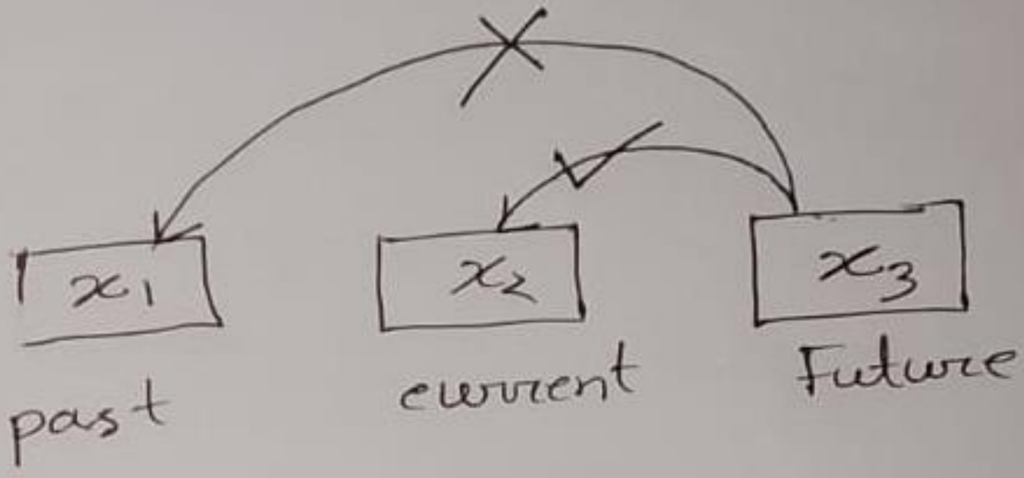
Markov Model

Introduction

- A markov model is a stochastic model used to model randomly changing system.
- It is assumed that future state depends on the current state, not on the events that occurred before.
- It is used for sequential data prediction.

