Mathematics for Computer Science
CSE 401
Markov Model

Shah Murtaza Rashid Al Masud, PhD
Department of CSE
University of Asia Pacific

Markov chain

a process with a finite number of states (or outcomes, or events) in which the probability of being in a particular state at step n + 1 depends only on the state occupied at step n.

Prof. Andrei A. Markov (1856-1922) , published his result in 1906.

Markov chain

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

Markov chains are useful in understanding Markov models, and, in particular, hidden Markov models used for data science applications.

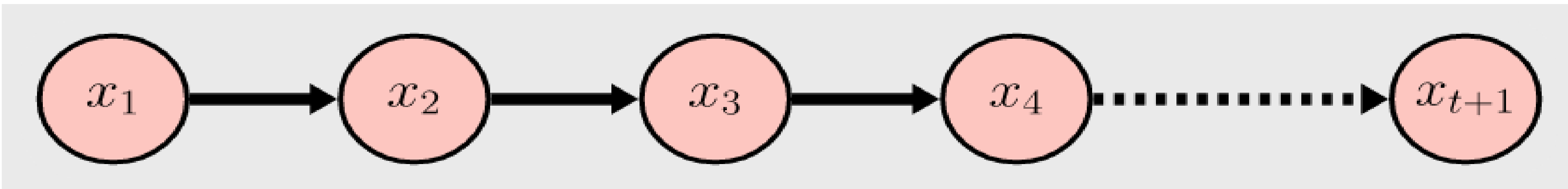
Markov chain

A Markov chain is one of the simplest Markov models. This chain assumes that an observation x_{t+1} at a future time $t+1$ is only dependent on the observation x_t at the current timestamp t.

In other words, given the present observation, the future is independent of the past.

$$P(x_{t+1}|x_1, x_2, \ldots, x_t) = P(x_{t+1}|x_t)$$

We use the following graphical model to denote a Markov chain.



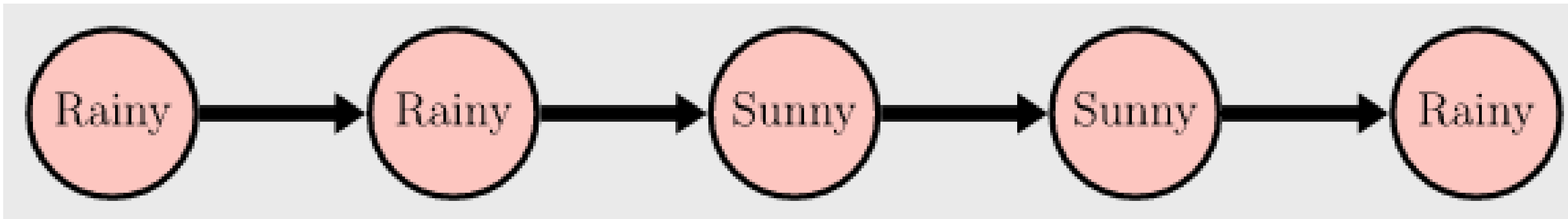
Markov chain graphical diagram

Nodes colored in pink denote observations. The arrows indicate the dependencies between the present and the past.

Using the rules of independence, we can calculate the joint probability of the sequence as:

$$P(x_1, x_2, \ldots, x_{t+1}) = P(x_1)P(x_2|x_1)P(x_3|x_2)\ldots P(x_t|x_{t-1})P(x_{t+1}|x_t)$$

Markov chain Example



Markov chain for Sunny and Rainy Weather

Assume a scenario where you observe the weather for a place. The above Markov chain denotes the change of weather. The weather of the next day depends on the previous day.

Parameters of Markov chain

Each observation x_t (at time t) can take a discrete value or "state." For example, in the case of weather, the states were *Rainy* or *Sunny*, in the case of coin swap, the states were *Biased* or *Fair*, and in the case of the air conditioner, the states were *ON* or *OFF*. We generally assume that the observation can take one of the K states.

Let us now understand the parameters for a Markov chain. We can rewrite the factorization of the above general Markov chain as:

$$P(x_1, x_2, \ldots, x_T) = P(x_1) \prod_{t=2}^{T} P(x_t|x_{t-1})$$

Markov chains assume that the conditional probability $P(x_t|x_{t-1})$ does not vary with time.

