DEEP LEARNING

HIDDEN MARKOV MODEL AND MARKOV CHAIN AND NETWORK MODEL

Session-31

CO - 4

AIM

To familiarize students with the concepts **HIDDEN MARKOV MODELS**

INSTRUCTIONAL OBJECTIVES

This Session is designed to:

1. Discussion on hidden Markov model
2. Demonstrate the hmm examples

LEARNING OUTCOMES

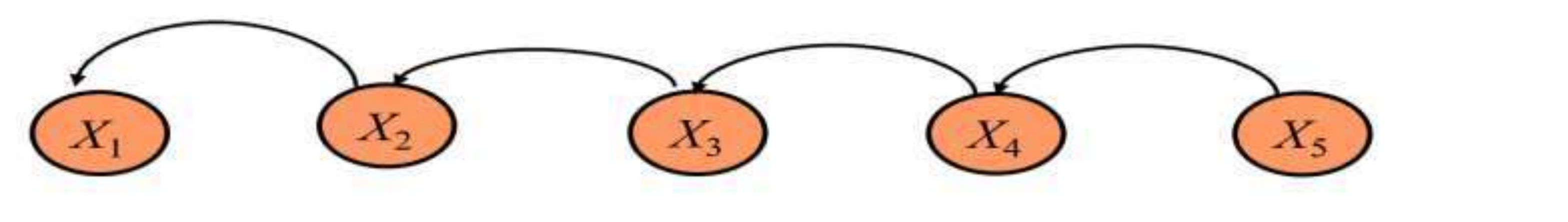
At the end of this session, you should be able to:

1. Able to apply hidden Markov model

MARKOV-CHAIN(MARKOV 1914) HAS BEEN APPLIED

- To short term market forecasting and business decision,
- On the future market the firms‘ future market share, given a consumer transition from one firm to the next.
- Market research problems (market share predictions)
- Markov text generators
- Asset pricing and other financial predictions
- Customer journey predictions
- Population genetics
- Algorithmic music composition
- Page ranks (google results)

MARKOV-CHAIN



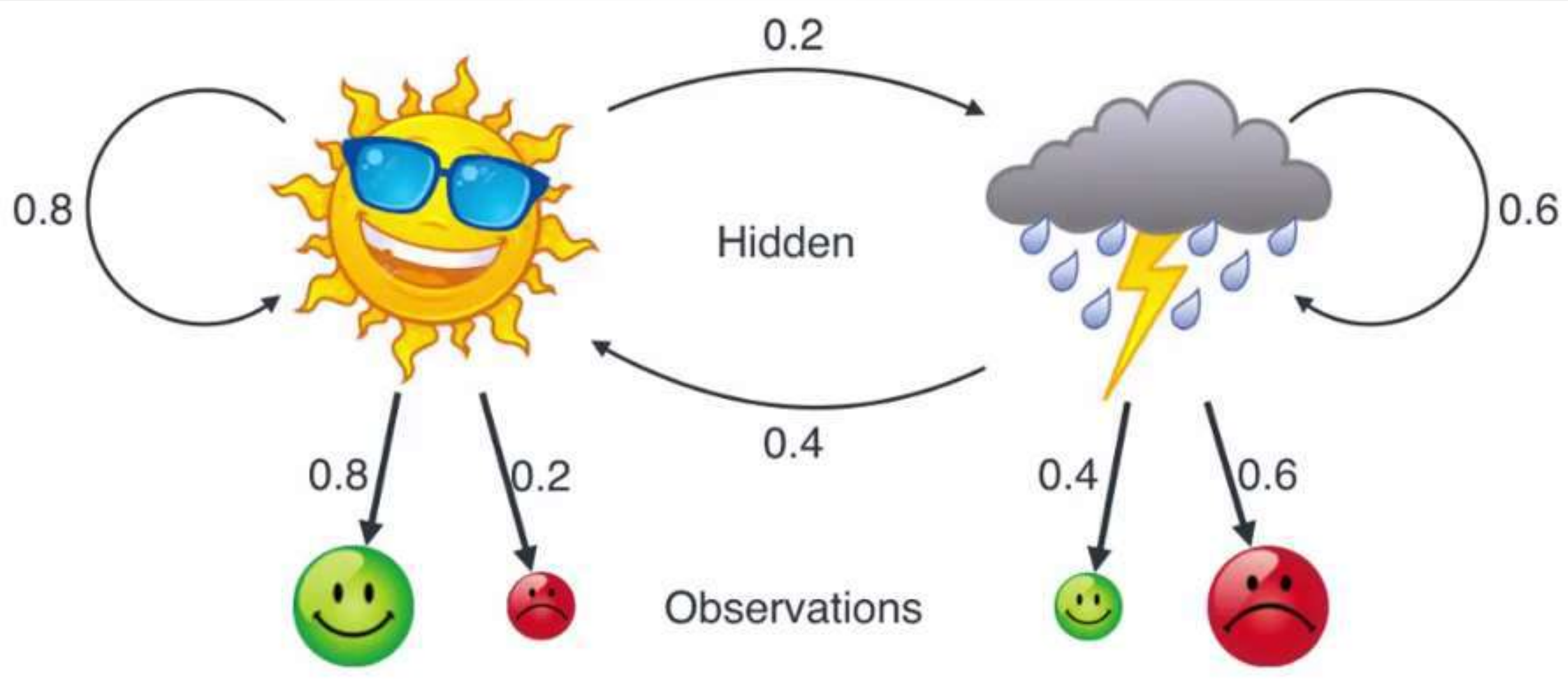
States: $\{X_1,X_2,X_3,X_4, X_5\}$,if x_3 is t be current state, then x_2 =t-1,x_1 =t-2,x_4 = t+1,x_5 = t+2

- Any Chain that follow Markov propery (i.e)$P(X_t \mid X_{t-1})$ then it is know as Markov chain.

HIDDEN MARKOV MODELS

- **Hidden Markov Models** (HMMs) are a class of probabilistic graphical **model** that allow us to predict a sequence of unknown (**hidden**) variables from a set of observed variables
- .A simple **example** of an **HMM** is predicting the weather (**hidden** variable) based on the type of clothes that someone wears (observed).
- A hidden Markov model (HMM) is a statistical approach that is frequently used for modelling biological sequences.
- in applying it, a sequence is modelled as an output of a discrete stochastic process, which progresses through a series of states that are 'hidden' from the observer.

HIDDEN MARKOV MODEL



CONTD..

- In order to compute the joint probability of a sequence of hidden states, we need to assemble three types of information.
- Generally, the term "states" are used to refer to the hidden states and "observations" are used to refer to the observed states.
- **Transition data** — the probability of transitioning to a new state conditioned on a present state.
- **Emission data** — the probability of transitioning to an observed state conditioned on a hidden state.
- **Initial state information** — the initial probability of transitioning to a hidden state. This can also be looked at as the prior probability.

