IT5005 Artificial Intelligence

Sirigina Rajendra Prasad
AY2025/2026: Semester 1

Probabilistic Reasoning over Time
(Markov Chains and HMM)

1. Identify Valid Probability Distributions

- $\boldsymbol{P}(X) = \begin{bmatrix} 0.1 \\ 0.3 \\ 0.6 \end{bmatrix}$
- $\boldsymbol{P}(X) = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.5 \end{bmatrix}$
- $\boldsymbol{P}(X) = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}$

1. Identify Valid Probability Distributions

- $\boldsymbol{P}(X) = \begin{bmatrix} 0.1 \\ 0.3 \\ 0.6 \end{bmatrix}$ Valid probability distribution

- $\boldsymbol{P}(X) = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.5 \end{bmatrix}$ Not a probability distribution
  Can convert it to probability distribution by normalizing the vector
  Normalization Constant:
  $$\alpha = \frac{1}{0.2 + 0.4 + 0.5}$$

Normalized Vector = $\alpha \begin{bmatrix} 0.2 \\ 0.4 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.1818 \\ 0.3636 \\ 0.4545 \end{bmatrix}$

- $\boldsymbol{P}(X) = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}$ Valid probability distribution

2. Identify Valid Transition Matrices

- $T_1 = \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.3 & 0.2 & 0.5 \\ 0.1 & 0.7 & 0.2 \end{bmatrix}$

- $T_2 = \begin{bmatrix} 0.1 & 0.2 & 0.8 \\ 0.1 & 0.2 & 0.7 \\ 0.0 & 0.7 & 0.3 \end{bmatrix}$

2. Identify Valid Transition Matrices

- $T_1 =$

	1	2	3
1	0.1	0.2	0.7
2	0.3	0.2	0.5
3	0.1	0.7	0.2

$\boldsymbol{P}(X_t|X_{t-1}=1)$

Valid transition matrix, because

$$\sum_{x_t=1}^{3} \boldsymbol{P}(X_t = x_t|X_{t-1} = c) = 1, \forall\ c \in \{1,2,3\}$$

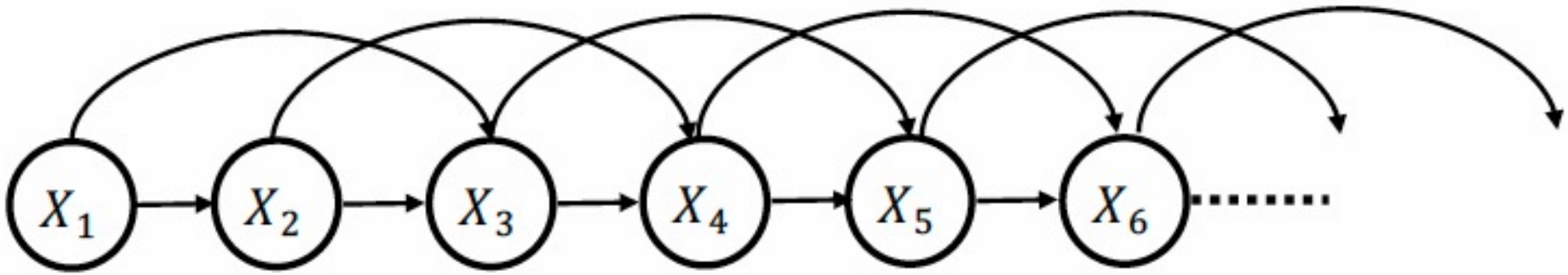
- $T_2 =$

	1	2	3
1	0.1	0.2	0.8
2	0.1	0.2	0.7
3	0.0	0.7	0.3

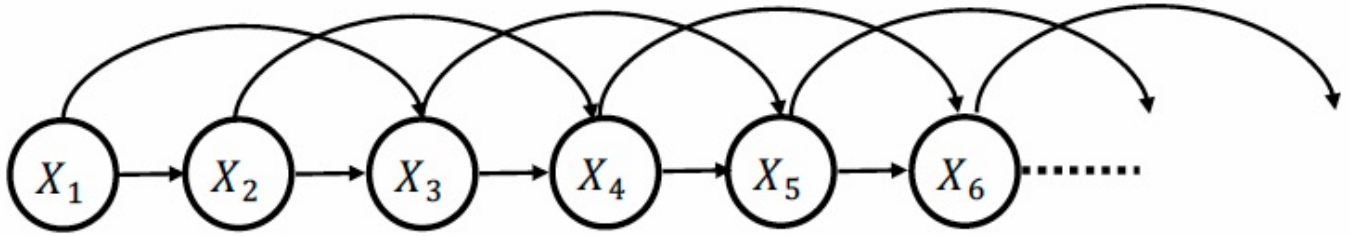
Not a valid transition matrix as the sum of the elements in the first row is not equal to 1

$$\sum_{x_t=1}^{3} \boldsymbol{P}(X_t = x_t|X_{t-1} = 1) \neq 1$$

3. Factorization of Markov Chain



Factorization of Markov Chain



- $P(X_1, X_2, \dots, X_n) = P(X_n|X_1, X_2, \dots, X_{n-1})P(X_1, X_2, \dots, X_{n-1})$

$$= P(X_n|X_{n-1}, X_{n-2})P(X_{n-1}|X_1, X_2, \dots, X_{n-2}) \\ P(X_1, X_2, \dots, X_{n-2})$$

$$= P(X_n|X_{n-1}, X_{n-2})P(X_{n-1}|X_{n-2}, X_{n-3}) \\ P(X_{n-2}|X_1, X_2, \dots, X_{n-3})P(X_1, X_2, \dots, X_{n-3})$$

$$= P(X_n|X_{n-1}, X_{n-2})P(X_{n-1}|X_{n-2}, X_{n-3}) \dots P(X_3|X_2, X_1) \\ P(X_2|X_1)P(X_1)$$

$$\vdots$$

$$= P(X_1)P(X_2|X_1)\prod_{i=3}^{i=n} P(X_i|X_{i-1}, X_{i-2})$$

4. Let X be the random variable that represent the stock price of a company. X can take on one of two possible values:

$$High(H), Low(L)$$

on any given day. Further, assume that X follows a first order Markov model that is shown in Figure 2.

Also, note that that the stock price of the company on the first day can be any of the two values with equal probability. Given these information, answer the questions below:

(a) Write the state transition model (transition matrix) and prior distribution.

(b) The sequence of stock prices of the company for ten days is

$$L \rightarrow L \rightarrow L \rightarrow H \rightarrow L \rightarrow H \rightarrow H \rightarrow H \rightarrow H \rightarrow L$$

Given the sequence, find the probability of stock price being high on day 11.

(c) For the above sequence, calculate the probability of stock prices for the first five days?

(d) Draw the trellis of the Markov model for three time steps and calculate the probability distribution of the stock price being high on day 3 i.e., $P(X_3 = H)$?

