Introduction to AI

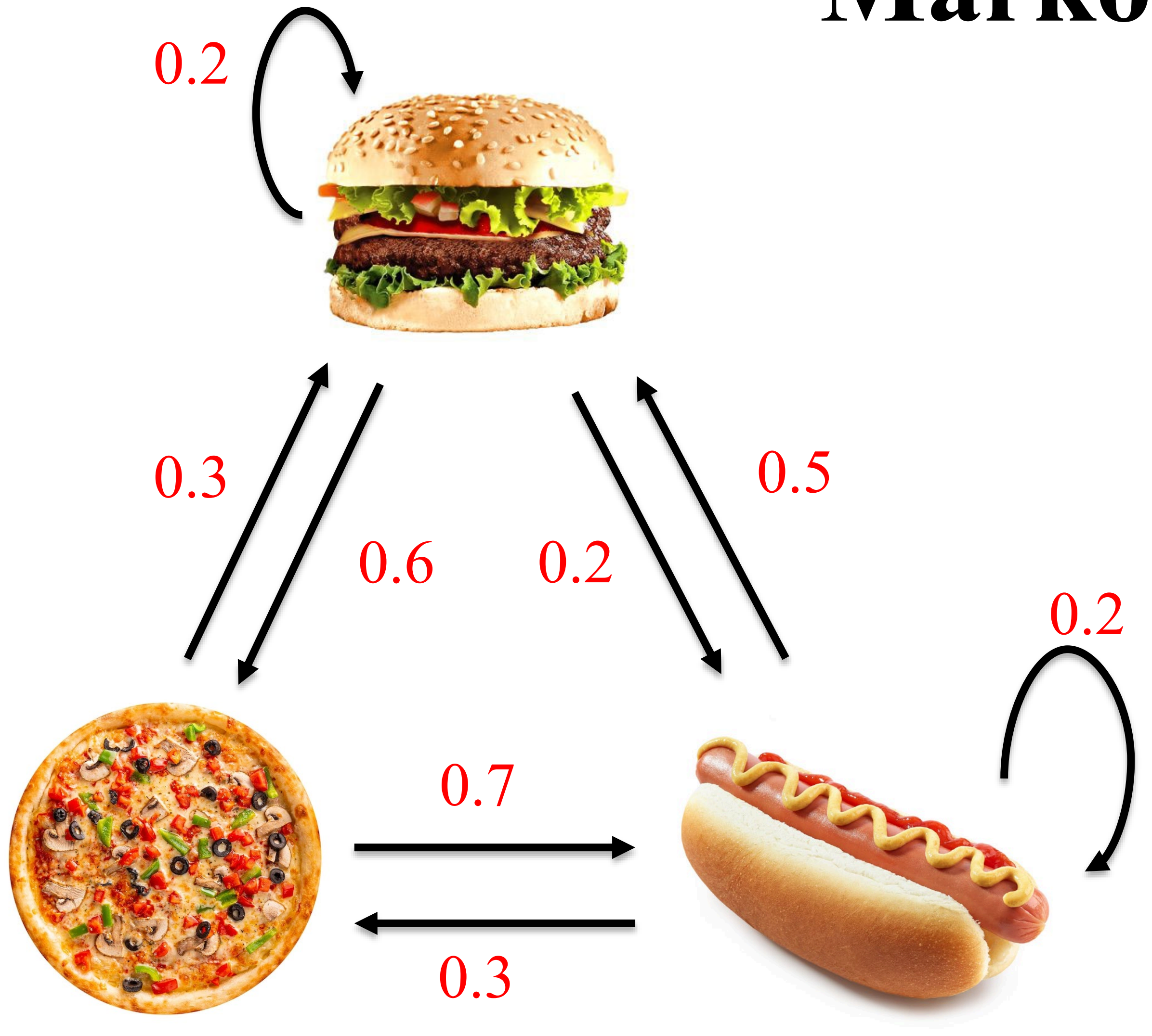
Lecture 20

Markov Chain and Hidden Markov Model

Dr. Tamal Ghosh
Department of CSE
Adamas University

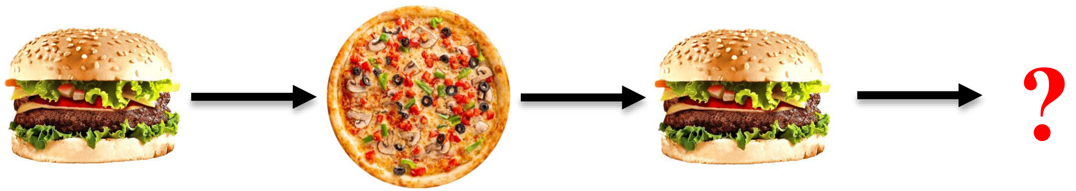
Markov Chain

☞ "Memoryless process."
☞ Only the **current state matters**, not how you got there