HIDDEN MARKOV MODEL

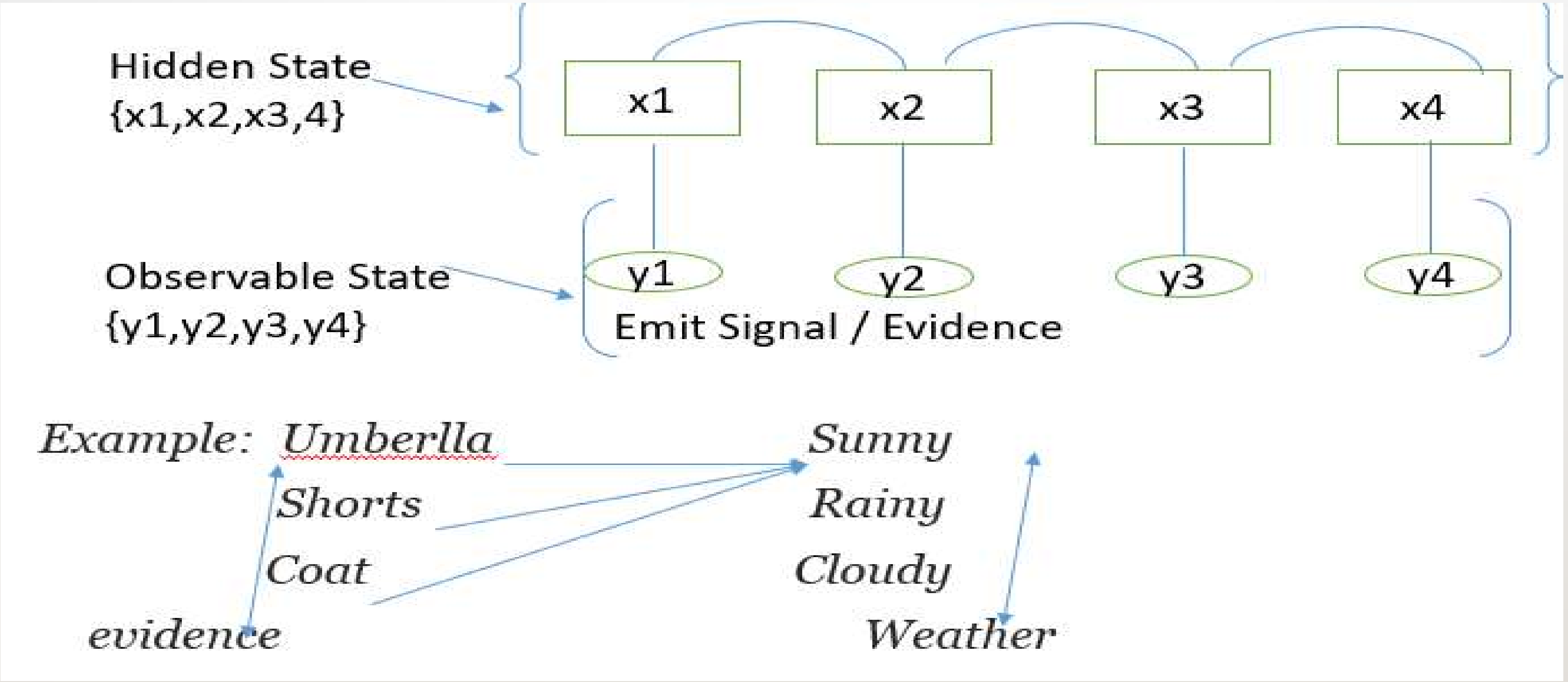
A Markov model consists of five elements:

1. A finite set of states $\Omega = \{s_1, \ldots, s_k\}$.
2. A finite signal alphabet $\Sigma = \{\sigma_1, \ldots, \sigma_m\}$.
3. Initial probabilities $P(s)$ (for every $s \in \Omega$) defining the probability of starting in state s.
4. Transition probabilities $P(s_i \mid s_j)$ (for every $(s_i, s_j) \in \Omega^2$) defining the probability of going from state s_j to state s_i.
5. Emission probabilities $P(\sigma \mid s)$ (for every $(\sigma, s) \in \Sigma \times \Omega$) defining the probability of emitting symbol σ in state s.

EMISSION PROBABILITY MATRIX:

- Probability of hidden state generating output v_i given that state at the corresponding time was s_j.

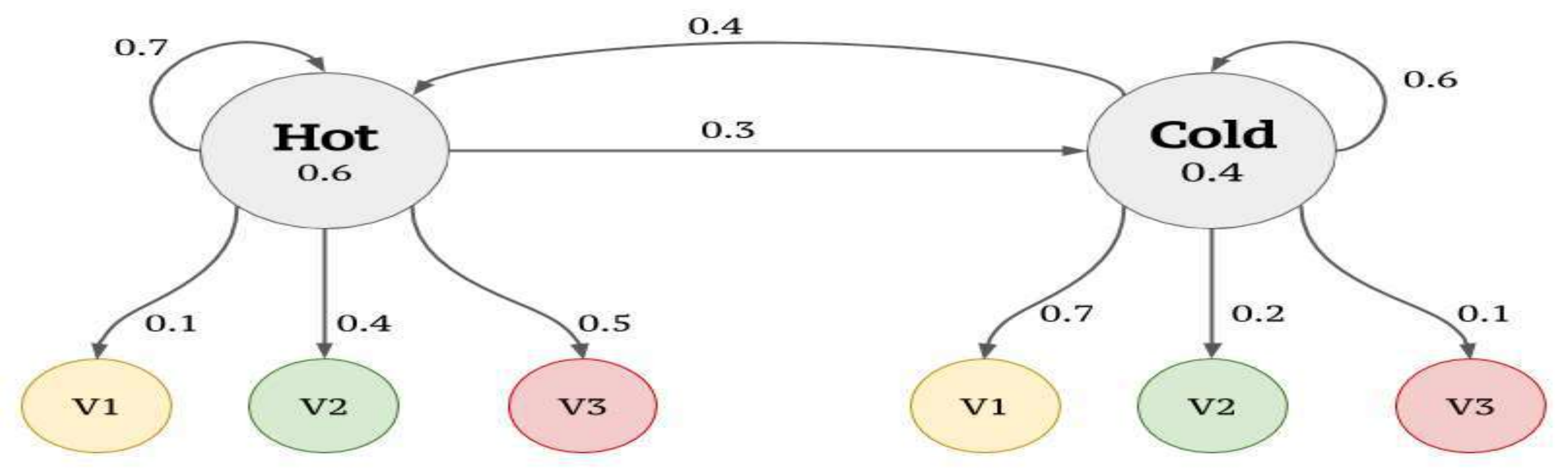
EXAMPLE



Situation:
If you are sitting in a room and unable to see the weather, based on the below evidence, you conclude the weather condition.

EXAMPLE OF HMM

- Consider the two given states Hot(H), Cold(C) are shown below, How likely is the sequence {3,1,3}?



SOLUTION

Given

- S = {hot , cold}
- v = {v1=1 ice cream ,

  v2=2 ice cream,

  v3=3 ice cream}

where V is the Number

of ice creams consumed on a day.

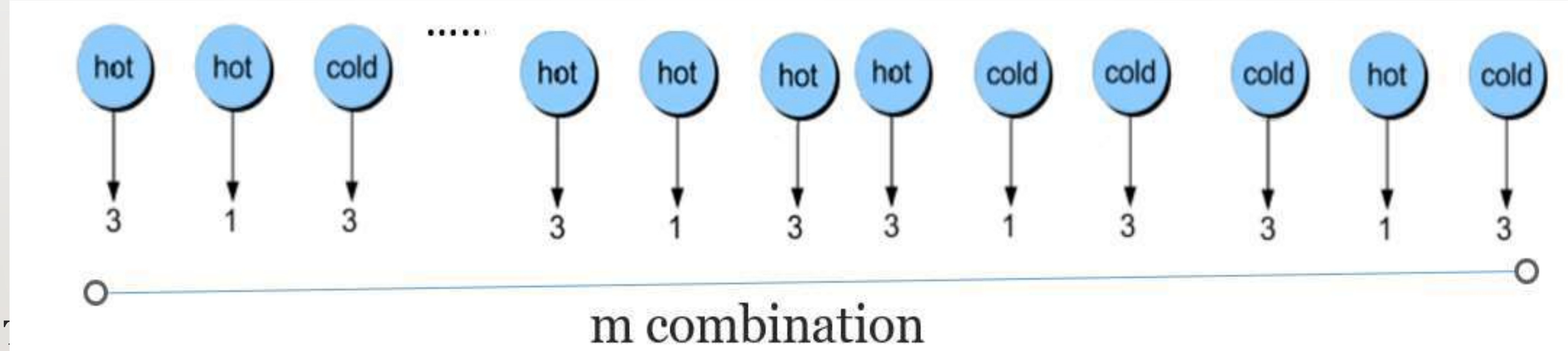
$$
\underset{\text{(State Transition Matrix)}}{A} = \begin{array}{c c} & \begin{array}{cc} H & C \end{array} \\ \begin{array}{c} H \\ C \end{array} & \left\{ \begin{array}{cc} 0.7 & 0.3 \\ 0.4 & 0.6 \end{array} \right\} \end{array}
$$

$$
\underset{\text{(Emission Matrix)}}{B} = \begin{array}{c c} & \begin{array}{ccc} V1 & V2 & V3 \end{array} \\ \begin{array}{c} H \\ C \end{array} & \left\{ \begin{array}{ccc} 0.1 & 0.4 & 0.5 \\ 0.7 & 0.2 & 0.1 \end{array} \right\} \end{array}
$$

$$
\underset{\text{(Initial state S0)}}{\pi} = \begin{array}{c} \begin{array}{cc} H & C \end{array} \\ \left\{ \begin{array}{cc} 0.6 & 0.4 \end{array} \right\} \end{array}
$$

COND..