(e) Find the probability distribution of the stock price on Day 4.

(f) What is the stationary distribution of the stock price i.e., $P(X)$ at time $t \rightarrow \infty$?

(g) Does the stationary distribution depend on the initial probability distribution? Further, assume that you are more of a long-term investor, will you choose to invest in the company?

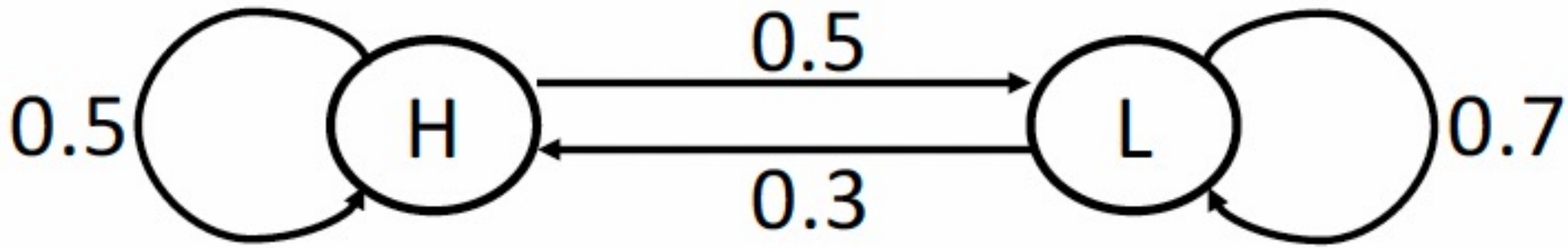
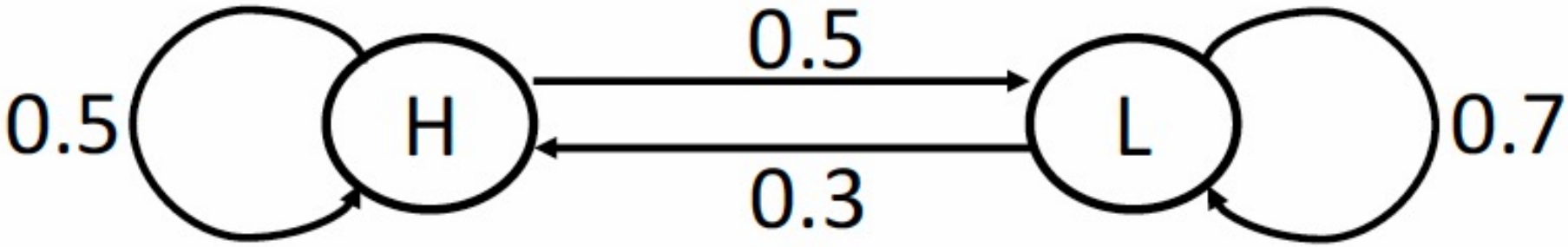


Figure 2: State Transition Model

4a. Write the state transition model and prior distribution



$$\mathbf{T} = \begin{bmatrix} P(X_t = H|X_{t-1} = H) & P(X_t = L|X_{t-1} = H) \\ P(X_t = H|X_{t-1} = L) & P(X_t = L|X_{t-1} = L) \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.3 & 0.7 \end{bmatrix}$$

$$\mathbf{P}(X_1) = \begin{bmatrix} P(X_1 = H) \\ P(X_1 = L) \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

4 b. The sequence of stock price being high on day 11

The sequence of stock price being high on day 11, given the following sequence of stock price:

$$L \to L \to L \to H \to L \to H \to H \to H \to H \to L$$

Solution:

$$P(X_{11} = H | X_1 = L, X_2 = L, \dots, X_{10} = L) = P(X_{11} = H | X_{10} = L)$$

$$= 0.3$$

$$\mathbf{T} = \begin{bmatrix} P(X_t = H | X_{t-1} = H) & P(X_t = L | X_{t-1} = H) \\ P(X_t = H | X_{t-1} = L) & P(X_t = L | X_{t-1} = L) \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.3 & 0.7 \end{bmatrix}$$

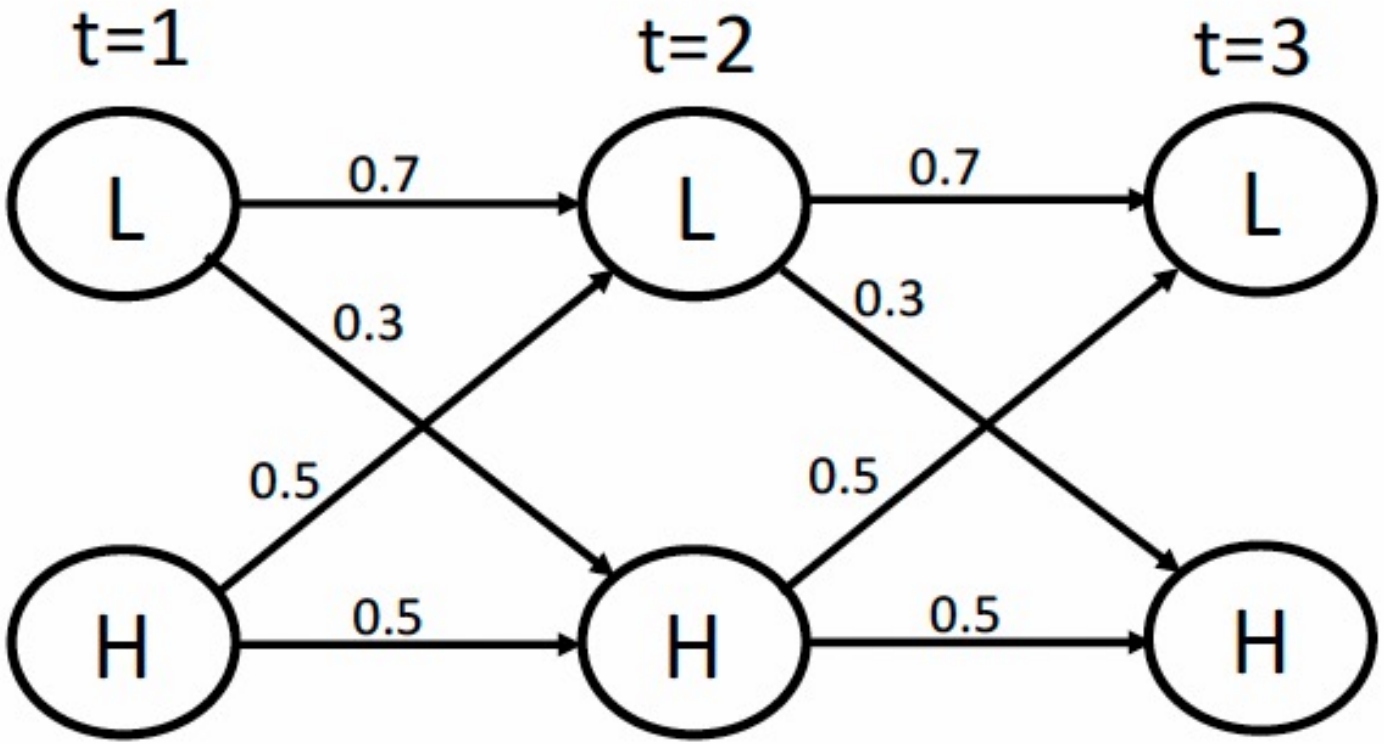
4c. Calculate the probability of the sequence for first five days