$$P(X_i|X_{i-1},........, X_1) = P(X_i\ |Parents(X_i\ ))$$

$$P(X_{n+1} = x\ |\ X_n = x)$$



$P(X_4=$  $|\ X_3=$  $) = 0.2$

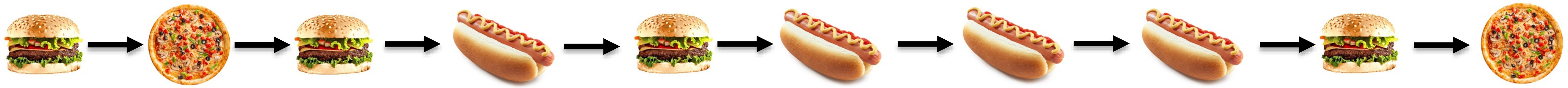
A Markov Chain is a stochastic (random) process that satisfies the Markov Property, meaning: The future state depends only on the present state, not on the sequence of past states.

Contd…

Components of a Markov Chain

1. **States** → Possible conditions (e.g., Sunny, Rainy)
2. **Transition Probabilities** → Probability of moving from one state to another
3. **Transition Matrix** → Table of all transition probabilities

- **Random Walk:**



long-term expected probabilities

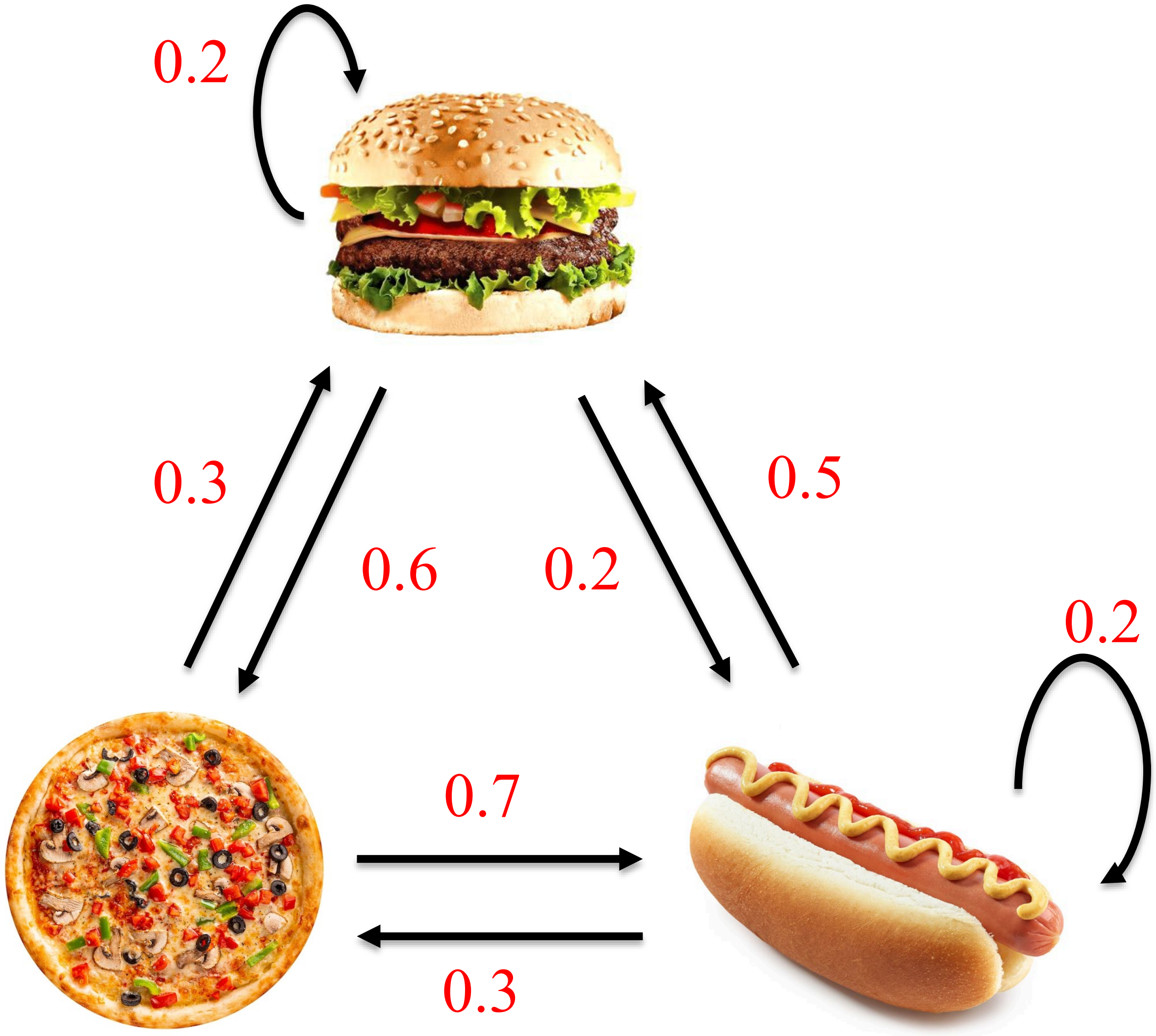
P() =4/10 = 0.4 P() =2/10=0.2 P() =4/10=0.4