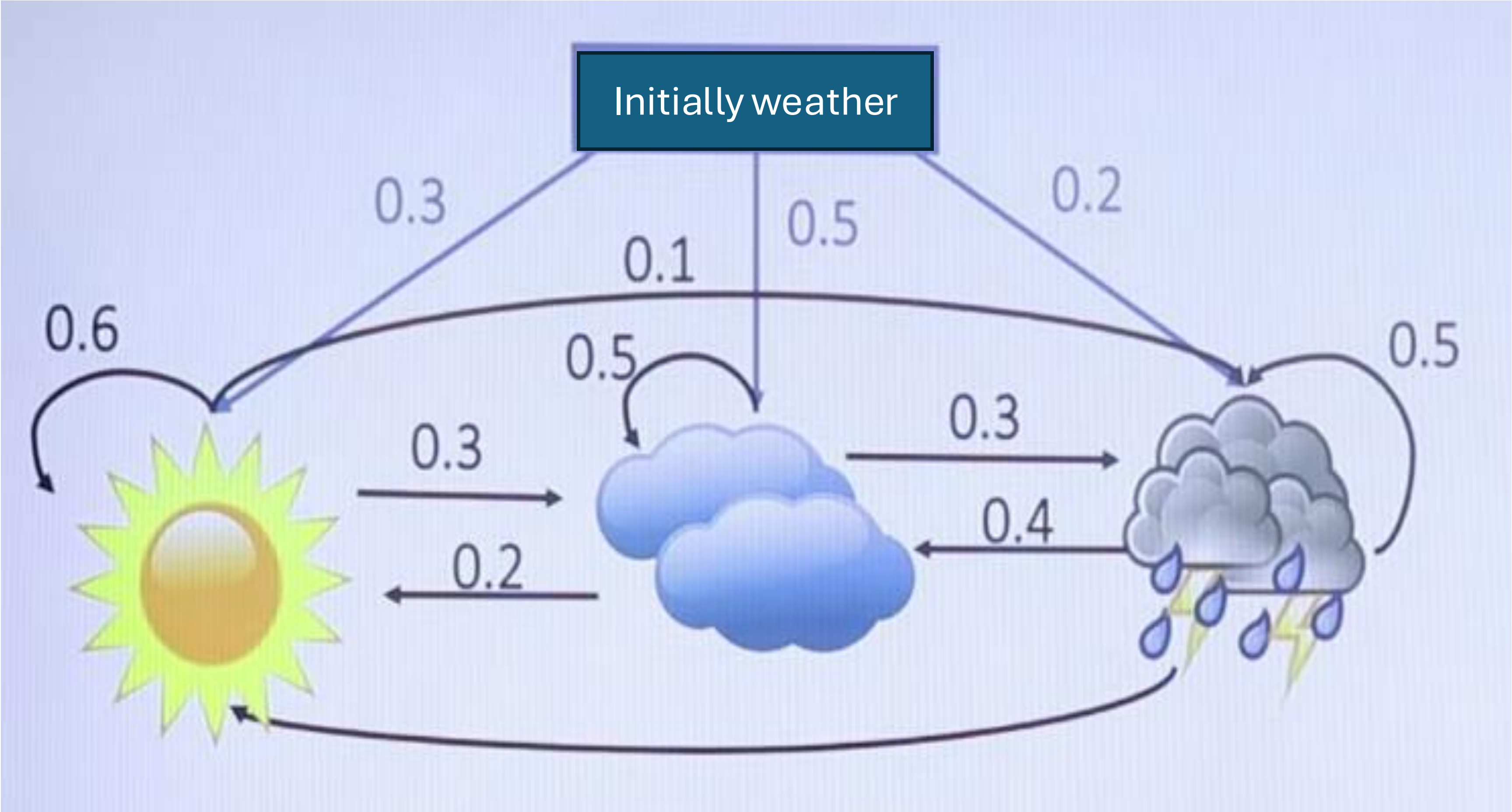


$$S = \{x_1, x_2, x_3\}$$

[$\because$ State Space, Set of states

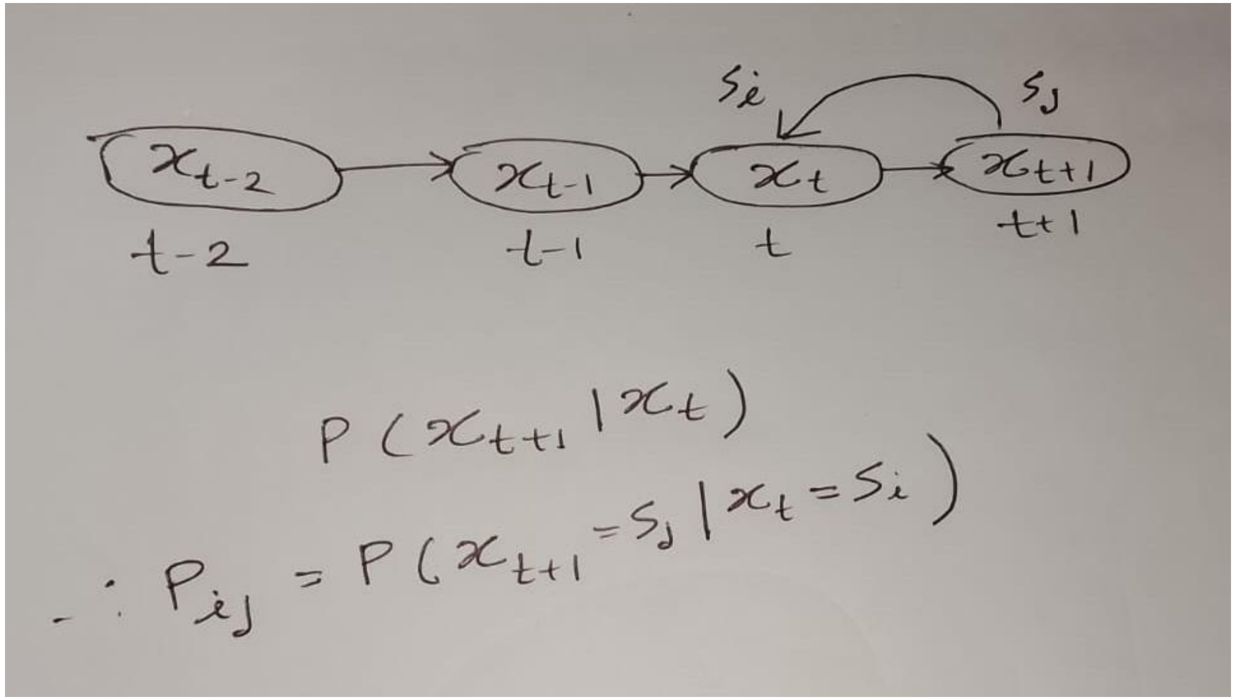
we have N distinct state

$$N = 3$$


Initially weather
0.3
0.5
0.2
0.1
0.6
0.5
0.5
0.3
0.3
0.4
0.2

Markov Chain

- Any chain of event that follow the markov property that chain is called markov chain.



Mathematical Definition

- Let $S = \{s_1, s_2, s_3, \ldots, s_n\}$ be the set of possible states.
- Then, for any time step t:

$$P(X_{t+1} = s_j \mid X_t = s_i, X_{t-1} = s_k, \ldots, X_0 = s_m) = P(X_{t+1} = s_j \mid X_t = s_i)$$

- This means:
  The probability of going to state s_j at time $t + 1$ depends only on the current state s_i.
- This is called first order markov model.

Transition Matrix Example

State Space = { Sunny, Rainy, Cloudy }
Initial state distribution π = { 0.3, 0.5, 0.2 }