$$L \to L \to L \to H \to L \to H \to H \to H \to H \to L$$

$$\begin{aligned} P(X_1 = L, X_2 = L, X_3 = L, X_4 = H, X_5 = L) \\ &= P(X_1 = L)P(X_2 = L|X_1 = L)\, P(X_3 = L|X_2 = L)P(X_4 = H|X_3 = L)P(X_5 = L|X_4 = H) \\ &= 0.5 * 0.7 * 0.7 * 0.3 * 0.5 \\ &= 0.03675 \end{aligned}$$

$$\mathbf{T} = \begin{bmatrix} P(X_t = H|X_{t-1} = H) & P(X_t = L|X_{t-1} = H) \\ P(X_t = H|X_{t-1} = L) & P(X_t = L|X_{t-1} = L) \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.3 & 0.7 \end{bmatrix}$$

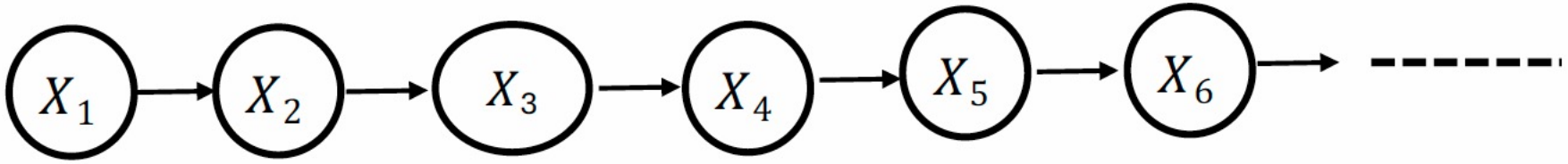
4d. Draw the trellis diagram for three time steps and calculate the probability of stock price being high on day 3, i.e., $P(X_3 = high)$.



4d. Find the probability of stock price being high on day 3

Solution:

Markov Chain:



Given:

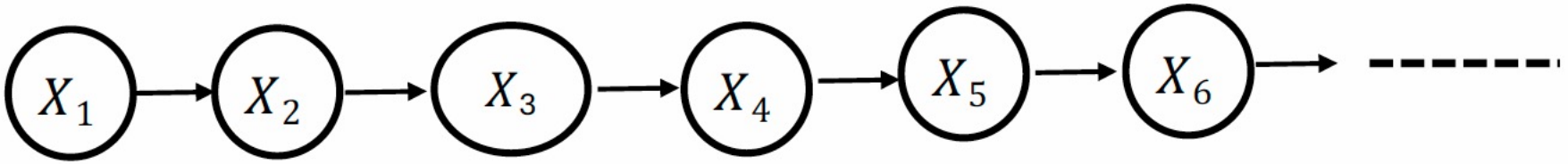
$$\mathbf{P}(X_1) = \begin{bmatrix} P(X_1 = H) \\ P(X_1 = L) \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

Direct Method:

$$\boldsymbol{P}(X_3) = \begin{bmatrix} P(X_3 = high) \\ P(X_3 = low) \end{bmatrix} = (\boldsymbol{T}^T)^2 \boldsymbol{P}(X_1)$$

4d. Find the probability of stock price being high on day 3

- Derivation:
  - Query Variable: X_3
  - Relevant variables: X_1, X_2



$$
\begin{aligned}
P(X_3 = H) &= \sum_{x_1}\sum_{x_2} P(X_1, X_2 X_3 = H) \\
&= \sum_{x_1}\sum_{x_2} P(X_1)P(X_2|X_1)P(X_3 = H|X_2) \\
&= \sum_{x_2} P(X_3 = H|X_2) \sum_{x_1} P(X_2|X_1)P(X_1) \\
&= P(H|H)\left[P(H|H)P(H) + P(H|L)P(L)\right] \\
&\quad + P(H|L)\left[P(L|H)P(H) + P(L|L)P(L)\right] \\
&= 0.5 * [0.5 * 0.5 + 0.3 * 0.5] + 0.3 * [0.5 * 0.5 + 0.7 * 0.5] \\
&= 0.38 \qquad (3)
\end{aligned}
$$

4e. Find the probability distribution of stock price on Day 4

$$\boldsymbol{P}(X_4) = \begin{bmatrix} P(X_4 = high) \\ P(X_4 = low) \end{bmatrix} = (\boldsymbol{T}^T)^3 \boldsymbol{P}(X_1) = \boldsymbol{T}^T \boldsymbol{P}(X_3)$$

$$\begin{aligned} \mathbf{P}(X_4) &= \mathbf{T}^T \mathbf{P}(X_3) \\ &= \begin{bmatrix} 0.5 & 0.3 \\ 0.5 & 0.7 \end{bmatrix} \begin{bmatrix} 0.38 \\ 0.62 \end{bmatrix} \\ &= \begin{bmatrix} 0.376 \\ 0.624 \end{bmatrix} \end{aligned}$$

4f. What is the stationary distribution of the stock price?

- Use the following properties of stationary distribution:
  - $\boldsymbol{P}(X_\infty) = T^T \boldsymbol{P}(X_\infty)$
  - $\sum P(X_\infty) = 1$

$$
\begin{aligned}
P(X_\infty = H) &= P(H|H)P(X_\infty = H) + P(H|L)P(X_\infty = L) \\
P(X_\infty = L) &= P(L|H)P(X_\infty = H) + P(L|L)P(X_\infty = L) \\
P(X_\infty = H) + P(X_\infty = L) &= 1
\end{aligned}
$$

$$
\begin{aligned}
P(X_\infty = H) &= 0.5 * P(X_\infty = H) + 0.3 * P(X_\infty = L) \\
P(X_\infty = L) &= 0.5 * P(X_\infty = H) + 0.7 * P(X_\infty = L)
\end{aligned}
$$

$P(X_\infty = high) = \frac{3}{8}$ and $P(X_\infty = low) = \frac{5}{8}$

4g. Stationary Distribution

- Does the stationary distribution depend on the initial probability distribution? Further, assume that you are more of a long-term investor, will you choose to invest in the company?

- Answer:
  - No, the initial distribution does not influence stationary distribution
  - No, we will choose NOT to invest, because the probability of low stock price is higher than the probability of high stock price.

