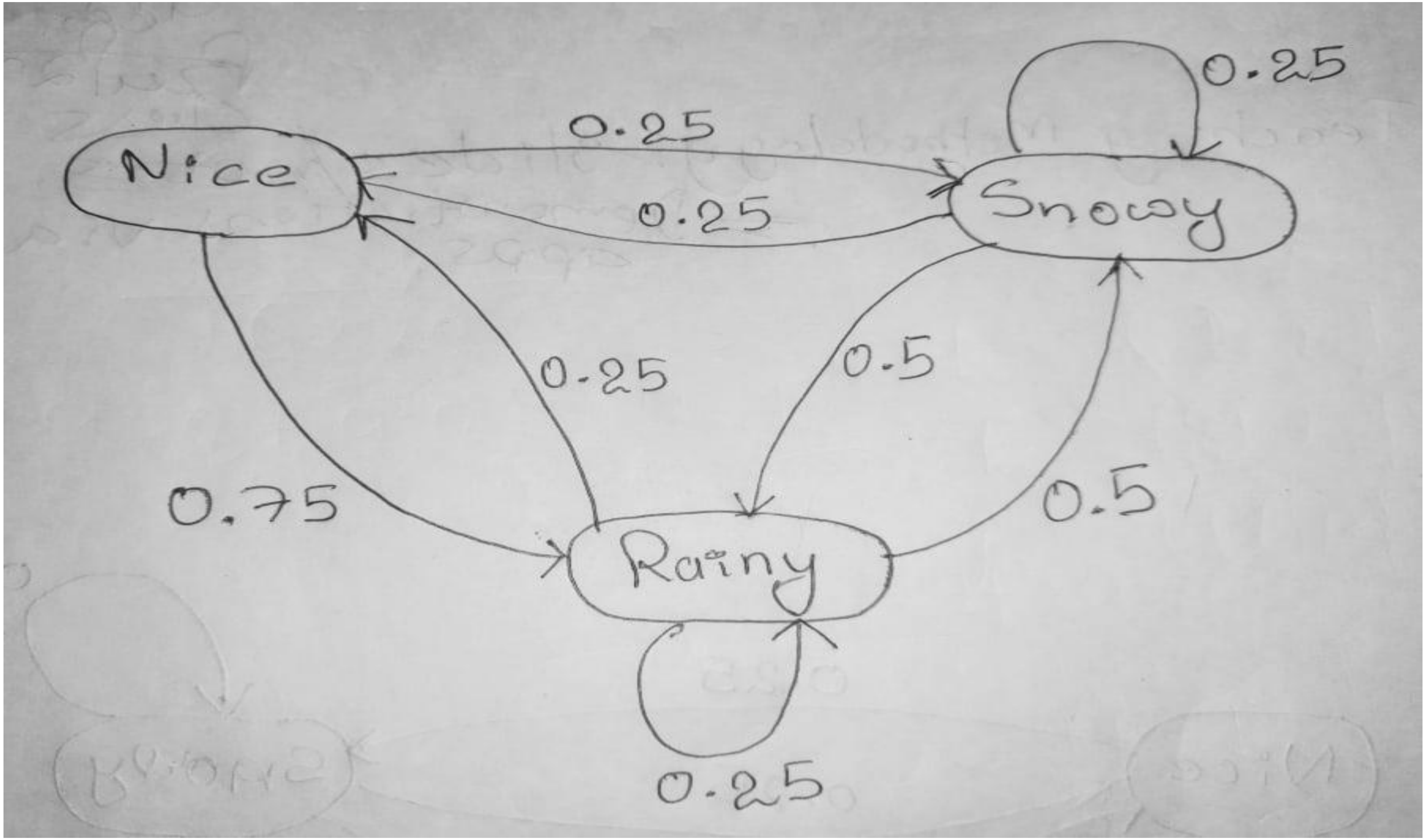
If its rainy today then tomorrow it will be:

a. rainy 25% of the time
b. nice 25% of the time
c. snowy 50%of the time

If its snowy today then tomorrow it will be:

a. rainy 50% of the time
b. nice 25% of the time
c. snowy 25% of the time

Make graph or stochastic FSM of above and construct Transition matrix.



Stochastic FSM

$$
\begin{array}{c|c}
 & \begin{array}{ccc} \text{Nice} & \text{Rainy} & \text{Snowy} \end{array} \\
\begin{array}{c} \text{Nice} \\ \text{Rainy} \\ \text{Snowy} \end{array} &
\begin{bmatrix} 0.0 & 0.75 & 0.25 \\ 0.25 & 0.25 & 0.5 \\ 0.25 & 0.5 & 0.25 \end{bmatrix}
\end{array}
$$

Transition Matrix

Numerical

Given that a person's last coca-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.