Sunny	Cloudy	Rainy
0.3	0.5	0.2

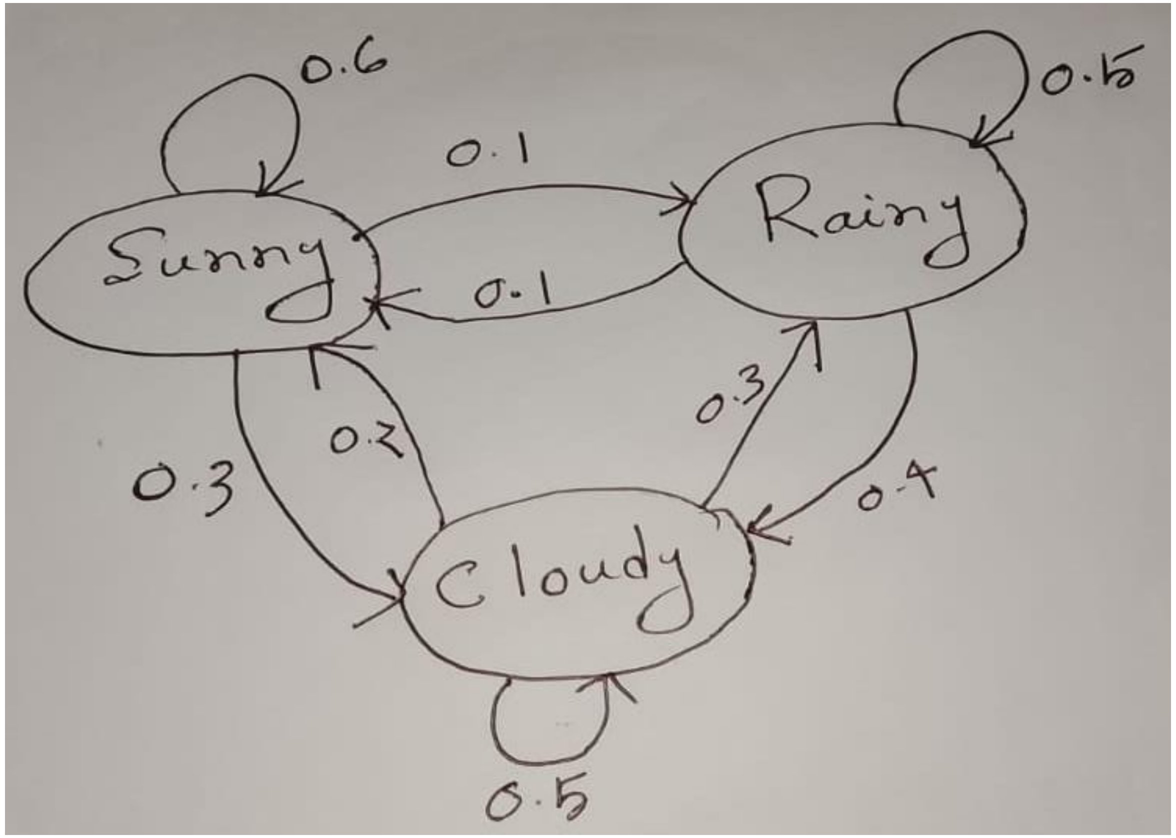
Transition Matrix

Future State

Present state

	Sunny	Cloudy	Rainy
Sunny	0.6	0.3	0.1
Cloudy	0.2	0.5	0.3
Rainy	0.1	0.4	0.5

Transitional probability



State Transition Diagram

Question 1

- Given that today the weather is sunny, what is the probability that tomorrow is sunny and day after is raning?
- Solution:

$t_1 = \text{Sunny}, \; t_2 = \text{Sunny}, \; t_3 = \text{Rainy}$

$P(t_3 = \text{Rainy}, t_2 = \text{Sunny} \mid t_1 = \text{Sunny})$

$= P(t_3 = \text{Rainy} \mid t_2 = \text{Sunny}) \times P(t_2 = \text{Sunny} \mid t_1 = \text{Sunny})$

$= 0.1 \times 0.6 = 0.06$

Question 2

- Given that today the weather is cloudy and yesterday was rainy, what is the probability that tomorrow would be sunny?

Question 3

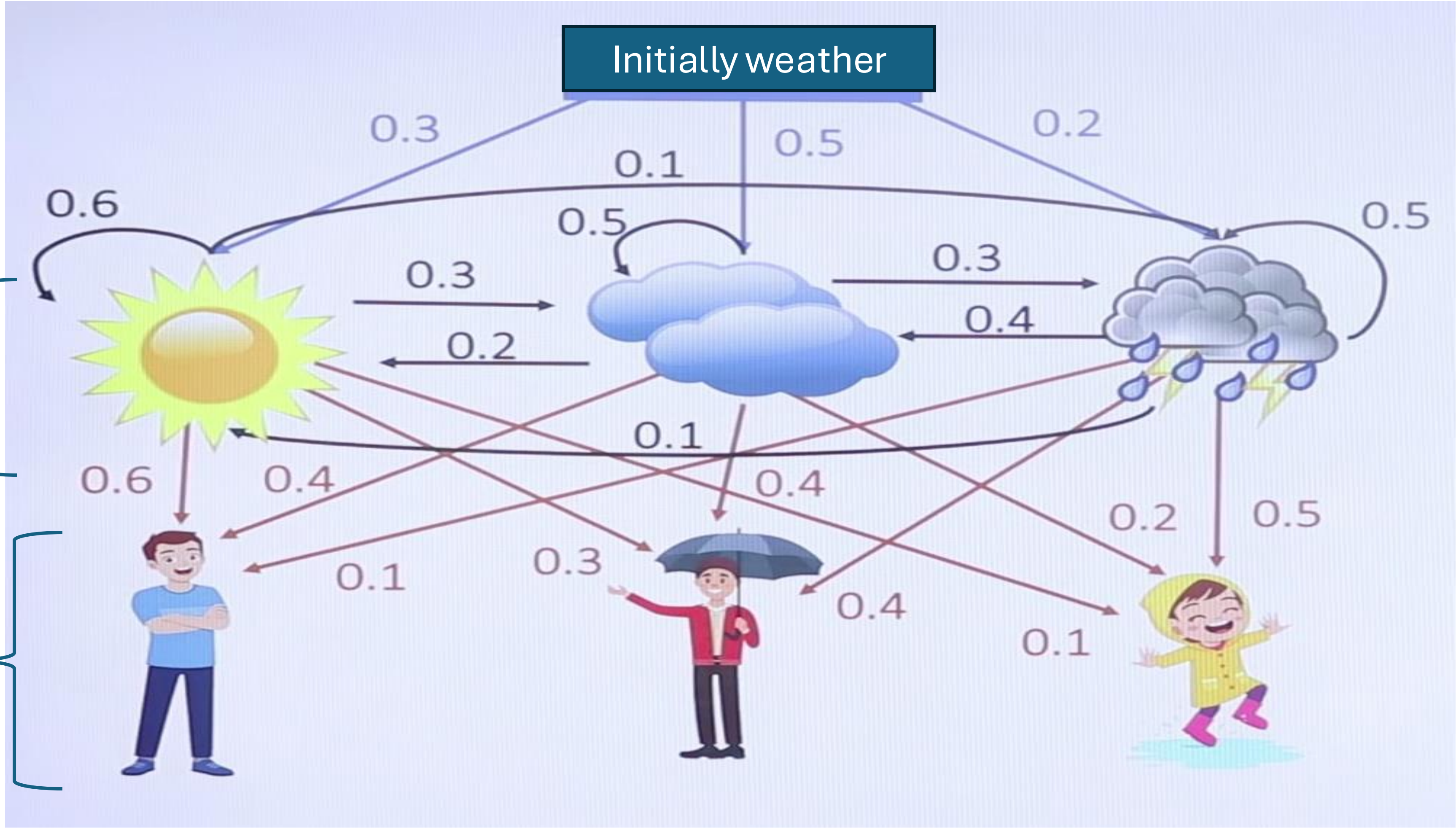
- What is the probability of given series?
  - Sunny -> Rainy -> Rainy -> Rainy -> Cloudy -> Cloudy
- Solution:
  - $P(S) * P(R \mid S) * P(R \mid R) * P(R \mid R) * P(C \mid R) * P(C \mid C)$