5. HMM: Sequence Generation

Consider a hidden Markov model with two hidden states $X = 1, 2$, and two possible output symbols $E = A, B$. The probability of starting in states $X = 1$ and $X = 2$ are 0.49 and 0.51, respectively. If X_t is the state at time t, the state transition probabilities are

$$\begin{aligned} P(X_{t+1} = 1 | X_t = 1) &= 1 \\ P(X_{t+1} = 1 | X_t = 2) &= 1 \end{aligned}$$

and the output probabilities are

$$\begin{aligned} P(E_t = A | X_t = 1) &= 1 \\ P(E_t = B | X_t = 2) &= 1 \end{aligned}$$

What is the sequence of three output symbols that has the highest probability of being generated from this HMM model? Explain your answer.

5. HMM Sequence

- Given Information:

$$P(X_1 = 1) = 0.49$$
$$P(X_1 = 2) = 0.51$$
$$P(X_{t+1} = 1 | X_t = 1) = 1$$
$$P(X_{t+1} = 1 | X_t = 2) = 1$$

- At $t = 1$, the system can be in either state 1 or state 2, and from $t = 2$, system goes to state 1 and stay in state 1.
- So only possible state sequences are:

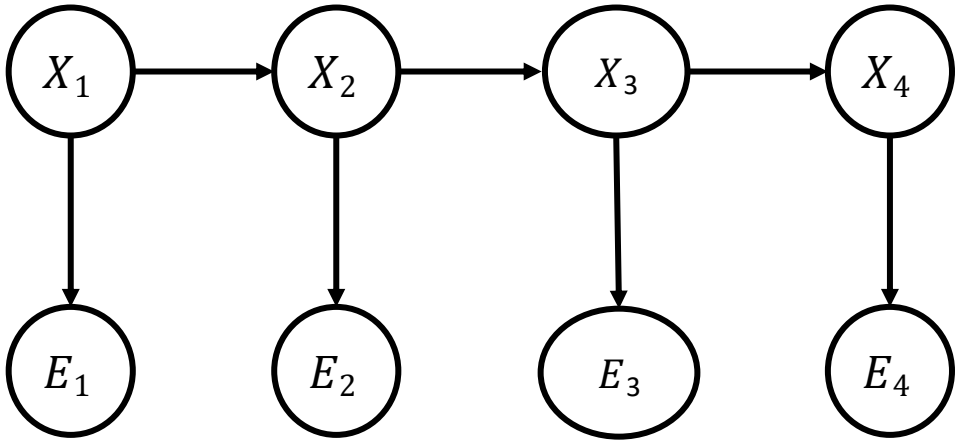
$$1 \to 1 \to 1$$
$$2 \to 1 \to 1$$

- Most probable output sequences:

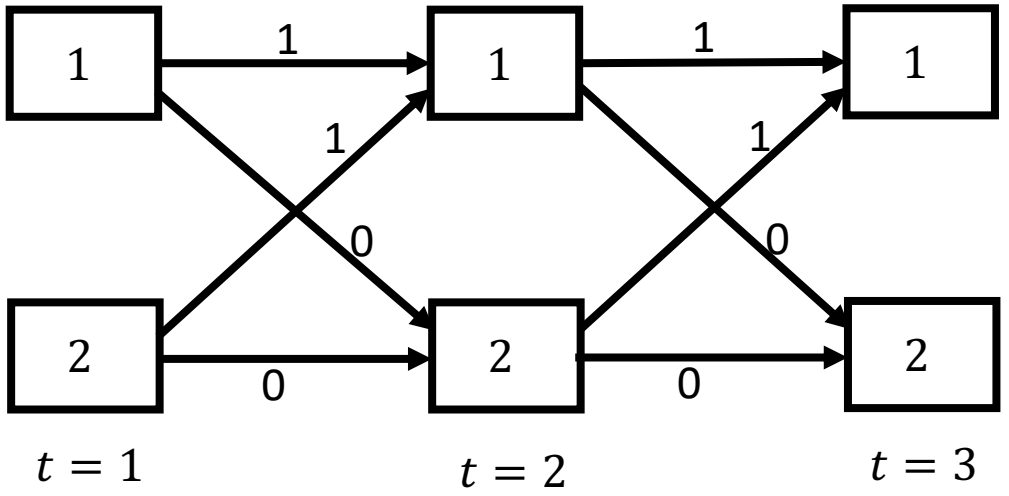
$$A \to A \to A$$
$$B \to A \to A$$

because:

$$P(E_t = A | X_t = 1) = 1$$
$$P(E_t = B | X_t = 2) = 1$$



Trellis



5. HMM: Probability of Sequences

$$
\begin{aligned}
P(E_1 = A, E_2 = A, E_3 = A) &= P(E_1 = A, E_2 = A, E_3 = A | X_1 = 1, X_2 = 1, X_3 = 1) \\
&\quad P(X_1 = 1, X_2 = 1, X_3 = 1) \\
&\quad + P(E_1 = A, E_2 = A, E_3 = A | X_1 = 2, X_2 = 1, X_3 = 1) \\
&\quad P(X_1 = 2, X_2 = 1, X_3 = 1) \\
&= P(X_1 = 1) P(E_1 = A | X_1 = 1) P(X_2 = 1 | X_1 = 1) \\
&\quad P(E_2 = A | X_2 = 1) P(X_3 = 1 | X_2 = 1) P(E_3 = A | X_3 = 1) \\
&\quad + P(X_1 = 2) P(E_1 = A | X_1 = 2) P(X_2 = 1 | X_1 = 2) \\
&\quad P(E_2 = A | X_2 = 1) P(X_3 = 1 | X_2 = 1) P(E_3 = A | X_3 = 1) \\
&= 0.49 + 0 \\
&= 0.49
\end{aligned}
$$

5. HMM: Probability of Sequences

$$
\begin{aligned}
P(E_1 = B, E_2 = A, E_3 = A) &= P(E_1 = B, E_2 = A, E_3 = A | X_1 = 1, X_2 = 1, X_3 = 1) \\
&\quad P(X_1 = 2, X_2 = 1, X_3 = 1) \\
&\quad + P(E_1 = B, E_2 = A, E_3 = A | X_1 = 2, X_2 = 1, X_3 = 1) \\
&\quad P(X_1 = 2, X_2 = 1, X_3 = 1) \\
&= 0.51 + 0 \\
&= 0.51
\end{aligned}
$$