- Suppose we want find how much John would eat , then the sequence will be



N=No of Hidden State={Hot, Clod}=2

T=No of Observations={3,1,3}=3   Therefore $M=N^T=2^3=8$ combination

COND..

By Joint Probability Distribution:

$P(O,Q)=P(O|Q).P(Q)$

Now consider the fist one in first seq i.e

Cold(C), Hot(H); $P(3,H)=P(3|H).P(H)$

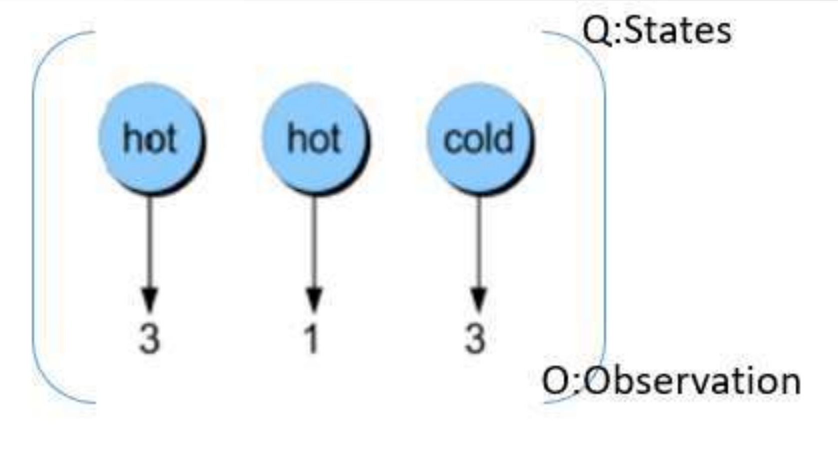
Then the first seq. becomes

$P(3,1,3\ ,\ H,H,C)=$

$P(3|H)\ P(1|H)\ P(3|C)$ *

$P(H)P(H|H)P(H|C)$

[P(H)-initial state]



CONTD..

P(3,1,3 , H,H,C)=

(Emission prob) * (transition Prob)

The formula becomes for HMM:

$$P(O,Q) = P(O|Q) \times P(Q) = \prod_{i=1}^{T} P(o_i|q_i) \times \prod_{i=1}^{T} P(q_i|q_{i-1})$$

Sub the values we get:

(i) P(3,1,3 , H,H,C)=0.063% (ii) P(313,HHH)=0.735%

(iii) P(313,HCH)=**1.26%** (IV)P(313,CHH)=0.042%

(V) P(313,CCC)=.1008% (VI)P(313,CCH)=0.2%

(VII) P(313,CHC)=0.24% (VIII) P(313,HCC)=0.5%

PROBLEM

- Find the max of all probabilities

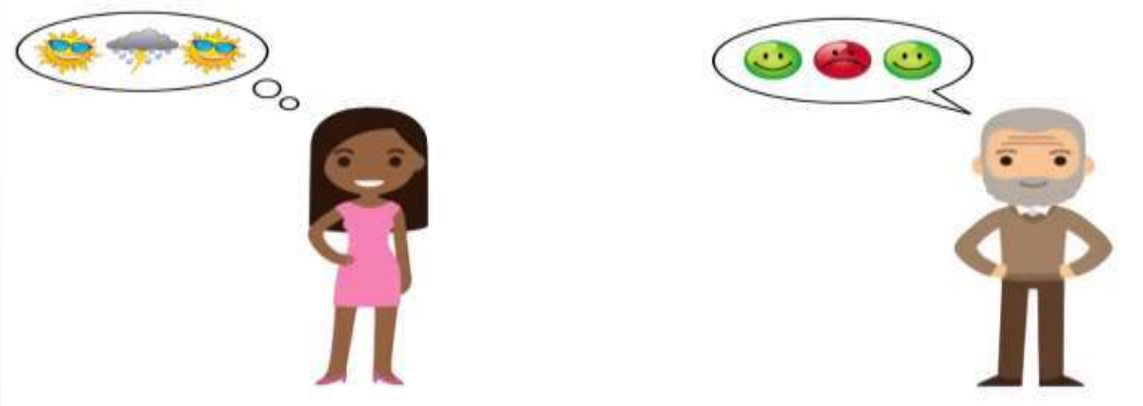
Max(P1,P2,P3,P4,P5,P6,P7,P8)

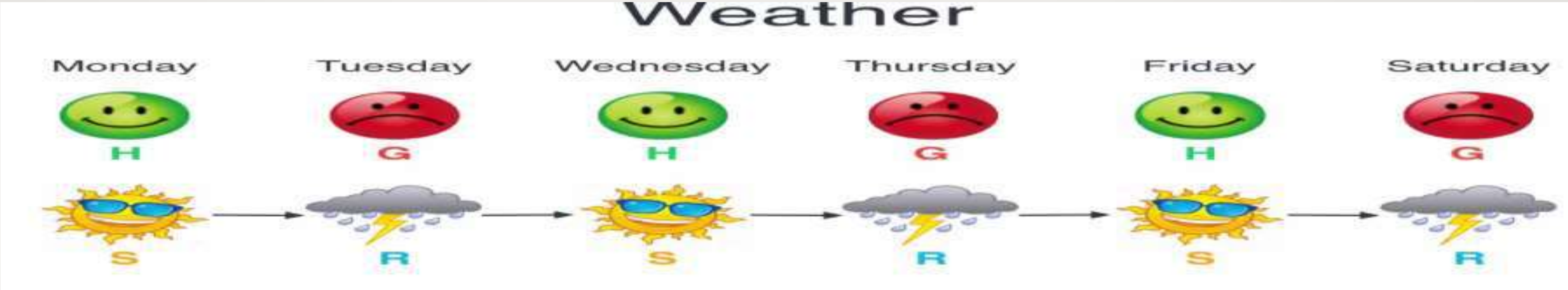
From the above example, We have P(313,HCH)=1.26% is maximum.

Thus we conclude that Hot,Cold Hot is maximum by using HMM

CONTD..