Parameters of Markov chain

Time-homogeneous Markov chains (or stationary Markov chains) are processes where

$$P[X_{t+1} = x |\, X_t = y] = P[X_t = x |\, X_{t-1} = y]$$

for all t. The probability of the transition is independent of t.

Parameters of Markov chain

Therefore, we can fully specify a Markov chain using two parameters as below:

- **Transition Matrix** (A): The transition matrix stores the probability of transition between the state i to state j. Thus, the transition matrix can be represented as a K x K matrix where the entry A_{ij} is given by $A_{ij} = P(x_t = j | x_{t-1} = i)$ where $i, j \in \{1, 2 \ldots K\}$.
- **Prior Probability** (π): The probability of starting from one of the available states. It is denoted by $\pi_i = P(x_1 = i)$ where $i \in \{1, 2 \ldots K\}$.

Markov Chains Transition Matrix

The value P_{ij} represents the probability that the process will, when in state i, next make a transition into state j. Since probabilities are nonnegative and since the process must make a transition into some state, we have

$$P_{ij} \geqslant 0, \qquad i, j \geqslant 0; \qquad \sum_{j=0}^{\infty} P_{ij} = 1, \qquad i = 0, 1, \ldots$$

Let P denote the matrix of one-step transition probabilities P_{ij}, so that

$$\mathrm{P} = \left\| \begin{array}{cccc} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ \vdots & \vdots & \vdots & \\ P_{i0} & P_{i1} & P_{i2} & \cdots \\ \vdots & \vdots & \vdots & \end{array} \right\|$$

Markov Chains Transition Matrix Features

It is square, since all possible states must be used both as rows and as columns.

All entries are between 0 and 1, because all entries represent probabilities.

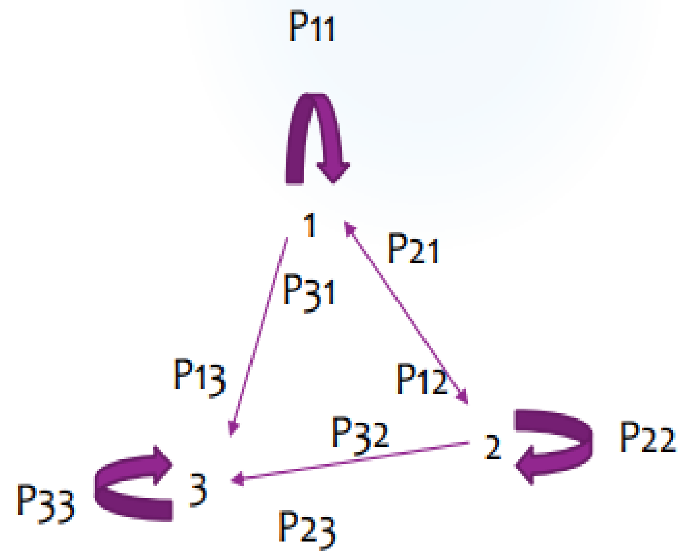
The sum of the entries in any row must be 1, since the numbers in the row give the probability of changing from the state at the left to one of the states indicated across the top.

Markov Chains Transition Matrix

Example 1:

X= Living status. State space, S= {1, 2, 3}; where, 1= Healthy, 2= Sick, 3 = Dead.

$$P = \begin{array}{cc} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} \end{array}$$