B, A, A is the most probable sequence

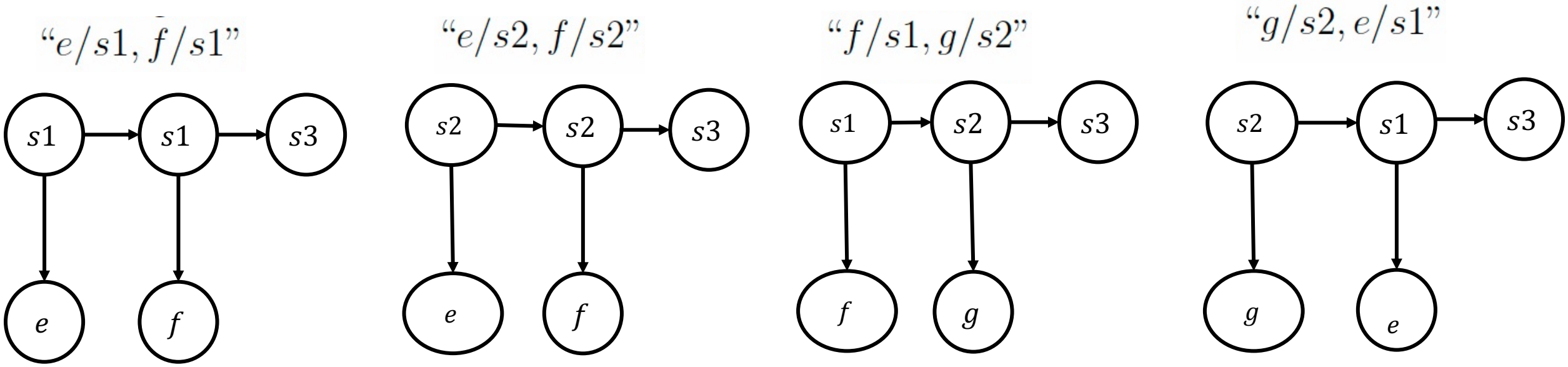
6. HMM: Parameter Learning

Let the set of output states be e, f, g and the set of hidden states be $s1, s2, s3$ where $s3$ is the stop state. Some data has been gathered regarding sequences of output states and their corresponding hidden states as follows: "$e/s1, f/s1$"; "$e/s2, f/s2$"; "$f/s1, g/s2$"; "$g/s2, e/s1$".

Assume T is the transition distribution, where each element T_{ij} represents a transition from state i to state j; O is the emission distribution, where each element $O_{o,j}$ represents the probability of output o in state j; and π_i is the initial probability distribution of states. State the maximum likelihood estimates for all possible π_i, T_{ij}, $O_{o,j}$ given the data.

6. HMM: Parameter Learning

Data:



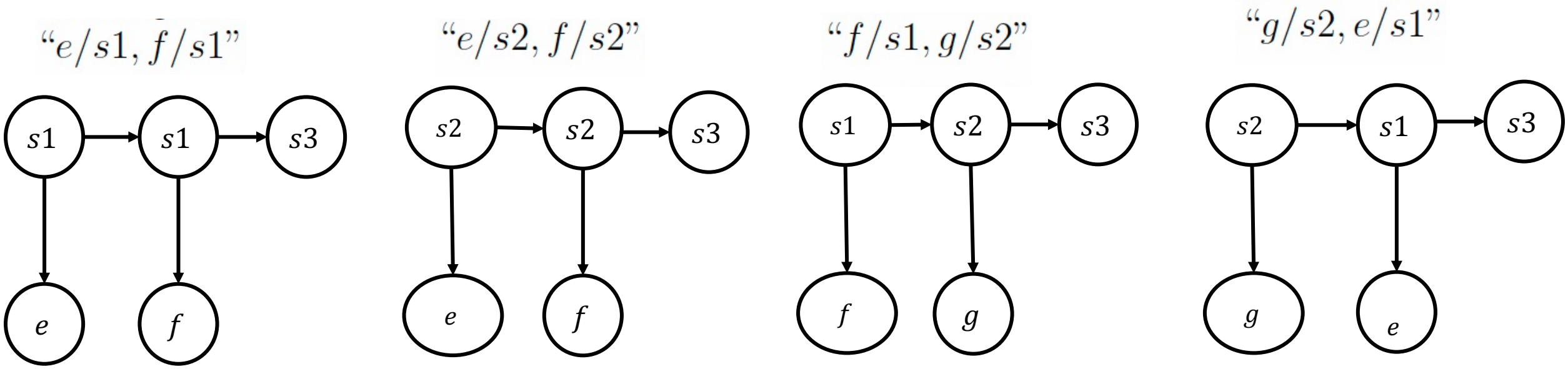
Let us index the states $s1$, $s2$ and $s3$ with $1, 2$ and 3, respectively.

$$P(X_1 = s_1) = \ P(X_1 = 1) = \pi_1 = \frac{2}{4} = \frac{1}{2}$$

$$P(X_1 = s_2) = \ P(X_1 = 2) = \pi_2 = \frac{2}{4} = \frac{1}{2}$$

6. HMM: Parameter Learning

Data:



Let us index the states $s1$, $s2$ and $s3$ with $1, 2$ and 3, respectively.
$s3$ is terminal state, no transitions from $s3$ to other states
So transition matrix contains two rows and three columns

$$T_{i,j} = \frac{count(i \to j)}{count(i)}$$

$$T_{1,1} = \frac{1}{4}; T_{1,2} = \frac{1}{4}; T_{1,3} = \frac{1}{2}$$

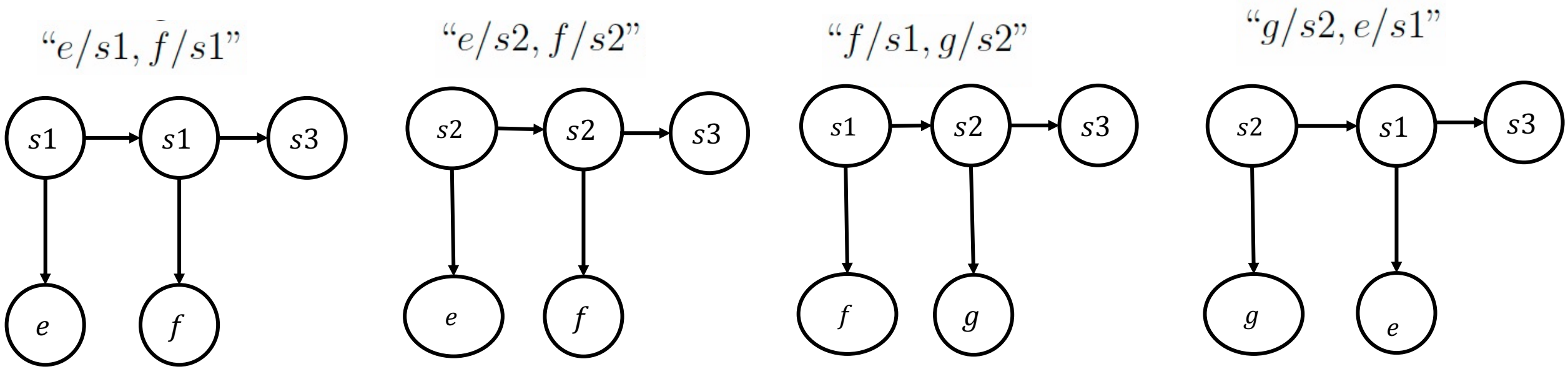
$$T_{2,1} = \frac{1}{4}; T_{2,2} = \frac{1}{4}; T_{2,3} = \frac{1}{2}$$

$$T = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

6. HMM: Parameter Learning

Given data:

Data:



Let us index the states $s1$, $s2$ and $s3$ with $1, 2$ and 3, respectively.

$$O_{j,o} = \frac{count(j \to o)}{count(j)}$$

$$O_{1,e} = \frac{1}{2}; O_{1,f} = \frac{1}{2}; O_{1,g} = 0;$$

$$O_{2,e} = \frac{1}{4}; O_{2,f} = \frac{1}{4}; O_{2,g} = \frac{1}{2};$$

7. HMM:

7. Let X_t be the hidden state at time t and it takes the values from the set S, A, B, C, D, where D is the $STOP$ state and the set of possible outputs be $E_1 = d$, $E_2 = p$, $E_3 = o$. You can assume 0 as the $START$ state. The following probabilities are given:

$$\pi_S = P(S|0) = 1, \pi_A = P(A|0) = 0, \pi_B = P(B|0) = 0, \pi_C = P(C|0) = 0$$

The transition and emission probabilities are shown in Table 1 and 2, respectively.

Table 1: Transition Probabilities

T_{ij}	$x_j = S$	$x_j = A$	$x_j = B$	$x_j = C$	$x_j = D$
$x_i = S$	0.3	0.2	0.1	0.1	0.3
$x_i = A$	0.25	0.1	0.05	0.15	0.45
$x_i = B$	0.05	0.4	0.15	0.2	0.2
$x_i = C$	0.5	0.2	0.2	0.05	0.05

Table 2: Emission Probabilities

O_{ij}	$e_j = d$	$e_j = p$	$e_j = o$
$x_i = S$	0.6	0.2	0.2
$x_i = A$	0.1	0.5	0.4
$x_i = B$	0.2	0.1	0.7
$x_i = C$	0.3	0.4	0.3

Calculate the following:

(a) $\mathbf{P}(X_1|E_1)$

(b) $\mathbf{P}(X_2|E_1, E_2)$

(c) $\mathbf{P}(X_3|E_1, E_2, E_3)$

Calculate the following:

(a) $\mathbf{P}(X_1|E_1)$

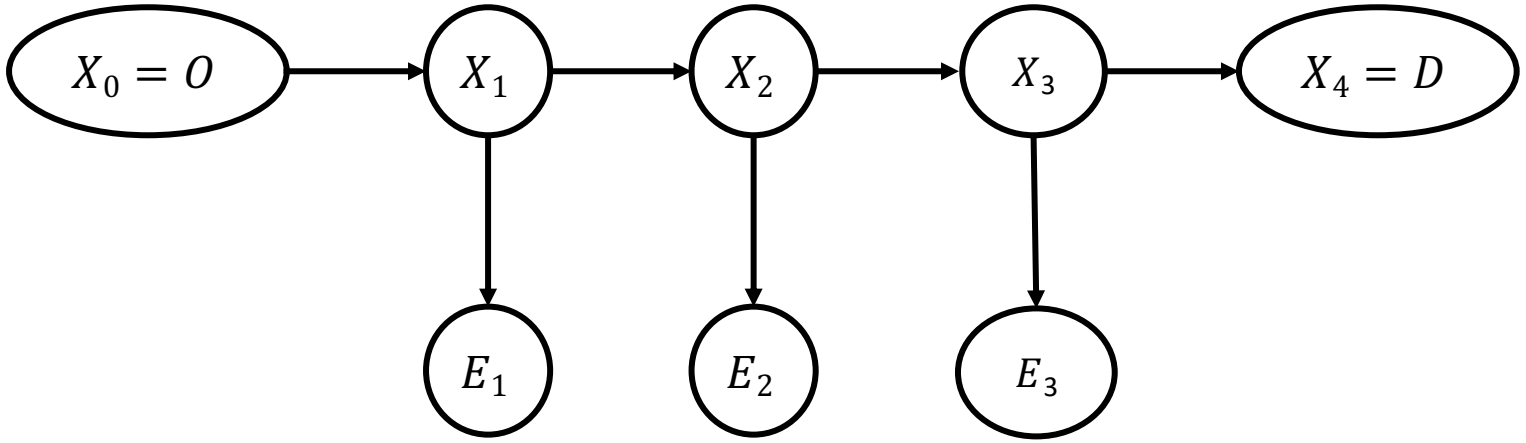
(b) $\mathbf{P}(X_2|E_1, E_2)$

(c) $\mathbf{P}(X_3|E_1, E_2, E_3)$

- Identify the inference problem?
  - Filtering
  - Prediction
  - Smoothing
  - Most Likely Sequence Estimation

HMM: Bayesian Filtering

- Sequence of evidence: d, p, o
- Calculation of $\boldsymbol{P}(X_1|E_1), \boldsymbol{P}(X_2|E_1, E_2), and\ \boldsymbol{P}(X_3|E_1, E_2, E_3)$ is a Bayesian filtering problem with
- $\boldsymbol{f_{1:1}} = \boldsymbol{P}(X_1|E_1)$
- $\boldsymbol{f_{1:2}} = \boldsymbol{P}(X_2|E_1, E_2)$
- $\boldsymbol{f_{1:3}} = \boldsymbol{P}(X_3|E_1, E_2, E_3)$



- Bayesian Filtering: $\boldsymbol{f}_{1:t+1} = \alpha\ \boldsymbol{O}_{t+1}\ \boldsymbol{T}^T\ \boldsymbol{f}_{1:t}$

Observation matrix changes with time

Transition matrix remains same (stationary MC)