  $= 0.3 \times 0.1 \times 0.5 \times 0.5 \times 0.4 \times 0.5 = 0.0015$

Hidden Markov Model

Hidden layer

Observation layer



Hidden Markov Model (Cont'd)

- It is a statistical model that is used to describe the probabilistic relationship between a sequence of observations and a sequence of hidden states.
- An HMM consists of two types of variables:
  - The **hidden states** are the underlying variables that generate the observed data, but they are not directly observable.
  - The **observations** are the variables that are measured and observed.
- HMMs are widely used in speech recognition, natural language processing, bioinformatics, and many other fields.

Hidden Markov Model (Cont'd)

- The Hidden Markov Model (HMM) is the relationship between the hidden states and the observations using two sets of probabilities: the transition probabilities and the emission probabilities.
    - The **transition probabilities** describe the probability of transitioning from one hidden state to another.
    - The **emission probabilities** describe the probability of observing an output given a hidden state.
-

Initial state distribution π

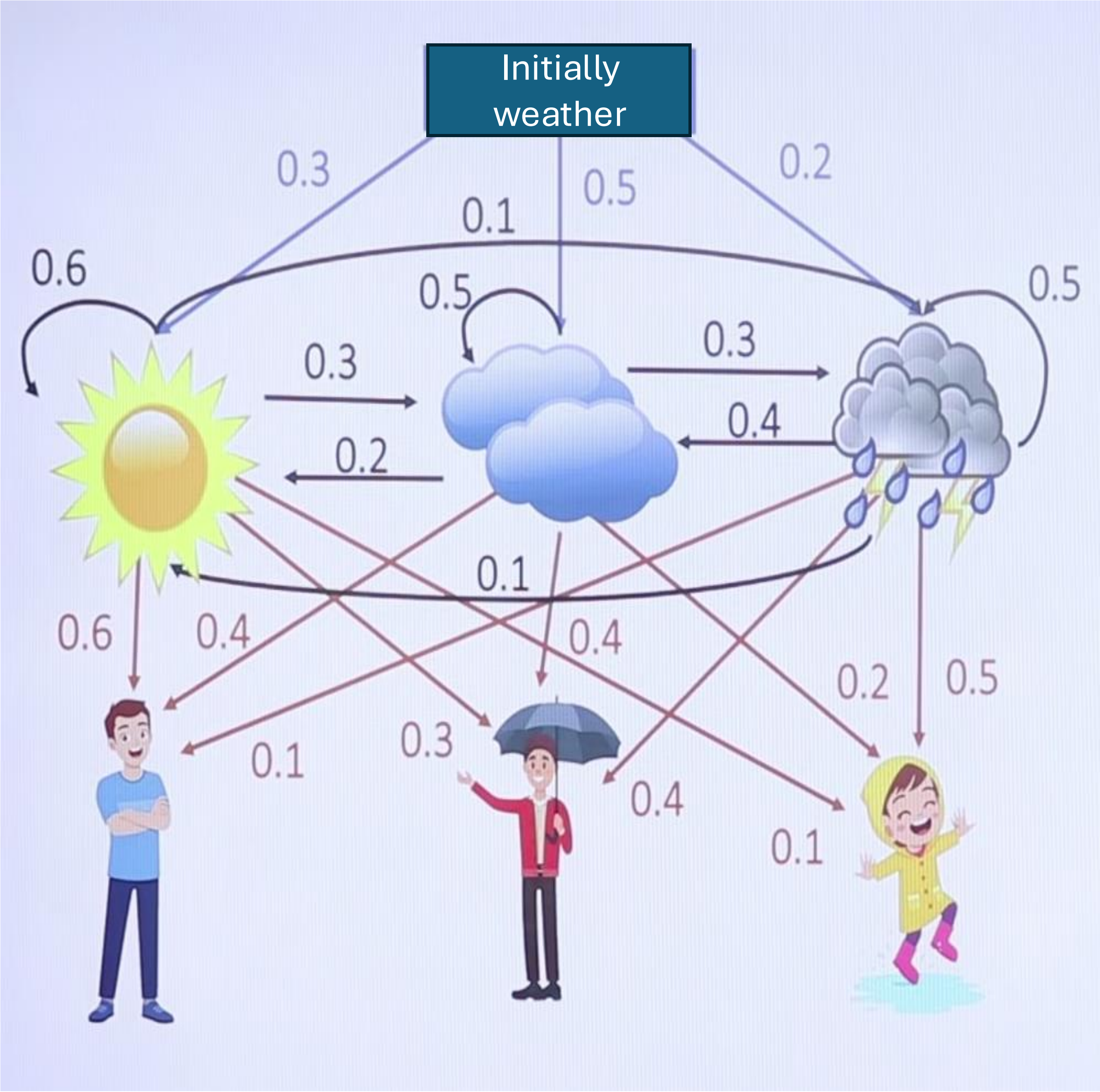
Sunny	Cloudy	Rainy
0.3	0.5	0.2

Transition Matrix

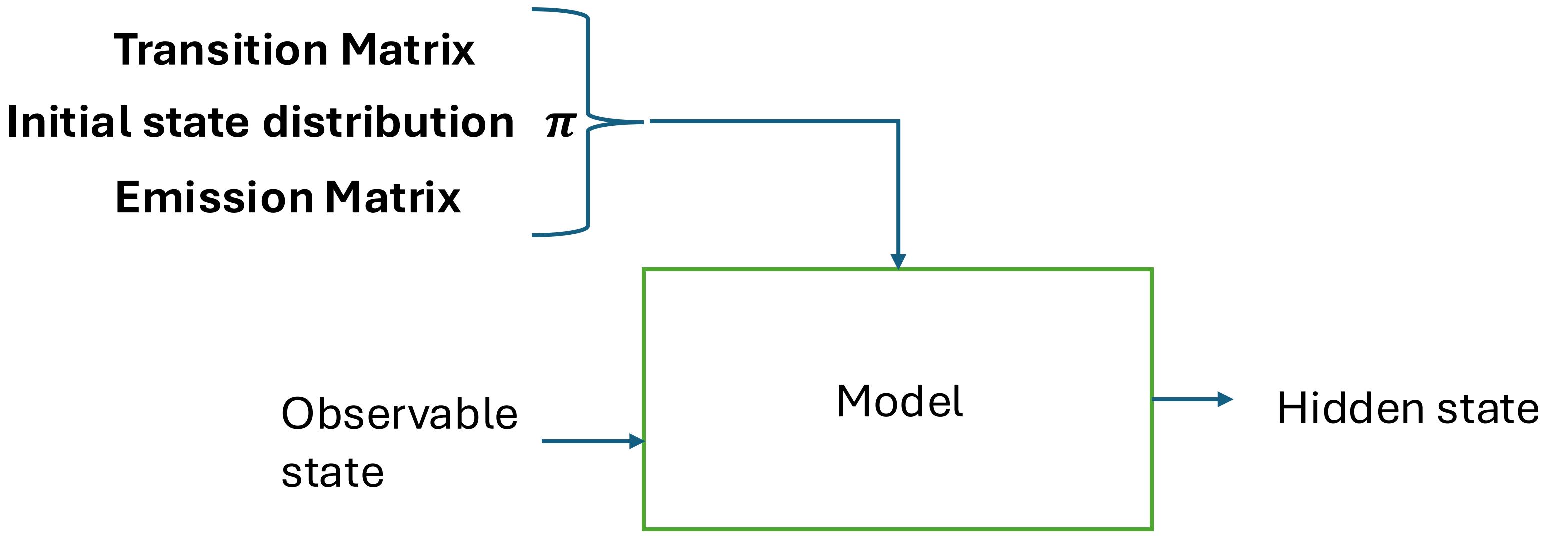
	Sunny	Cloudy	Rainy
Sunny	0.6	0.3	0.1
Cloudy	0.2	0.5	0.3
Rainy	0.1	0.4	0.5

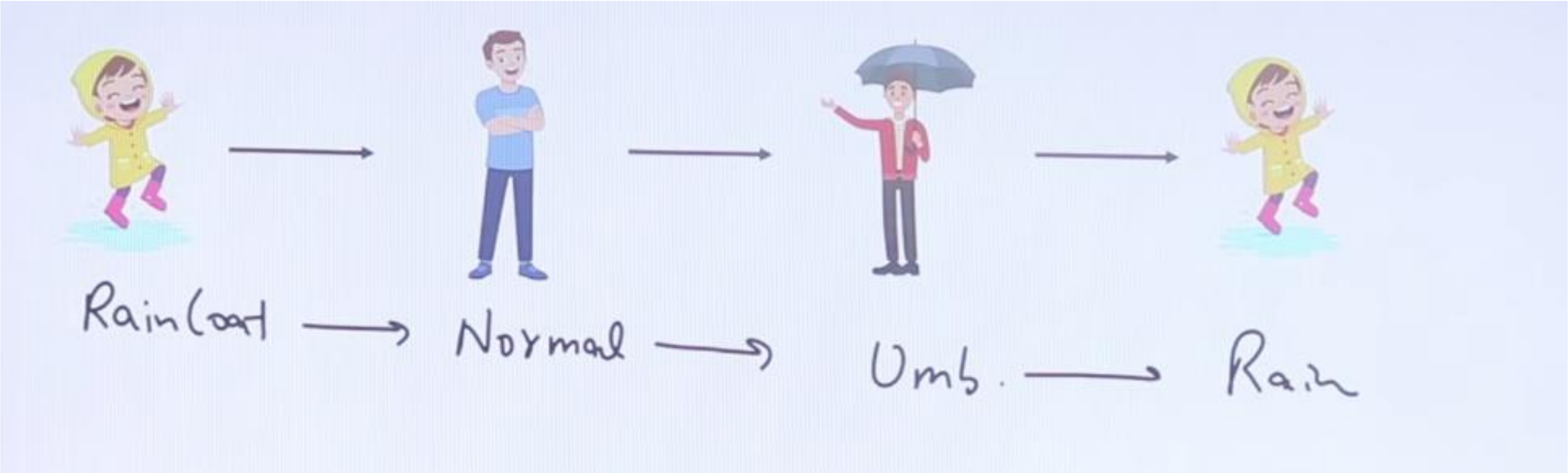
Emission Matrix

	Normal	Umbrella	Raincoat
Sunny	0.6	0.3	0.1
Cloudy	0.4	0.4	0.2
Rainy	0.1	0.4	0.5



Hidden Markov Model (Cont'd)