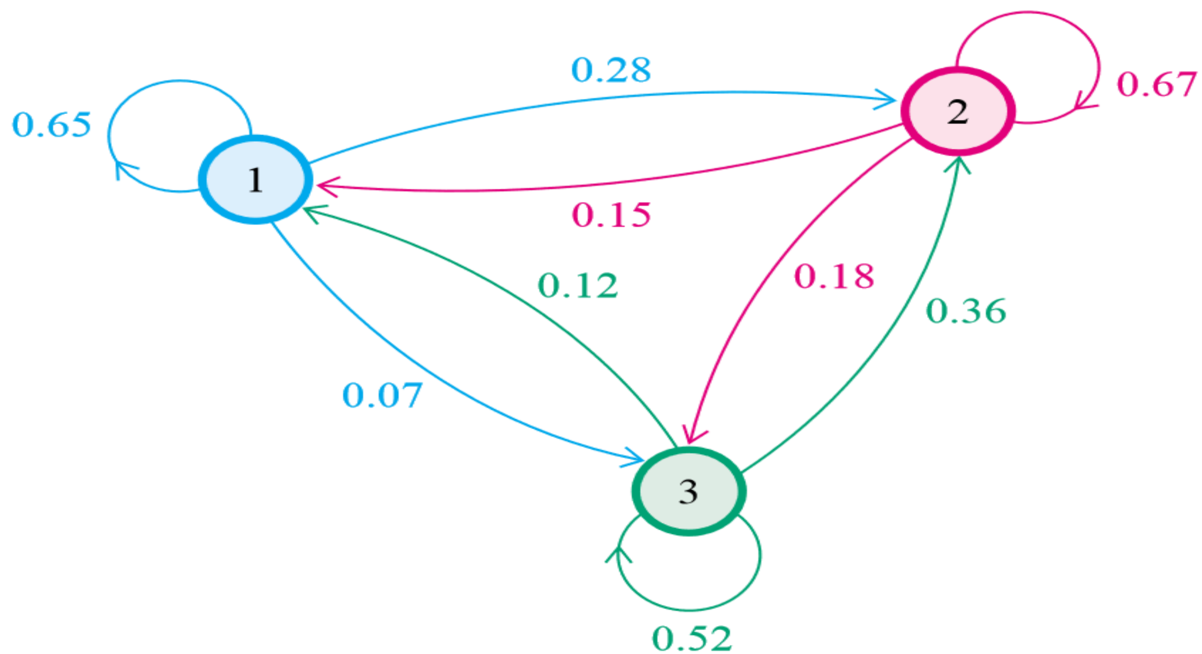


Here, $P_{11} > 0, P_{12} > 0, P_{13} > 0, P_{21} > 0, P_{22} > 0, P_{23} > 0,$
$P_{31} = 0, P_{32} = 0, P_{33} = 1$

Example:

$$\text{State } n+1 \left\{ \begin{array}{c} S_1 \\ S_2 \\ S_3 \end{array} \right. \overset{\displaystyle \overbrace{\begin{array}{ccc} S_1 & S_2 & S_3 \end{array}}^{\text{State } n}}{\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}}$$

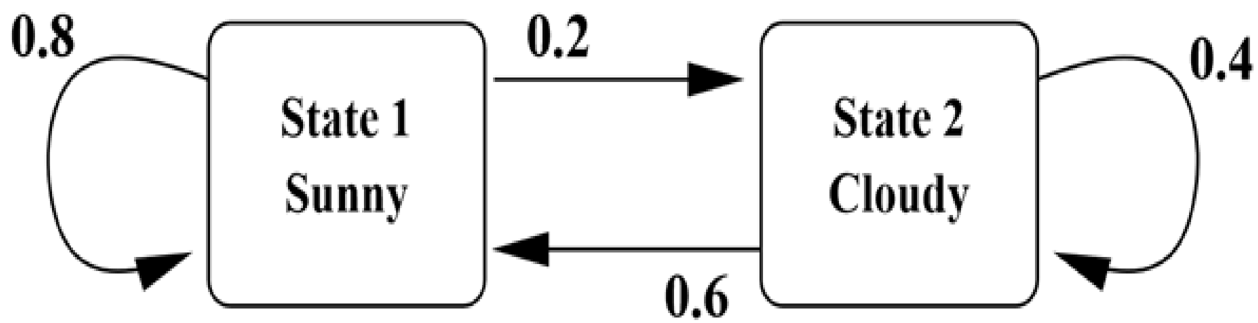
$$\begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 0.65 & 0.28 & 0.07 \\ 0.15 & 0.67 & 0.18 \\ 0.12 & 0.36 & 0.52 \end{bmatrix} \end{array} = P.$$



Example: Sunny or Cloudy. A meteorologist studying the weather in a region decides to classify each day as simply *sunny* or *cloudy*. After analyzing several years of weather records, he finds:

- the day after a sunny day is sunny 80% of the time, and cloudy 20% of the time; and
- the day after a cloudy day is sunny 60% of the time, and cloudy 40% of the time.