Weather
0.8
0.2
0.6
0.4

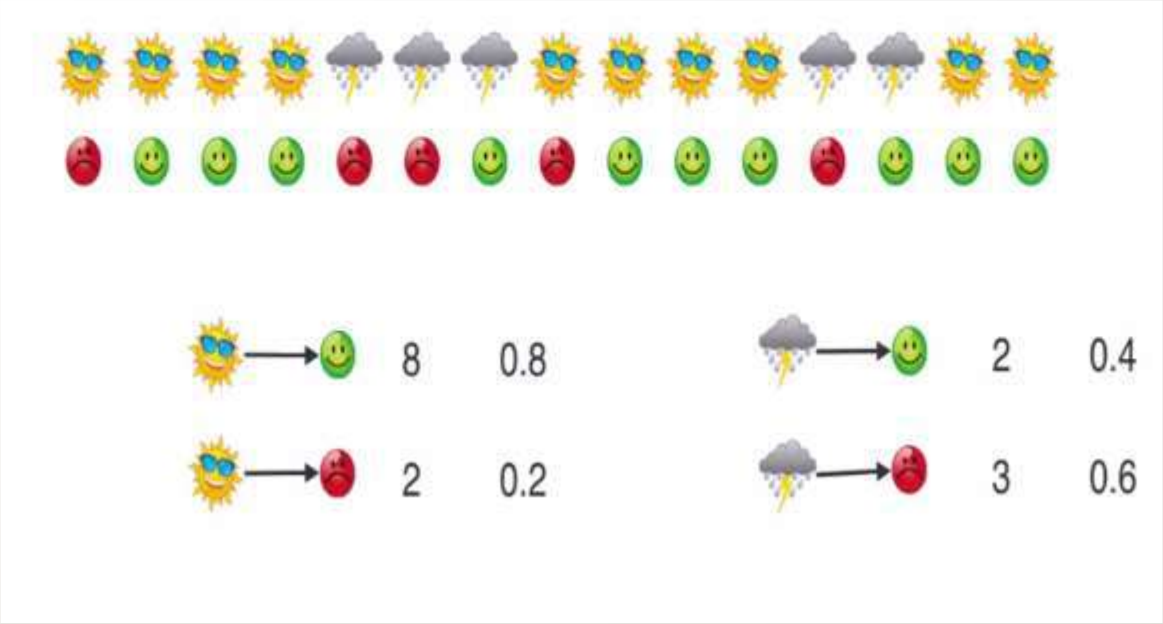



Weather
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
H
G
H
G
H
G
S
R
S
R
S
R

QUESTIONS ON PROBLEM

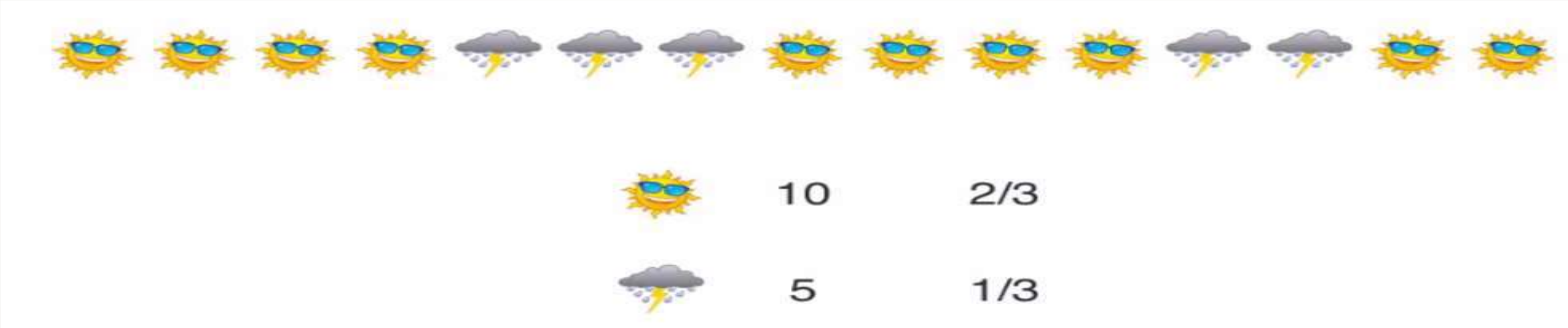
- 1.What is the probability that a random day is Sunny or Rainy?
- 2. If Bob is Happy today .What's the probability that its Sunny or Rainy?
- If for three days Bob is Happy, Grumpy, Happy, What was the weather?

HOW DID WE FIND THE PROBABILITIES



PROBABILITY THAT A DAY IS RAINY OR SUNNY:

- Probability of hidden state generating output v_i given that state at the corresponding time was s_j.



Using bayes theorem, sunny probability is 10/15 = 2/3
rainy probability is 5/15 = 1/3

MARKOV ASSUMPTIONS

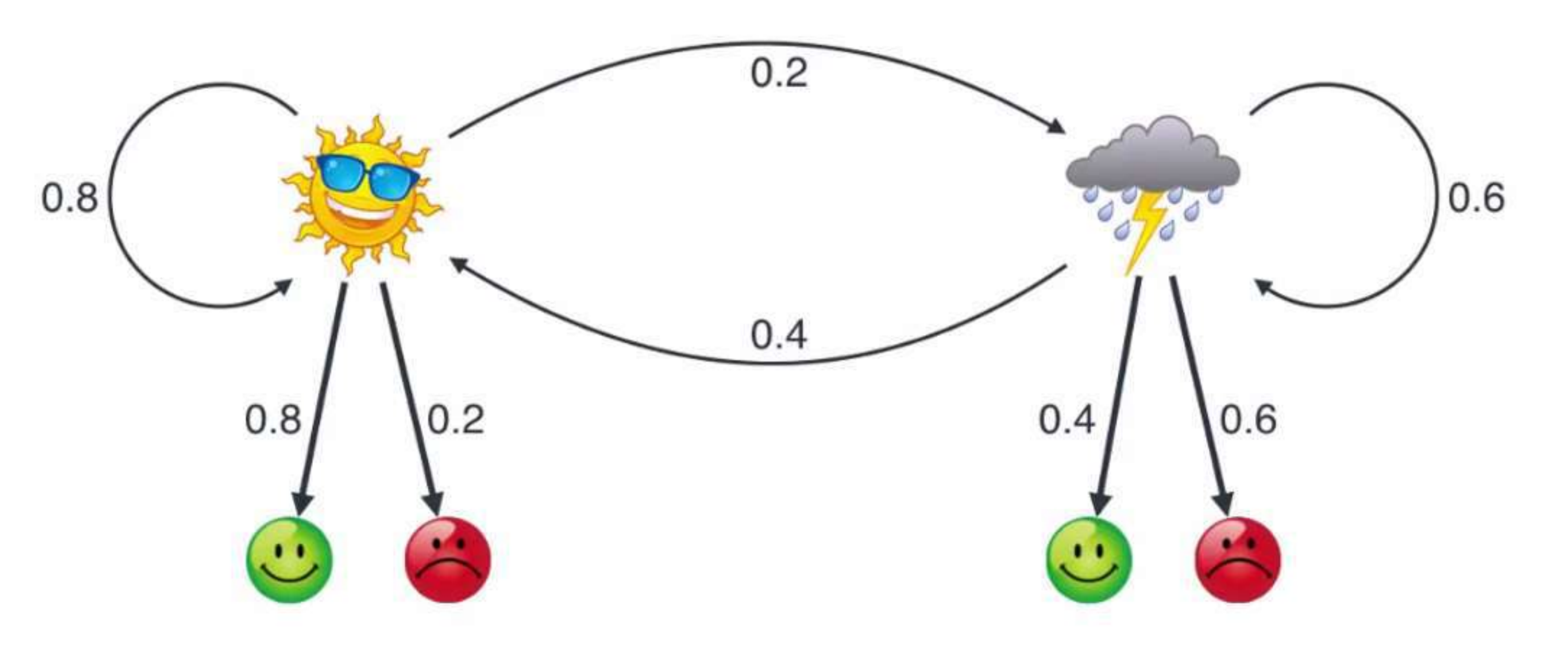
- State transitions are assumed to be independent of everything except the current state:

$$P(s_1, \ldots, s_n) = P(s_1) \prod_{i=1}^{n-1} P(s_{i+1} \mid s_i)$$

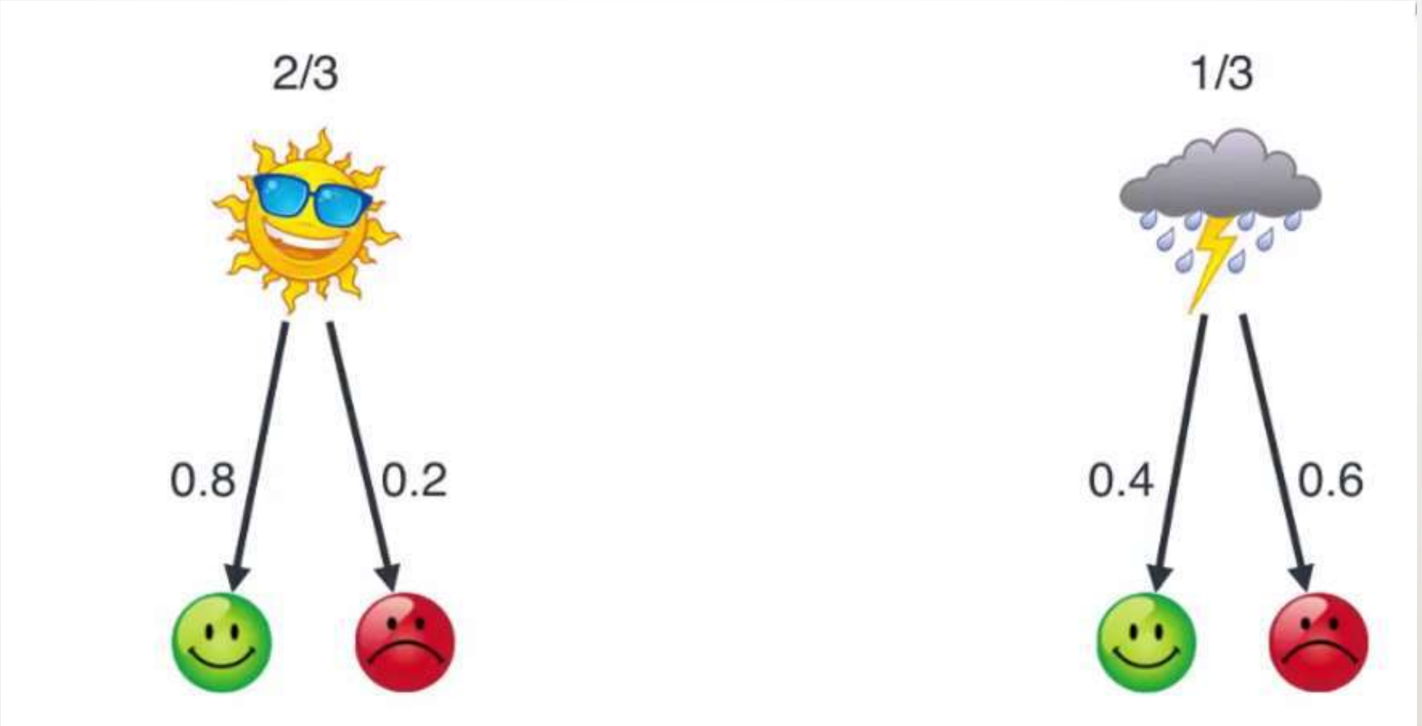
- Signal emissions are assumed to be independent of everything except the current state:

$$P(s_1, \ldots, s_n, \sigma_1, \ldots, \sigma_n) = P(s_1, \ldots, s_n) \prod_{i=1}^{n} P(\sigma_i \mid s_i)$$

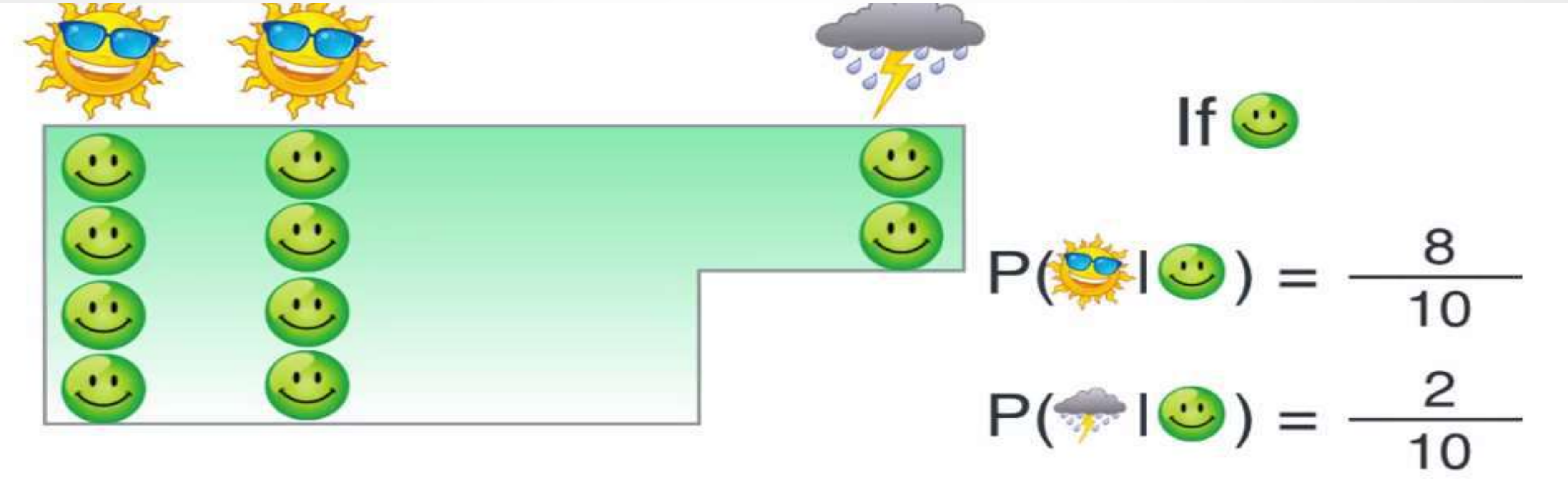
EMISSION PROBABILITY MATRIX:



BAYES THEOREM

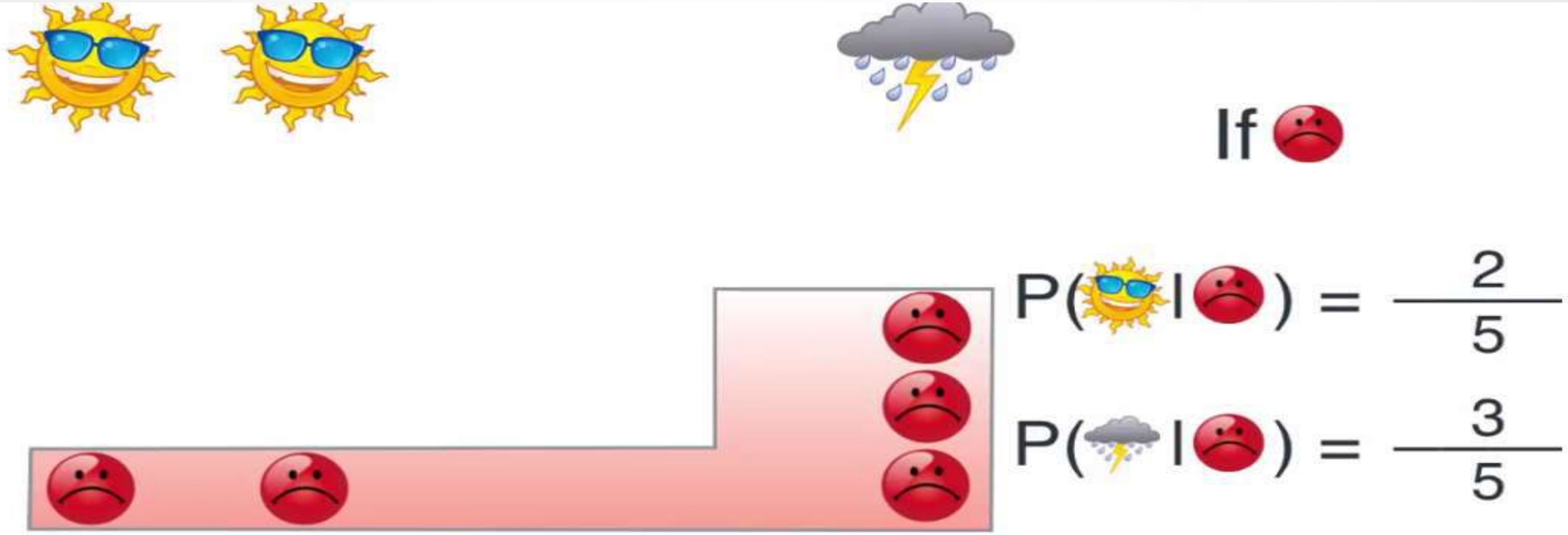


BAYES THEOREM



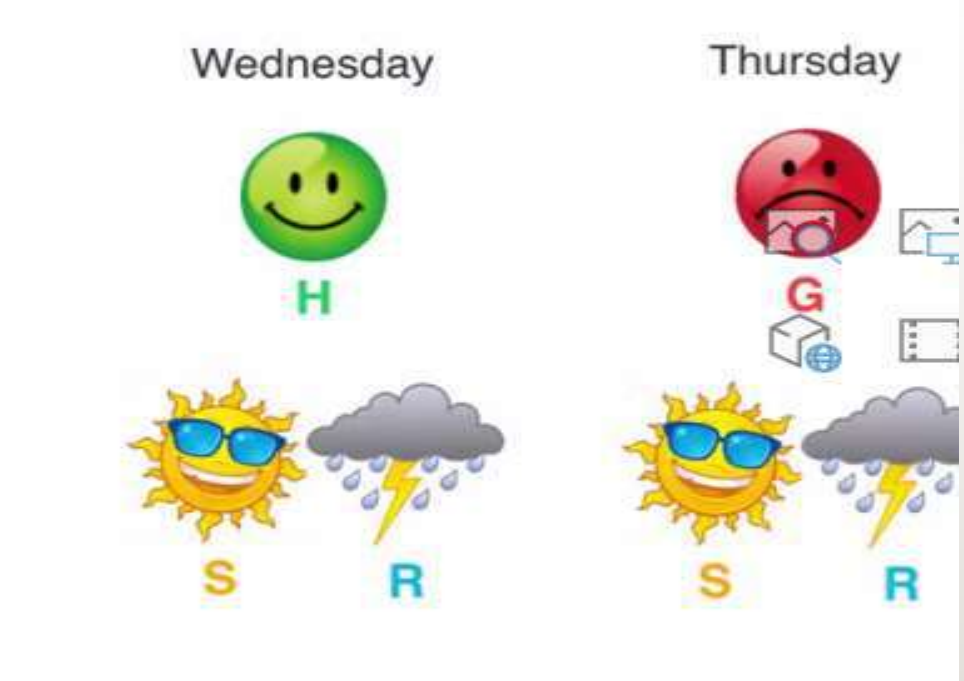
Probability of sunny, given that bob is happy =8/10
Probability of rainy given that bob is happy =2/10

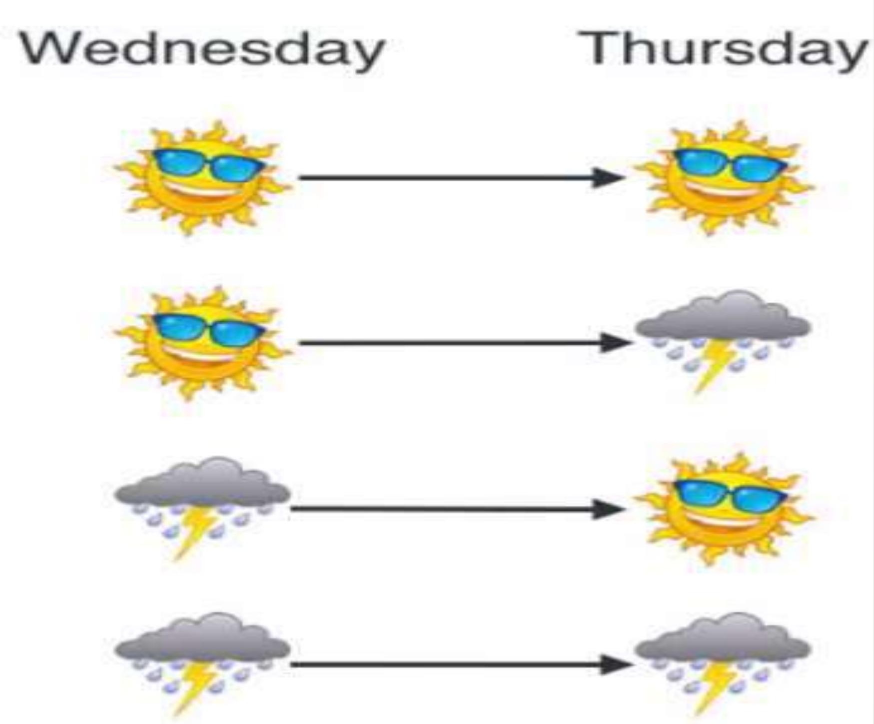
CONTD..



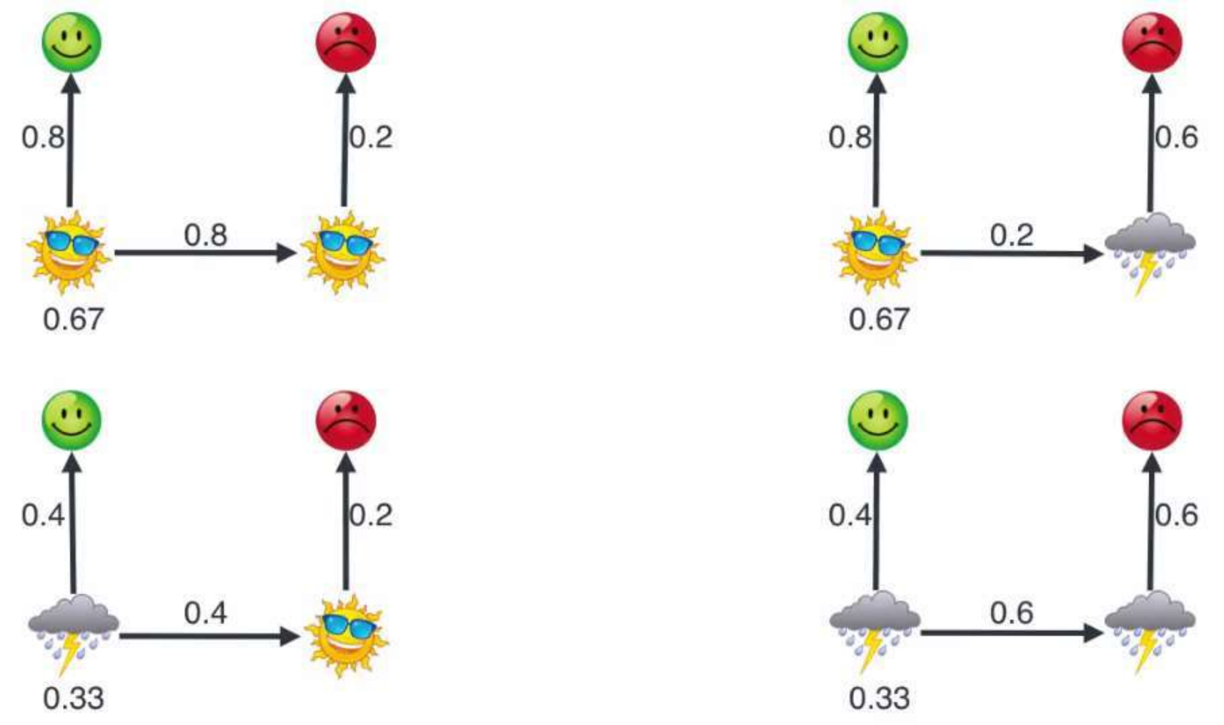
Probability of sunny, given that bob is sad =2/5
Probability of rainy given that bob is sad =3/5

EXAMPLE: IF HAPPY-GRUMPY, WHAT IS THE WEATHER.