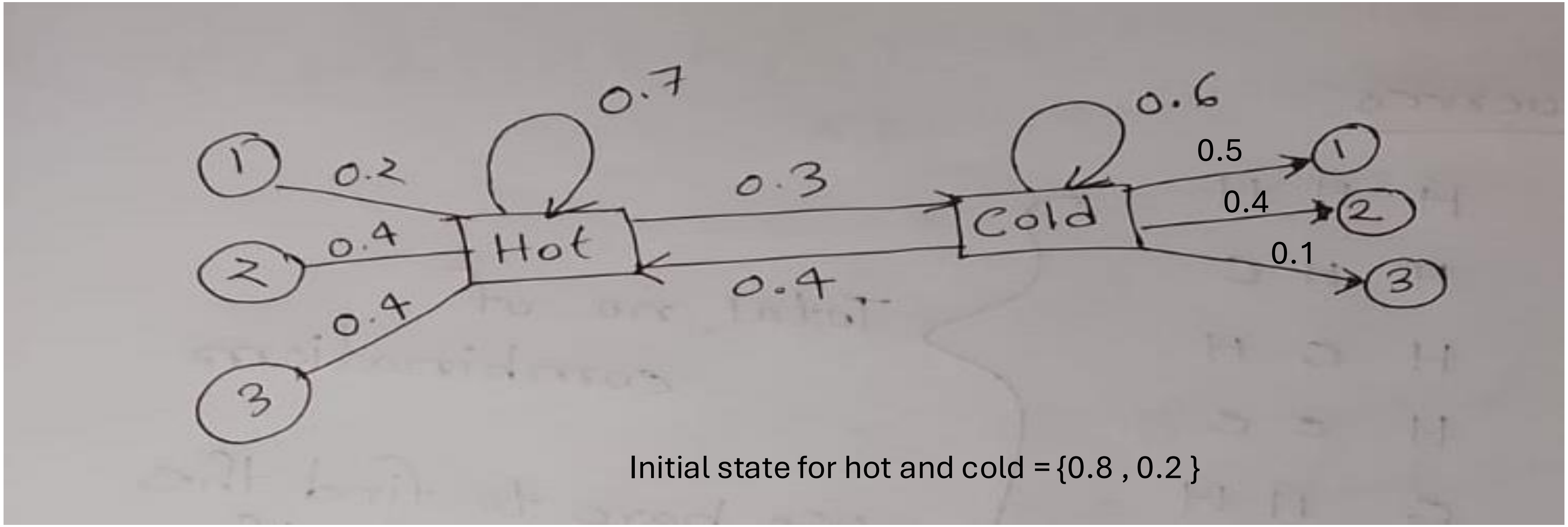
N = No of hidden state = 3

T = Given no of observation = 4

Total no of combination = N^T = 3^4 = 81

Question 1

- If the observable state is (1, 3, 1) find the hidden state



Transition Matrix

$$\begin{array}{c|cc} & HOT & COLD \\ \hline HOT & 0.7 & 0.3 \\ COLD & 0.4 & 0.6 \end{array}$$

$$\pi = [\underset{HOT}{0.8} \quad \underset{COLD}{0.2}]$$

Emission Matrix

$$\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline HOT & 0.2 & 0.4 & 0.4 \\ COLD & 0.5 & 0.4 & 0.1 \end{array}$$

∴ Joint Probability = $P(a,b) = P(a|b)\,P(b)$

$$
\begin{tikzcd}
\boxed{H} \arrow[r] \arrow[d] & \boxed{C} \arrow[r] \arrow[d] & \boxed{H} \arrow[d] \\
1 & 3 & 1
\end{tikzcd}
$$

$P(1,H) = P(1|H)\,.\,P(H)$

∴ $P(O|Q) = P(O|Q)\,P(Q)$

- Let's assume,
  Hidden state Q
  Observable " O

Sequences:

H	H	H
H	H	C
H	C	H
H	C	C
C	H	H
C	H	C
C	C	H
C	C	C

Total no of combinations

∴ We have to find the sequence with the highest probability

$$\therefore P(1,3,1,H,C,H) = P(1|H) * P(3|C) * P(1|H) * P(H) * P(C|H) * P(H|C)$$