We can setup up a Markov chain to model this process. There are just two states: $S_1 = \textit{sunny}$, and $S_2 = \textit{cloudy}$. The transition diagram is



and the transition matrix is

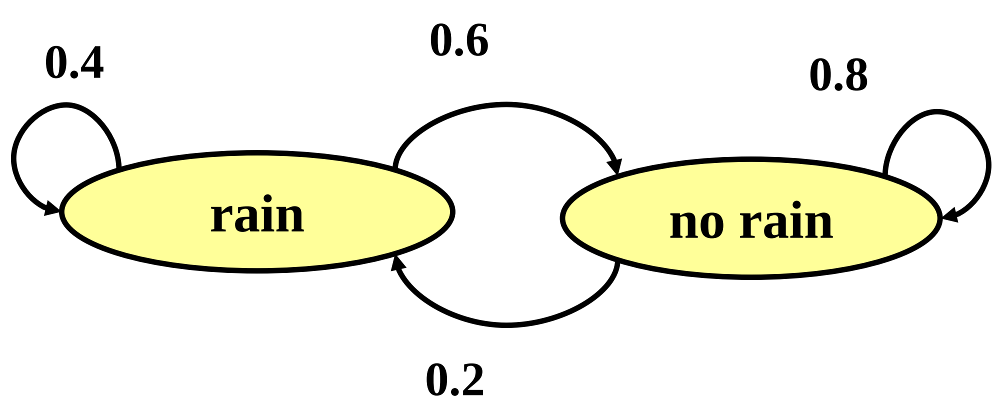
$$P = \begin{bmatrix} 0.8 & \mathbf{0.2} \\ \mathbf{0.6} & 0.4 \end{bmatrix}.$$

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 Finite State Machine (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

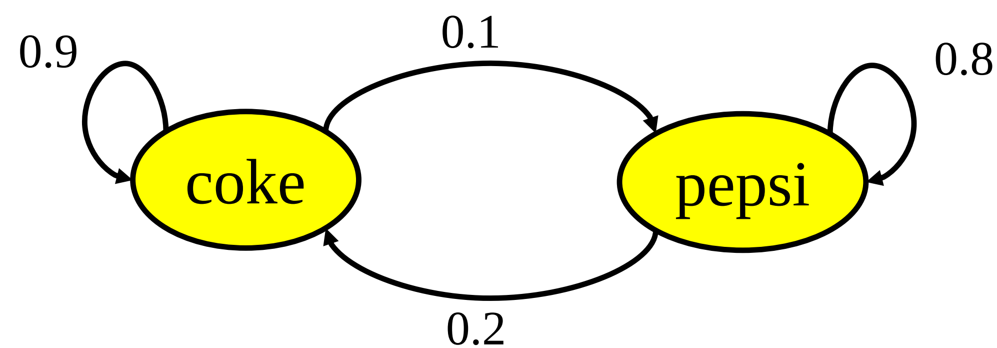
Markov Process Coke vs. Pepsi Example

- 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.

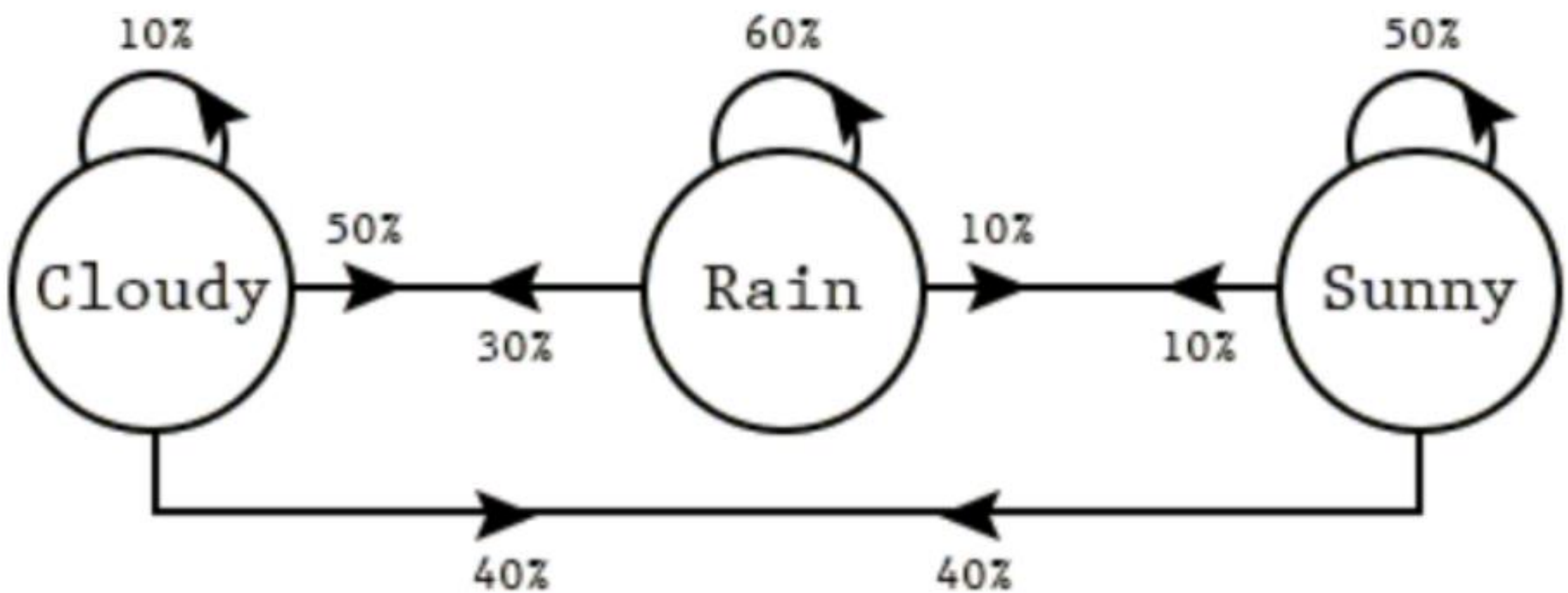
transition matrix:

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

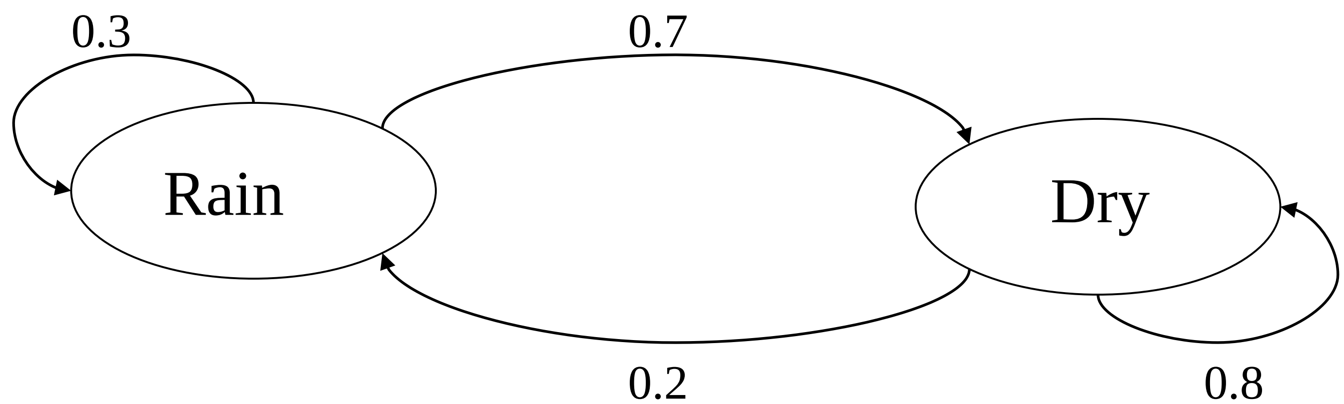


Markov chain Example

The diagram below represents a Markov chain where there are three states representing wheather of the day (cloudy, rainy and sunny). And, there are transitiion probabilities representing the weather of the next day given the weather of current day.



Example of Markov Model/Chain



- Two states : ‘Rain’ and ‘Dry’.
- Transition probabilities: P(‘Rain’|‘Rain’)=0.3 , P(‘Dry’|‘Rain’)=0.7 , P(‘Rain’|‘Dry’)=0.2, P(‘Dry’|‘Dry’)=0.8
- Initial probabilities: say P(‘Rain’)=0.4 , P(‘Dry’)=0.6 .