Transition Probability

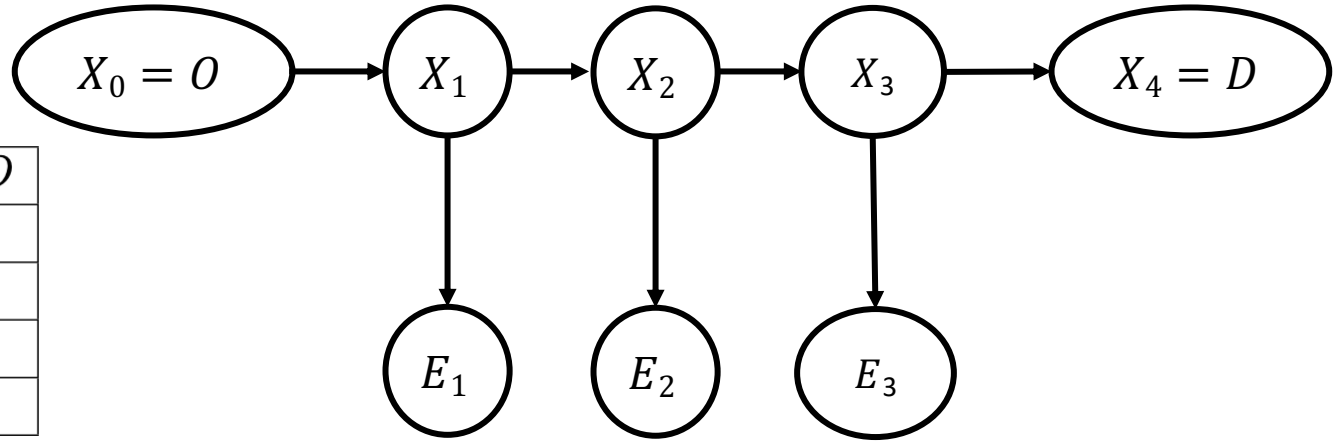


Table 1: Transition Probabilities

T_{ij}	$x_j = S$	$x_j = A$	$x_j = B$	$x_j = C$	$x_j = D$
$x_i = S$	0.3	0.2	0.1	0.1	0.3
$x_i = A$	0.25	0.1	0.05	0.15	0.45
$x_i = B$	0.05	0.4	0.15	0.2	0.2
$x_i = C$	0.5	0.2	0.2	0.05	0.05

Transition probabilities from START state

$$\boldsymbol{T} = \begin{array}{c} \\ O \\ S \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{cccccc} S & A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.3 & 0.2 & 0.1 & 0.1 & 0.3 \\ 0.25 & 0.1 & 0.05 & 0.15 & 0.45 \\ 0.05 & 0.4 & 0.15 & 0.2 & 0.2 \\ 0.5 & 0.2 & 0.2 & 0.05 & 0.05 \end{bmatrix} \end{array}$$

This matrix includes transition from START state ($X_0 = O$) and transitions to STOP state ($X_4 = D$)

Transition probabilities to STOP state

$$\boldsymbol{T} = \begin{array}{c} \\ S \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{cccc} S & A & B & C \end{array} \\ \begin{bmatrix} 0.3 & 0.2 & 0.1 & 0.1 \\ 0.25 & 0.1 & 0.05 & 0.15 \\ 0.05 & 0.4 & 0.15 & 0.2 \\ 0.5 & 0.2 & 0.2 & 0.05 \end{bmatrix} \end{array}$$

This matrix excludes transitions from START state ($X_0 = O$) and transitions to STOP state ($X_4 = D$)

Observation Matrix

$E_1 = \mathrm{d}$

$\boldsymbol{O}_1 = diag\,[P(d|S), P(d|A), P(d|B), P(d|C)]$

$= diag\,[0.6, 0.1, 0.2, 0.3]$

$$= \begin{bmatrix} 0.6 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0.3 \end{bmatrix}$$

$E_2 = p$

$\boldsymbol{O}_2 = diag\,[P(p|S), P(p|A), P(p|B), P(p|C)]$

$= diag\,[0.6, 0.1, 0.2, 0.3]$

$$= \begin{bmatrix} 0.6 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0.3 \end{bmatrix}$$

$E_3 = o$

$\boldsymbol{O}_3 = diag\,[P(o|S), P(o|A), P(o|B), P(o|C)]$

$= diag\,[0.2, 0.4, 0.7, 0.3]$

$$= \begin{bmatrix} 0.2 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 0.3 \end{bmatrix}$$

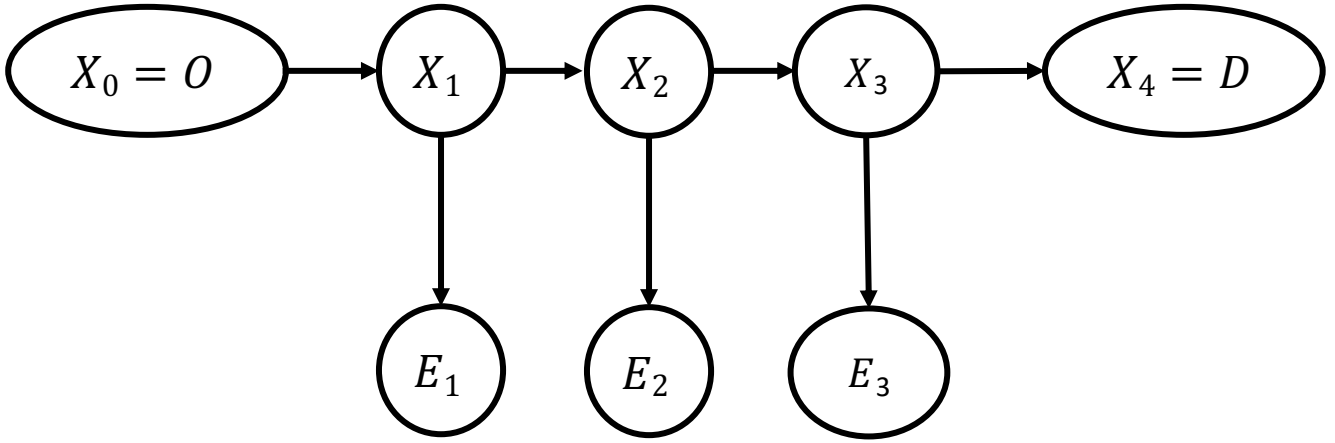
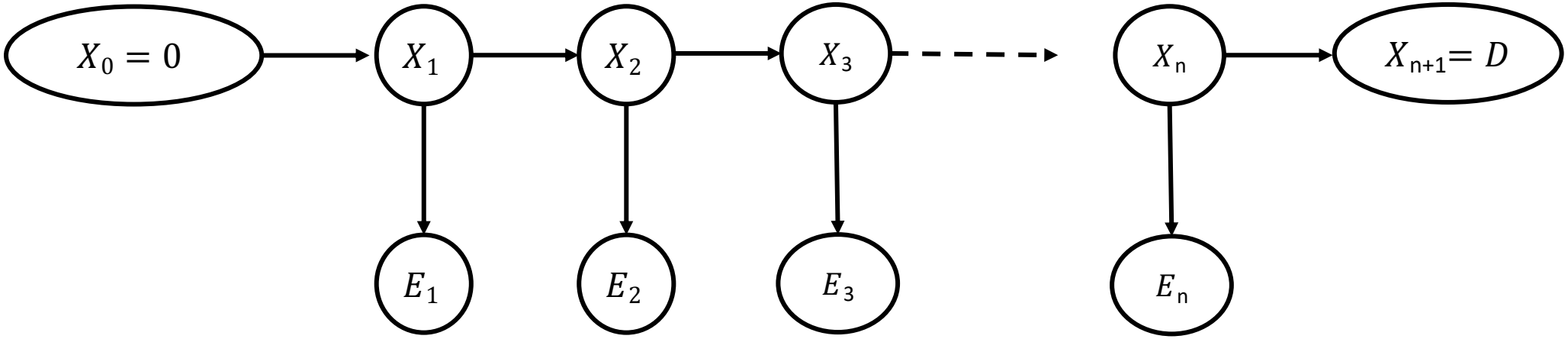