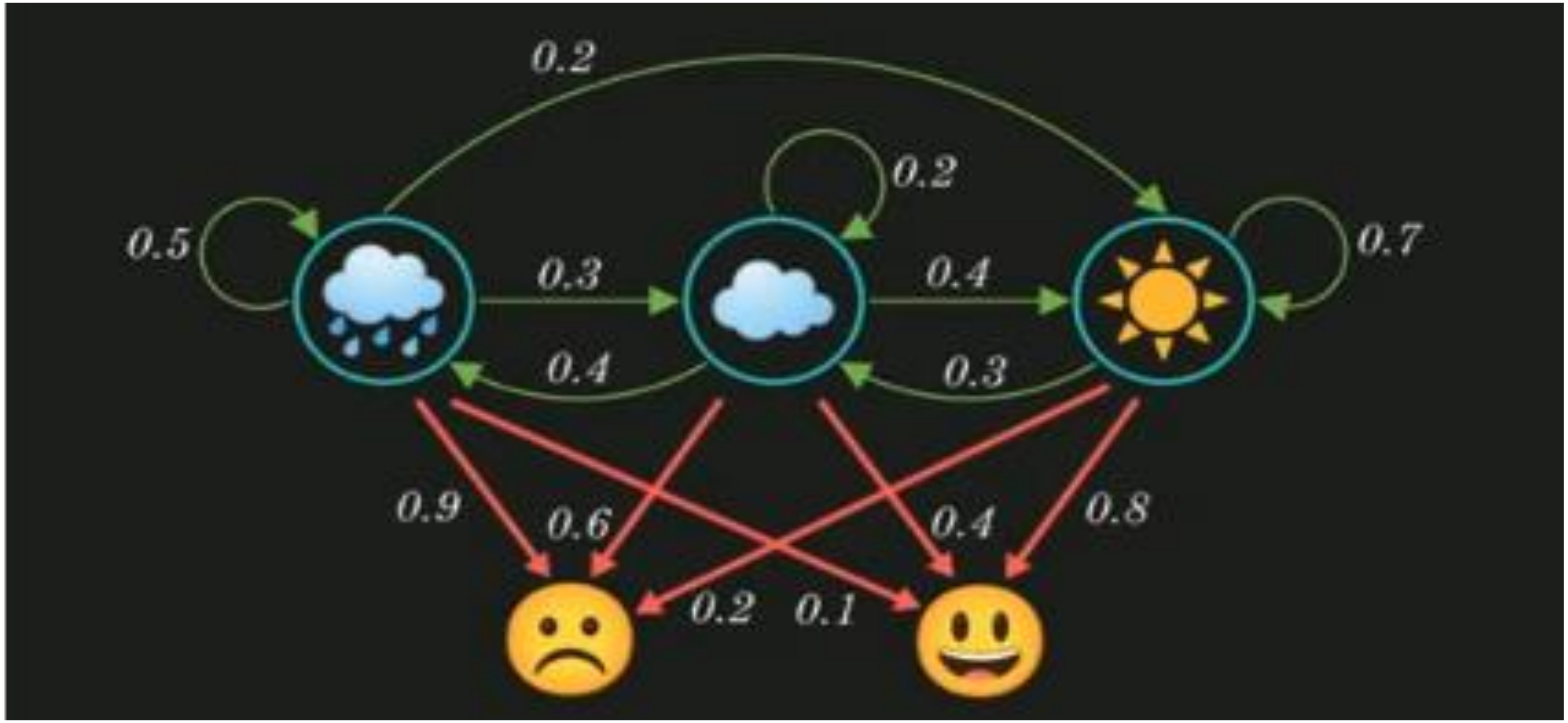
$$= 0.2 * 0.1 * 0.2 * 0.8 * 0.3 * 0.4$$

$$= 3.84 \times 10^{-4}$$

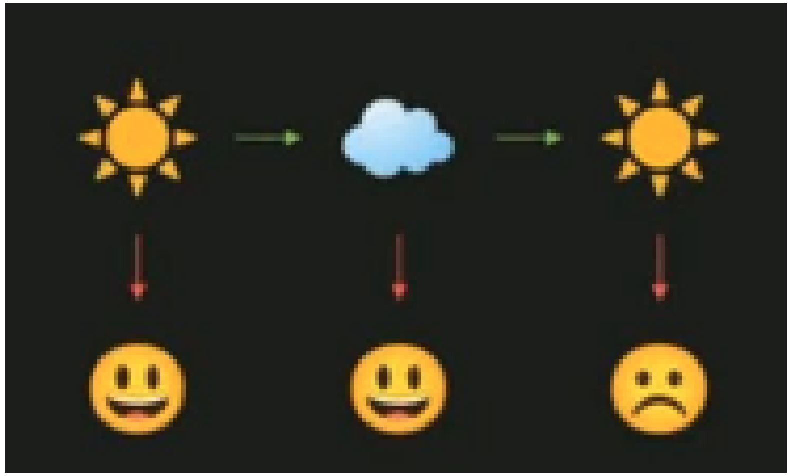
∴ like this calculate for the rest of the sequence and take the highest one.

Question 2

- Suppose, the state of emotion(Happy or Sad) is dependent on the weather(Sunny, Cloudy, Rain)



- Now if probability of weather being Sunny is 50.90% then what is the probability of occurrence of the following series of event?



- Given P(S) = 50.90%, then what is the joint probability of the observed mood given the weather sequence below?
- P(Y=H,H,S , X=S,C,S)

Therefore,

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

Need for Viterbi Algorithm

- It is based on dynamic programming.
- Dynamic programming breakdowns the problems into sub-problems.
- It saves the result for future purposes therefore don’t need to compute the result again.