Calculation of sequence probability

- By Markov chain property, probability of state sequence can be found by the formula:

$$P(s_{i1}, s_{i2}, \ldots, s_{ik}) = P(s_{ik} \mid s_{i1}, s_{i2}, \ldots, s_{ik-1}) P(s_{i1}, s_{i2}, \ldots, s_{ik-1})$$
$$= P(s_{ik} \mid s_{ik-1}) P(s_{i1}, s_{i2}, \ldots, s_{ik-1}) = \ldots$$
$$= P(s_{ik} \mid s_{ik-1}) P(s_{ik-1} \mid s_{ik-2}) \ldots P(s_{i2} \mid s_{i1}) P(s_{i1})$$

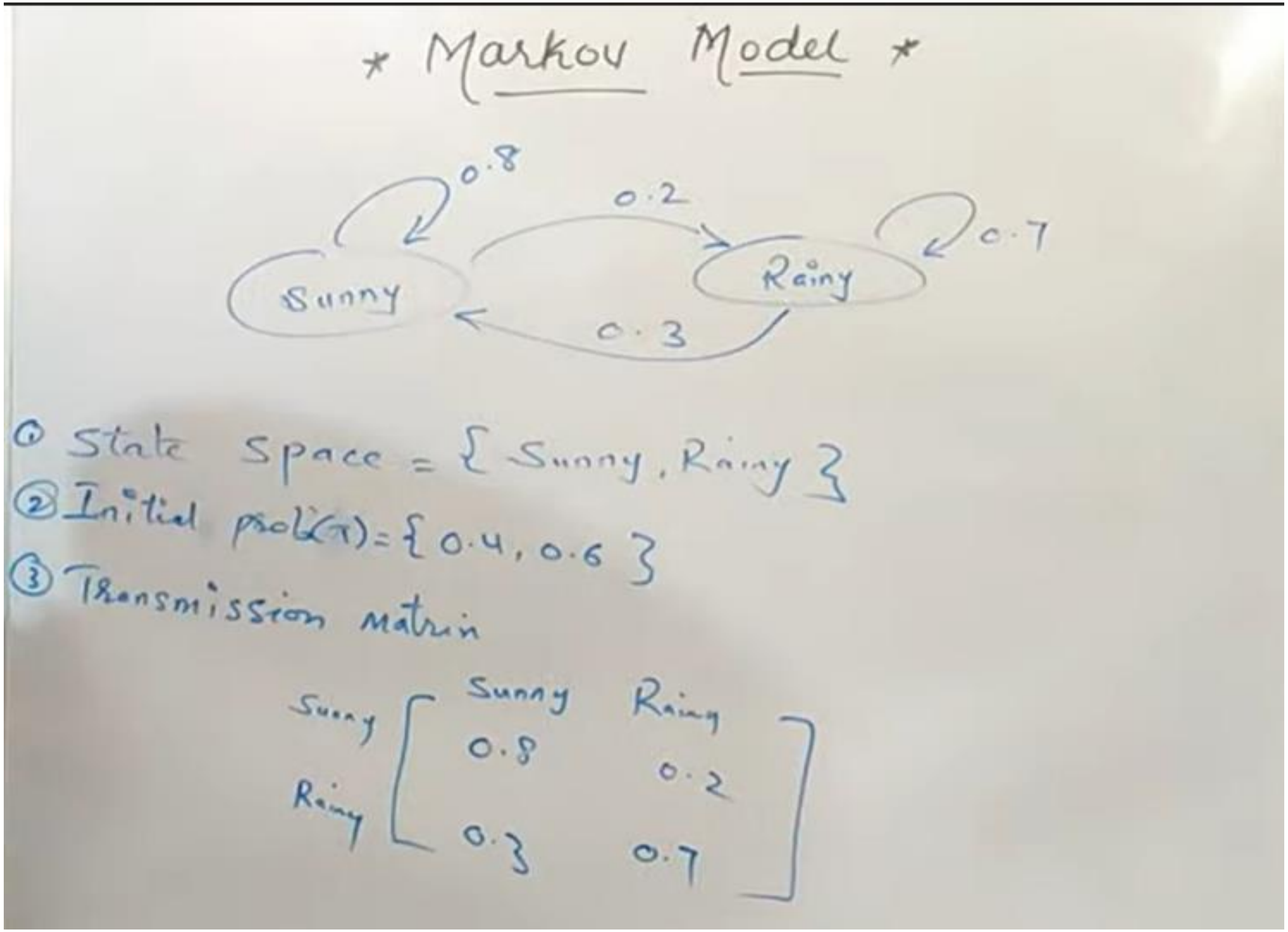
- Suppose we want to calculate a probability of a sequence of states in our example, {‘Dry’,’Dry’,’Rain’,Rain’}.