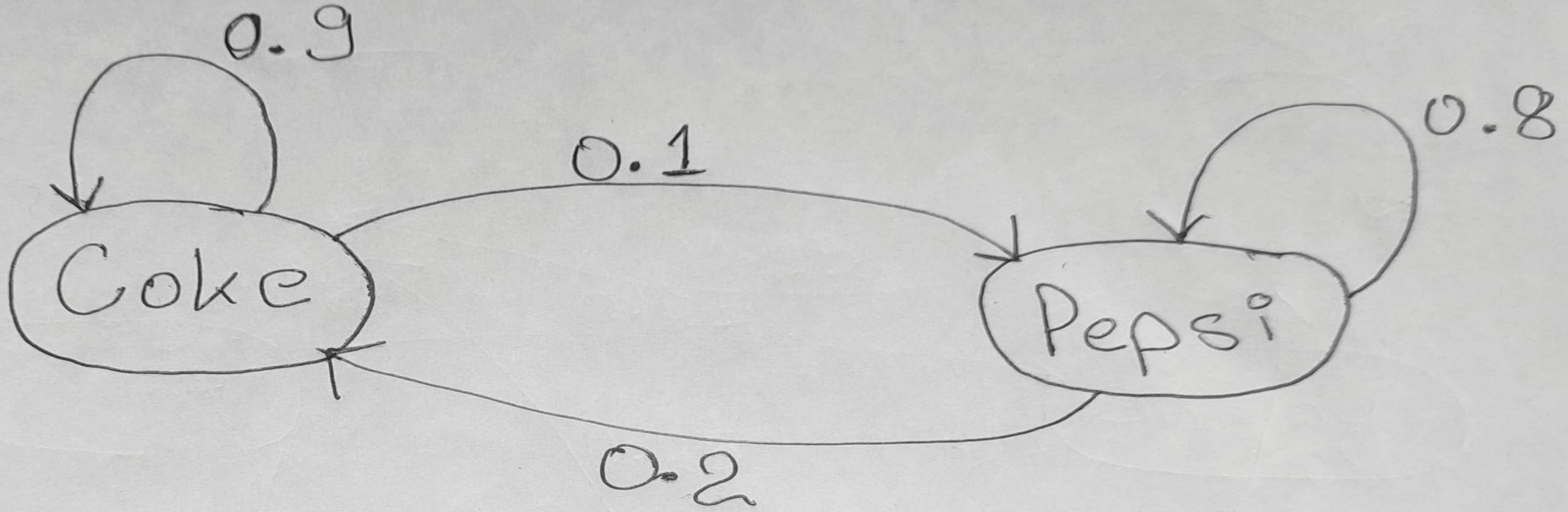
a. Make graph of above problem and construct a transition matrix.

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

c. Given that a person is currently a Coke purchaser, what is the probability that he will purchase pepsi three purchase from now?

d. 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?

a.



Transition Matrise

$$\begin{array}{c c} & \begin{array}{cc} \text{Coke} & \text{Pepsi} \end{array} \\ \begin{array}{c} \text{Coke} \\ \text{Pepsi} \end{array} & \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \end{array}$$

b.

Method 1

Pepsi → [ ] → Coke

⇓

Either Coke/Pepsi

Possibilities

Pepsi → Coke → Coke   Pepsi → Pepsi → Coke

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

= Pr[Pepsi → Coke and Coke → Coke] +

Pr[Pepsi → Pepsi and Pepsi → Coke]

$= 0.2 * 0.9 + 0.8 * 0.2$

$= 0.34$

b.

Method 2 (Using Current Status Distribution Matrix)

$$P = \begin{array}{c|cc} & Coke & Pepsi \\ \hline Coke & 0.9 & 0.1 \\ Pepsi & 0.2 & 0.8 \end{array}$$

Current Status Distribution Matrix (Q_n)

$$Q_0 = \begin{array}{cc} Coke & Pepsi \\ \end{array}$$
$$Q_0 = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

Since question is asking for two purchases from now,

$$\boxed{Q_2 = Q_0 * P^2}$$

$$= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}^2$$

$$= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

$$= \begin{bmatrix} 0.34 & 0.66 \end{bmatrix}$$

Coke Pepsi

∴ Probability is 0.34

c.

Currently Coke Purchaser

So, Current status distribution matrix $(Q_0) = \begin{array}{cc} \text{Coke} & \text{Pepsi} \\ \left[1 \right. & \left. 0 \right] \end{array}$

Probability he will buy (pepsi) three purchase from now $Q_3 = Q_0 * P^3$

$$= [1 \quad 0] \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}^3$$

$$= [1 \quad 0] \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

$$= [\underbrace{0.781}_{\text{Coke}} \quad \underbrace{0.219}_{\text{Pepsi}}]$$

$\therefore$ Probability he will buy Pepsi three purchases from now is 0.219.

d.

Current Status Distribution

$$\text{Matrix } (Q_0) = \begin{array}{cc} \text{Coke} & \text{Pepsi} \\ \left[\begin{matrix} 0.6 \end{matrix}\right. & \left.\begin{matrix} 0.4 \end{matrix}\right] \end{array}$$

$$Q_3 = Q_0 * P^3$$

$$= \begin{bmatrix} 0.6 & 0.4 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}^3$$

$$= [\underbrace{0.6438}_{\text{Coke}} \quad \underbrace{0.3562}_{\text{Pepsi}}]$$

$\therefore$ Probability $= 0.6438$

Numerical

Given that a chance of Ford car user to buy a Ford car in next purchase is 70% and that his next purchase is will be a Scorpio is 30% and chance of Scorpio car user to buy Scorpio car at the next purchase is 80% and chance that his next purchase will be Ford car is 20%. What is the probability to buy a Scorpio car after three purchase of a current Ford user? If 70% user use Ford car today, what percentage of user will use Scorpio after 3 purchase?

Numerical

Given that chance of a Honda bike user to buy Honda bike at next purchase is 70% and that his next purchase will be Yamaha is 30%. The chance of Yamaha bike user to buy Yamaha bike at next purchase is 80% and that his next purchase will be Honda is 20%. What is the probability to buy Yamaha bike after three purchase of a current Honda Bike user?

- Ans: 0.525