Viterbi Algorithm Steps

- **Step 1: Initialization**
  - Start at the first observation
  - For each possible state, calculate its probability
  - Remember which state you're in
- **Step 2: Forward Pass (Moving Through Time)**
  - Move to the next observation
  - For each possible state at this time:
    - Look at all states from the previous time
    - Find which previous state gives the highest probability of reaching the current state
    - Multiply by the probability of producing the current observation
    - Keep track of the best previous state (this is your "breadcrumb")

Viterbi Algorithm Steps

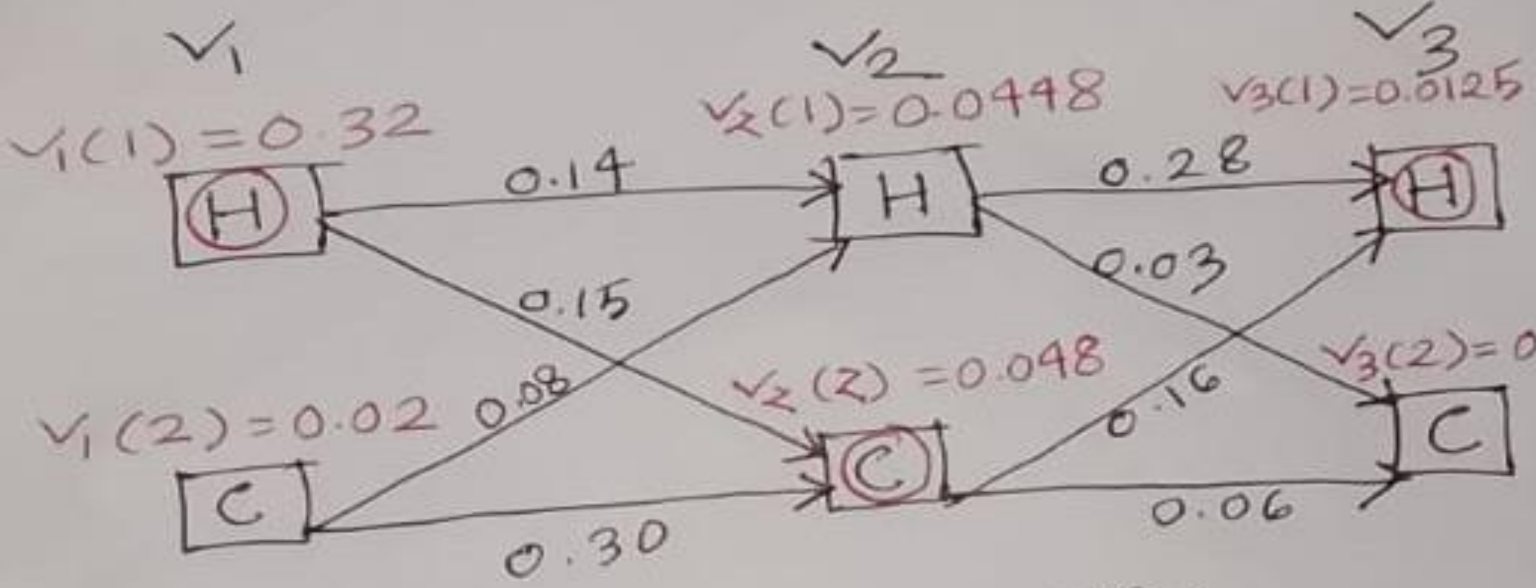
- **Step 3: Repeat**
  - Repeat Step 2 for all remaining observations
  - At each time step, always keep only the best path to each state
  - Continue until you reach the last observation
- **Step 4: Find the Best Final State**
  - Look at all possible states at the final time
  - Choose the one with the highest probability
  - This is the end of your best path

Viterbi Algorithm Steps

- **Step 5: Backtrack to Find Full Path**
  - Start from the best final state you just found
  - Follow the breadcrumbs backward
  - At each step, look at which previous state you marked as "best"
  - Continue until you reach the beginning
  - This gives you the complete most-likely sequence of states

$0.14 \times 0.32 = 0.0448$

$0.08 \times 0.02 = 0.0016$



$0.0448 \times 0.28 = 0.0125$

$0.048 \times 0.16 = 0.0077$

$0.0448 \times 0.03 = 0.0013$

$0.048 \times 0.06 = 0.0029$

$.15 \times .32 = 0.048$

$.02 \times .30 = 0.006$

3 1 3

$\therefore P(3,H) = P(3|H)\,P(H) = 0.4 \times 0.8 = 0.32$

$\therefore P(3,C) = P(3|C)\,P(C) = 0.1 \times 0.2 = 0.02$

$\therefore P(1,H) = P(1|H)\,P(H|H) = 0.2 \times 0.7 = 0.14$

$\therefore P(1,C) = P(1|C)\,P(C|H) = 0.5 \times 0.3 = 0.15$

$\therefore P(1,H) = P(1|H)\,P(H|C) = 0.2 \times 0.4 = 0.08$

$\therefore P(1,C) = \cancel{P(1|H)}\,P(1|C)\,P(C|C) = 0.5 \times 0.6$

$= 0.30$

$\therefore \cancel{P(3|H)}$

$\therefore P(3,H) = P(3|H)\,P(H|H) = 0.4 \times 0.7 = 0.28$

$\therefore P(3,C) = P(3|C)\,P(C|H) = 0.1 \times 0.3 = 0.03$

$\therefore P(3,H) = P(3|H)\,P(H|C) = 0.4 \times 0.4 = 0.16$

$\therefore P(3,C) = P(3|C)\,P(C|C) = 0.1 \times 0.6 = 0.06$

So, the ans is HCH

Computational Complexity

Brute Force: $O(N^T) = O(2^3) = 8$

Viterbi: $O(N^2 \times T) = O(2^2 \times 3) = 12$

N = 2 Hidden States, T = 3 Observations

Brute Force: $O(N^T) = O(2^{100}) = \boxed{1.26 \times 10^{30}}$

Viterbi: $O(N^2 \times T) = O(2^2 \times 100) = 400$

HUGE! Number!

N =2 Hidden States, T = 100 Observations