P({‘Dry’,’Dry’,’Rain’,Rain’}) =

P(‘Rain’|’Rain’) P(‘Rain’|’Dry’) P(‘Dry’|’Dry’) P(‘Dry’)=

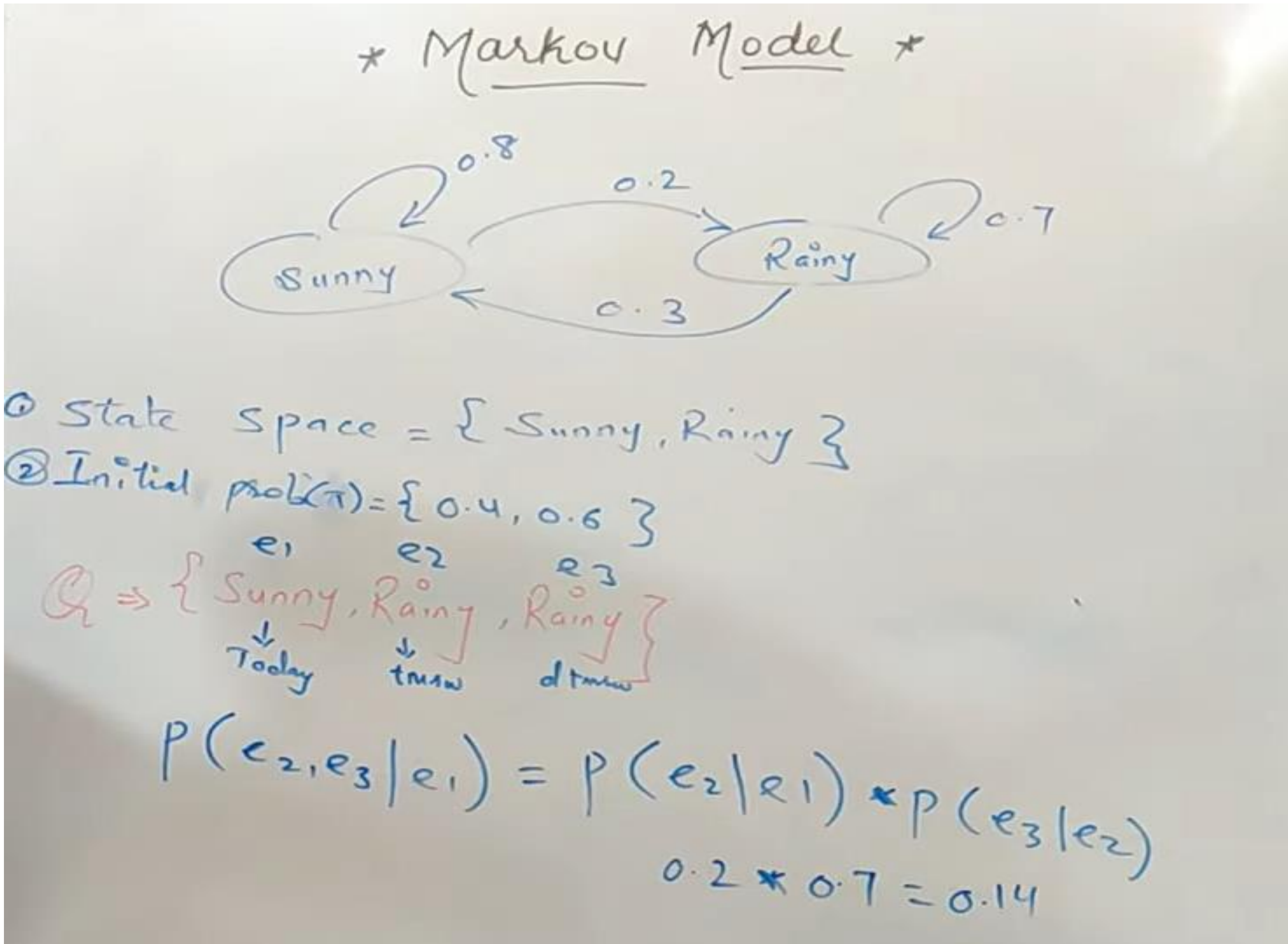
= 0.3*0.2*0.8*0.6

Example: Draw the transition matrix



Example: Find the probability of any given sequence.