For 10^5 random walk: 0.3519 0.2124 0.4356

This is also known as Stationary Distribution, so these states are also stationary and they will not change after an infinite number of random walks.

Reason why expected probabilities change 1. This is **empirical estimation** (simulation-based) 2. With finite steps (even 100k), there is still **sampling noise** 3. With more steps → it will get closer to the true stationary distribution

Markov Chain Representation



Adjacency/Transition matrix:

$$A = \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0 & 0.7 \\ 0.5 & 0.3 & 0.2 \end{bmatrix}$$

Finding Probability of States:

$$\pi_0 = [0 \quad 1 \quad 0]$$

$$\pi_0 A = [0.3 \quad 0 \quad 0.7] = \pi_1$$

$$\pi_1 A = [0.41 \quad 0.18 \quad 0.41] = \pi_2$$

$$\pi_2 A = [0.35 \quad 0.25 \quad 0.41] = \pi_1$$

Contd…

Formula:

$$\pi A = \pi$$

EigenVector Eqn:

$$Av = \lambda v$$

If we put λ =1 and reverse v's order of multiplication, we get the exact same equation which is the equilibrium state of the equation. Hence π is the eigenvector of matrix A.
This eigenvector also satisfies,
π[1] + π[2] + π[3]=1 *as it represents probability distribution.*