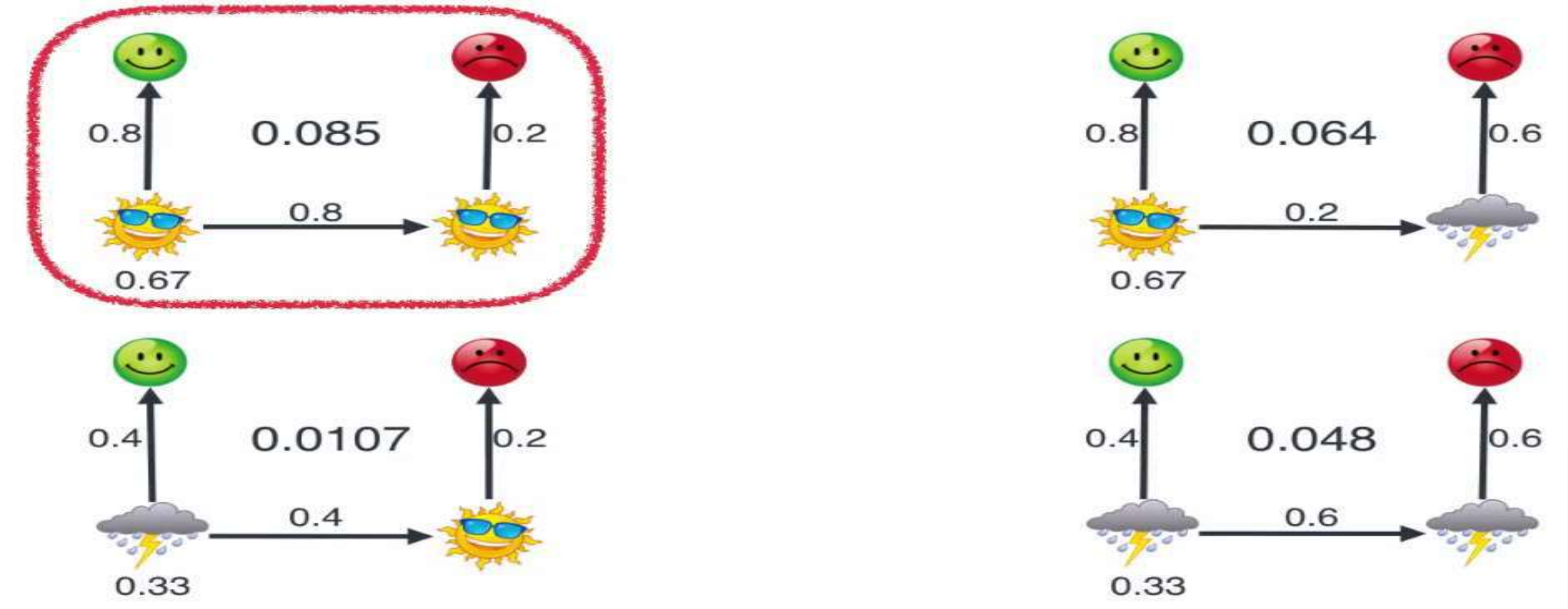




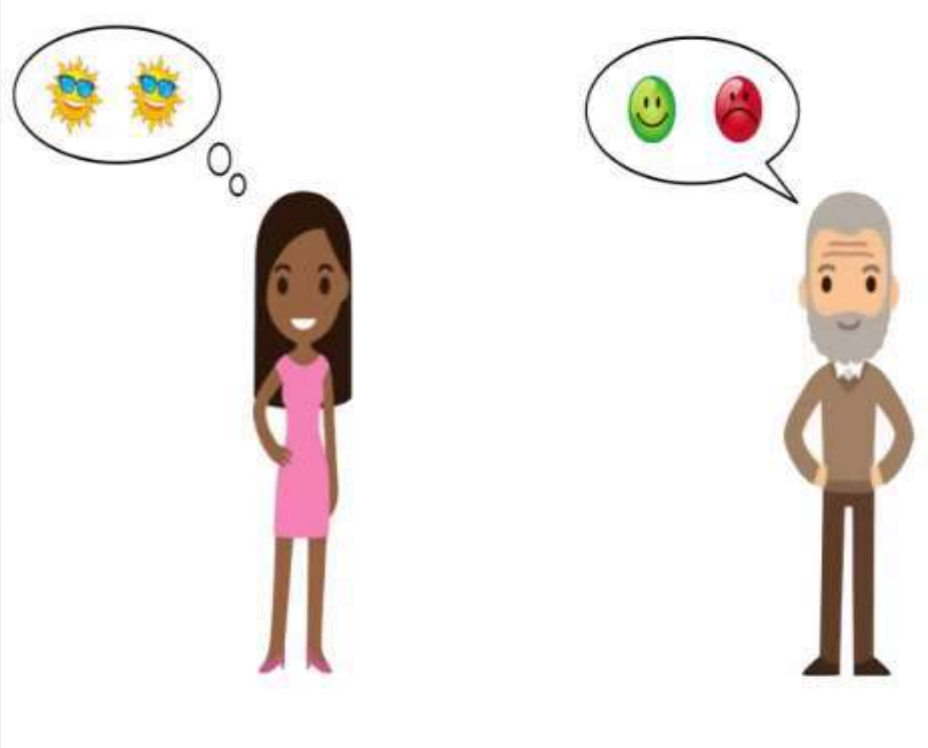
ALL POSSIBLE SEQUENCES

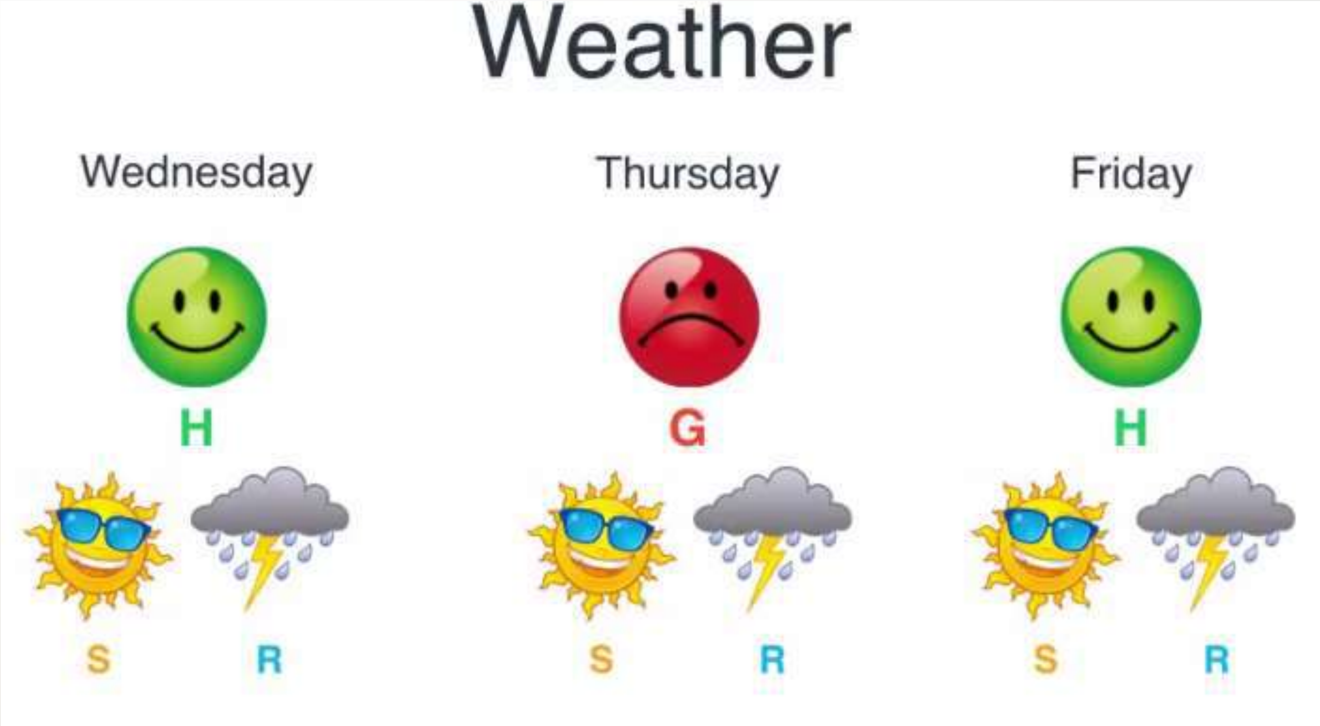


ALL POSSIBLE SEQUENCES

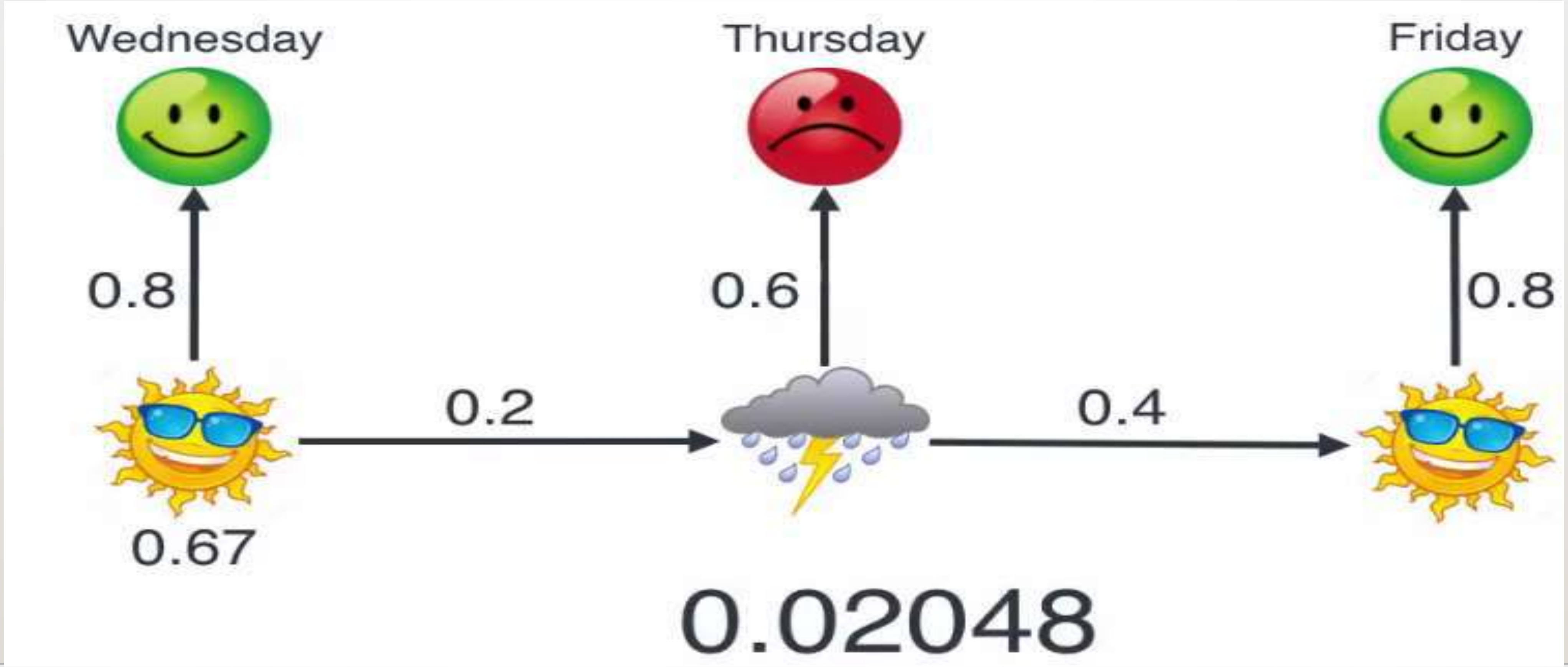


ALL POSSIBLE SEQUENCES





ALL POSSIBLE SEQUENCES



EMISSION PROBABILITY MATRIX:

- Probability of hidden state generating output v_i given that state at the corresponding time was s_j.

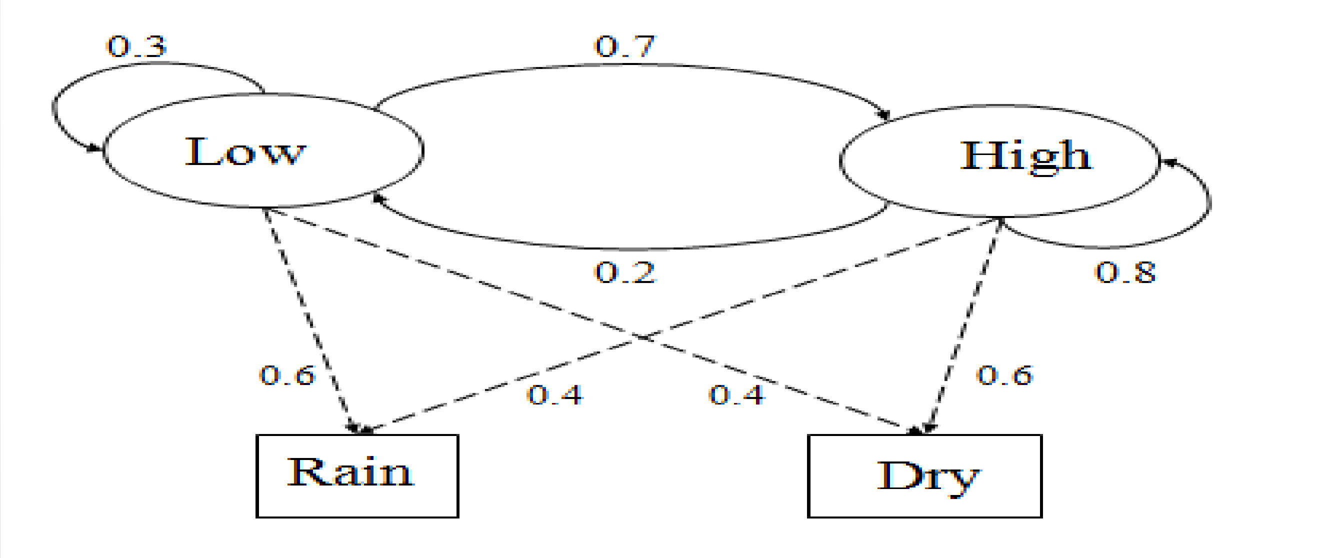
Self-Assessment Questions

1. **Where does Markov models used**
   a. speech recognization
   b. understanding real world
   c. none mentioned

2. Which algorithm works by first running the standard forward pass to compute?
a) Smoothing
b) Modified smoothing
c) HMM
d) DFS algorithm

TERMINAL QUESTIONS

1.Consider the two given states Low, High and two given observations Rain and Dry. The initial probabilities for Low and High is (0.4 and 0.6). Find the probability of a sequence of observations, i.e., {Dry, Rain}



REFERENCES FOR FURTHER LEARNING OF THE SESSION

Books:

1 Ian Goodfellow and Yoshua Bengio and Aaron Courville (2016) Deep Learning Book

Resources:

https://towardsdatascience.com/hidden-markov-models-simply-explained-d7b4a4494c50

THANK YOU

Team -Deep Learning