Today is sunny, what is the probability that tomorrow and day after tomorrow will be rainy and rainy

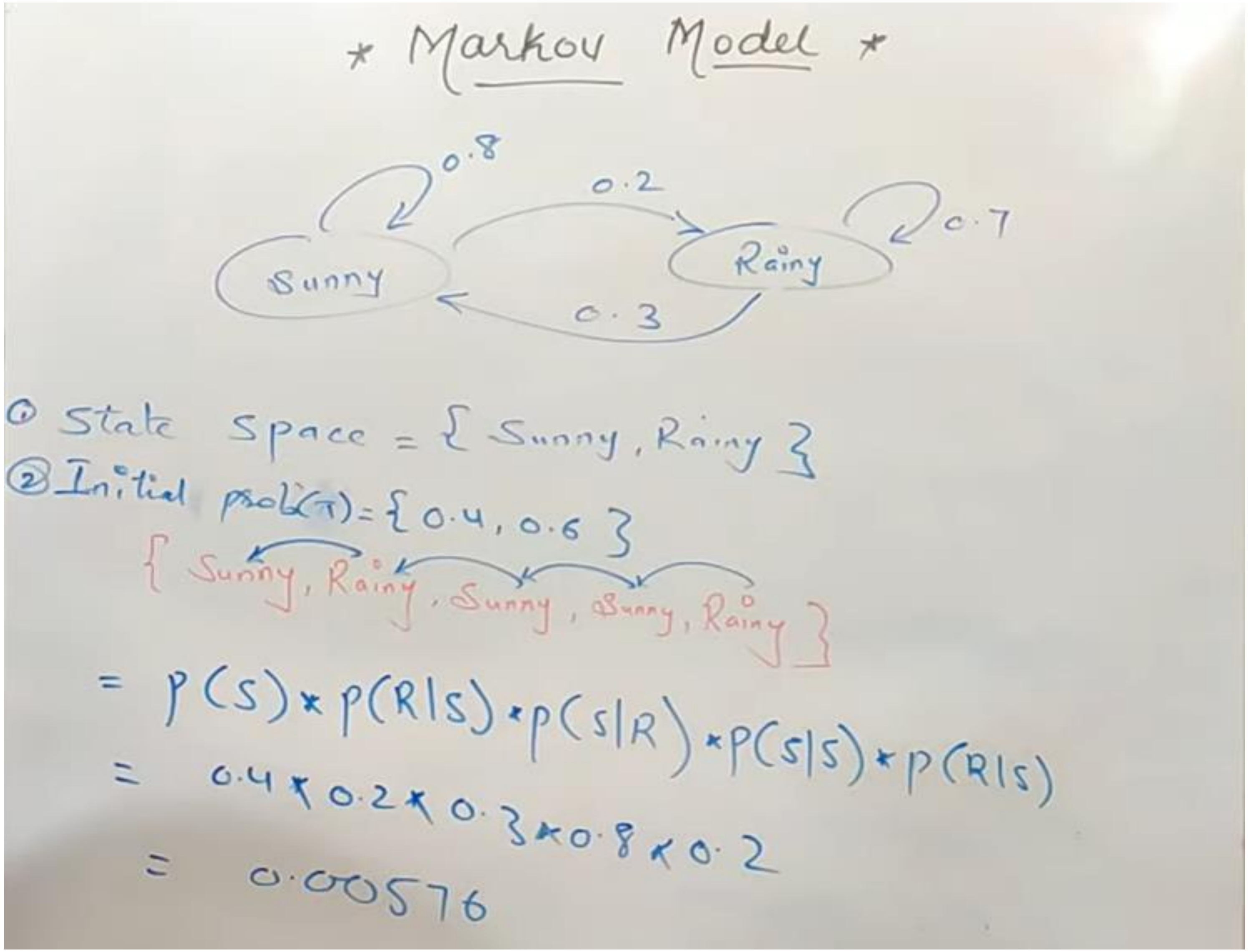


Event e2 and event e3 will happen when Event e1 already happened

Today is sunny,

Example: **A series is given, nothing is said that today's weather is given and followed by other elements of the series.**

What is the probability that the series will happen



The first event does not depend on anything. We write P(s) which is the initial probability. For the second transition: write probability of Rainy when sunny already happened and so on. 0.00576 is the answer, means this is the probability that the series will happen

Thank you