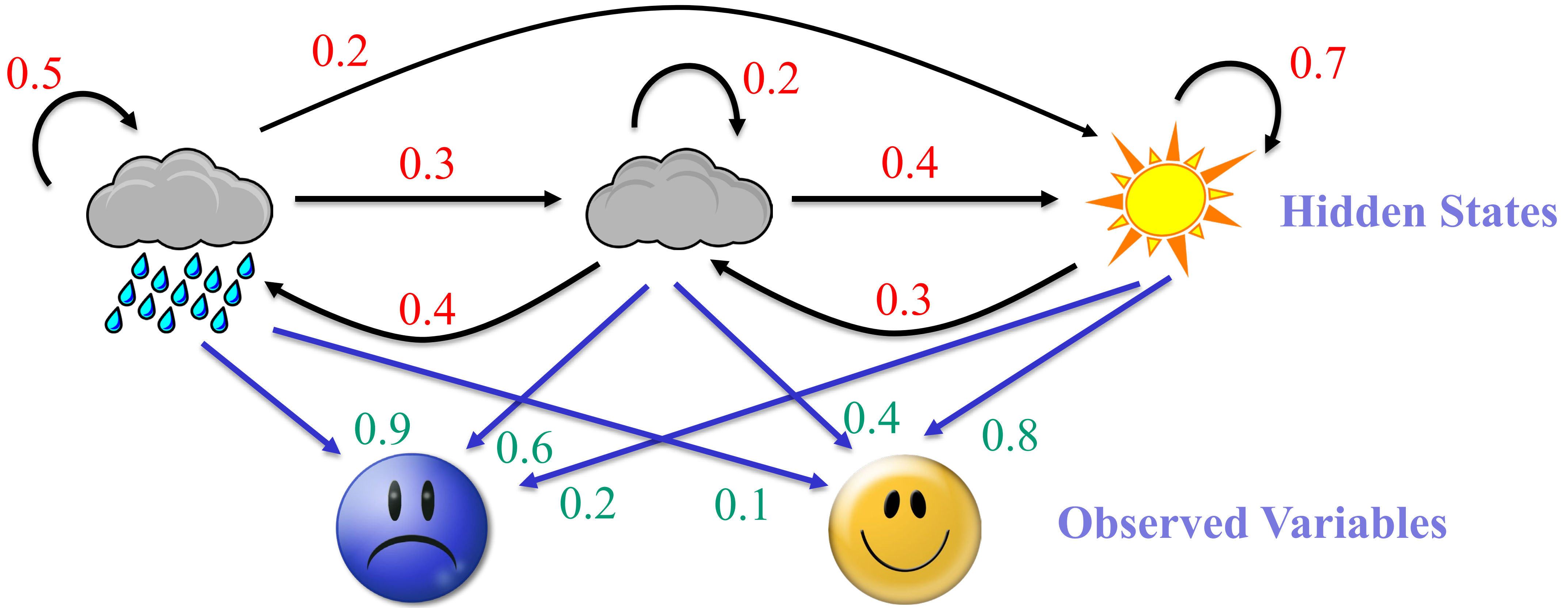
Contd…

- After solving these two equations,

$$\boldsymbol{\pi} = \begin{bmatrix} \frac{25}{71} & \frac{15}{71} & \frac{31}{71} \end{bmatrix} = \begin{bmatrix} 0.3521 & 0.2112 & 0.4366 \end{bmatrix}$$

- **This result is close to the simulation result.**
- **And we have the probability of hamburger being served 35%, pizza 21% and hotdogs 44% of the days.**

Hidden Markov Model

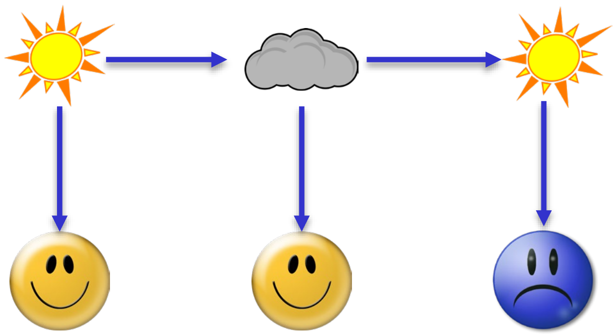


Hidden Markov Model = Hidden States of MC + Observed Variables

$$\begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.2 & 0.4 \\ 0.0 & 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix}$$

Transition Matrix Emission Matrix

Scenario:



Q1. What is the probability of this scenario occurring? Or what is the joint probability of observed mood sequence and weather sequence?

P(Y= , X=)

Using left eigenvector:

$$\pi A = \pi$$

[0.218 0.273

0.509]

? 0.8

P(X1=) P(Y1= , X1=)

0.3 0.4

P(X2= , X1=) P(Y2= , X2=)

0.4 0.2

P(X3= , X2=) P(Y3= , X3=)

Contd…

- Calculate the product:

0.509 x 0.8 x 0.3 x 0.4 x 0.4 x 0.2 = 0.00391

Q2. What is the most likely weather sequence given the observed mood sequence?

It could be ANY.

It is found that  **contributes to the maximum probability value. Which is 0.04105**

If the hidden states are represented by Xi and observed variables are Yi. The HMM can be presented as,

$$\underset{X=X1,\, X2,\, \ldots,\, Xn}{\text{argmax}} \; P(X1, X2, \ldots Xn \mid Y= Y1, Y2, \ldots, Yn)$$

Apply Bayes' Theorem

$$\underset{X=X1,\, X2,\, \ldots,\, Xn}{\text{argmax}} \; P(X1, X2, \ldots Xn \mid Y = Y1, Y2, \ldots, Yn)$$ can be written as,

$$\underset{X=X1,\, X2,\, \ldots,\, Xn}{\text{argmax}} \; P(Y|X)P(X)/P(Y)$$

$P(Y|X) = \Pi\ P(Y_i \mid X_i)$

$P(X) = \Pi\ P(X_i \mid X_{i-1})$

- So finally we have,

$$Argmax_{X=x1,x2,\ldots,xn} \prod P(Yi \mid X_i) P(X_i \mid X_{i-1})$$