Table 2: Emission Probabilities

O_{ij}	$e_j = d$	$e_j = p$	$e_j = o$
$x_i = S$	0.6	0.2	0.2
$x_i = A$	0.1	0.5	0.4
$x_i = B$	0.2	0.1	0.7
$x_i = C$	0.3	0.4	0.3

Diagonal elements of the matrix $\boldsymbol{O}_{t+1}$ indicates probability of observing the evidence for a given state

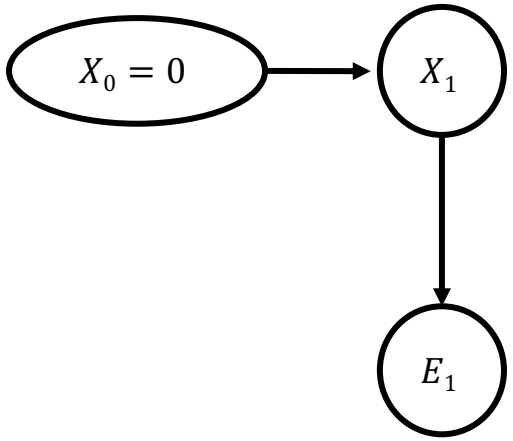
7. HMM



- $X_t \in \{S, A, B, C, D\},\ E_t \in \{d, p, o\}$

- $\boldsymbol{P}(X_1|O) = \begin{bmatrix} P(X_1 = S|O) \\ P(X_1 = A|O) \\ P(X_1 = B|O) \\ P(X_1 = C|O) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$

a. Calculate $\boldsymbol{P}(X_1|E_1)$



- $\boldsymbol{P}(X_1|E_1 = d) = \begin{bmatrix} P(X_1 = S|E_1 = d) \\ P(X_1 = A|E_1 = d) \\ P(X_1 = B|E_1 = d) \\ P(X_1 = C|E_1 = d) \end{bmatrix} = ?$

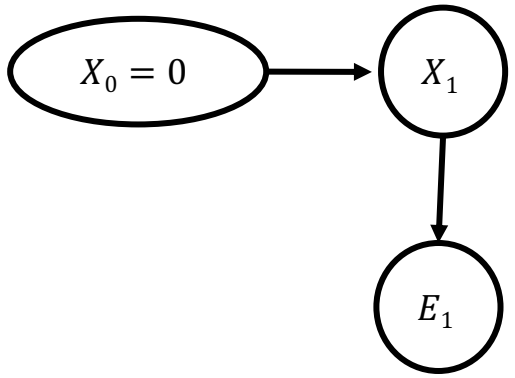
- Relevant Variables: X_0, X_1, E_1

$$\boldsymbol{P}(X_1|E_1 = d) = \frac{\boldsymbol{P}(X_1, E_1 = d)}{P(E_1 = d)}$$
$$= \alpha\, \boldsymbol{P}(X_1, E_1 = d)$$
$$= \alpha \boldsymbol{P}(X_0 = 0, X_1, E_1 = d)$$
$$= \alpha P(X_0 = 0)\boldsymbol{P}(X_1|X_0 = 0)\boldsymbol{P}(E_1 = d|X_1)$$

$$= \alpha.1. \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \circ \begin{bmatrix} 0.6 \\ 0.1 \\ 0.2 \\ 0.3 \end{bmatrix} = \alpha \begin{bmatrix} 0.6 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Hadamard product (element-wise product) of vectors

a. Calculate $\boldsymbol{P}(X_1|E_1 = \boldsymbol{d})$



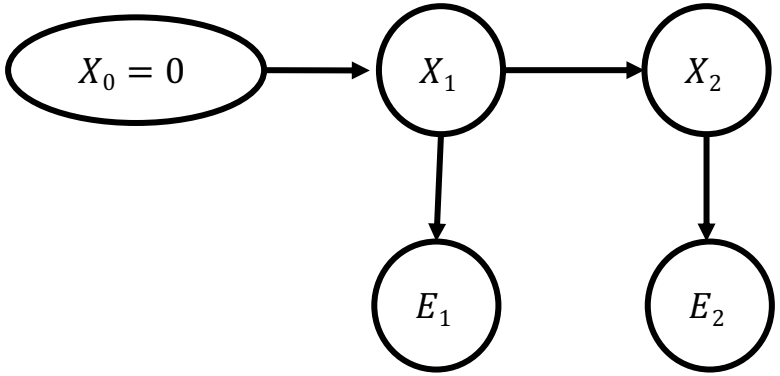
- $\boldsymbol{P}(X_1|E_1 = d) = \begin{bmatrix} P(X_1 = 0|E_1 = d) \\ P(X_1 = S|E_1 = d) \\ P(X_1 = A|E_1 = d) \\ P(X_1 = B|E_1 = d) \\ P(X_1 = C|E_1 = d) \end{bmatrix}$ =?

$$\boldsymbol{P}(X_1|E_1) = \boldsymbol{f}_{1:1} = \alpha \boldsymbol{O}_1 \boldsymbol{T}^T \boldsymbol{f}_0$$

$$\boldsymbol{O}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0 & 0.3 \end{bmatrix} \quad \boldsymbol{T} = \begin{array}{c} \\ O \\ S \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccccc} O & S & A & B & C \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0.3 & 0.2 & 0.1 & 0.1 \\ 0 & 0.25 & 0.1 & 0.05 & 0.15 \\ 0 & 0.05 & 0.4 & 0.15 & 0.2 \\ 0 & 0.5 & 0.2 & 0.2 & 0.05 \end{bmatrix} \end{array} \quad \boldsymbol{f}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\boldsymbol{f}_{1:1} = \alpha \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0 & 0.3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0.3 & 0.2 & 0.1 & 0.1 \\ 0 & 0.25 & 0.1 & 0.05 & 0.15 \\ 0 & 0.05 & 0.4 & 0.15 & 0.2 \\ 0 & 0.5 & 0.2 & 0.2 & 0.05 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \alpha \begin{bmatrix} 0 \\ 0.6 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

b. Calculate $\boldsymbol{P}(X_2|E_1, E_2)$



- $\boldsymbol{P}(X_2|E_1 = d, E_2 = p) = \begin{bmatrix} P(X_2 = S|E_1 = d, E_2 = p) \\ P(X_2 = A|E_1 = d, E_2 = p) \\ P(X_2 = B|E_1 = d, E_2 = p) \\ P(X_2 = C|E_1 = d, E_2 = p) \end{bmatrix} = ?$

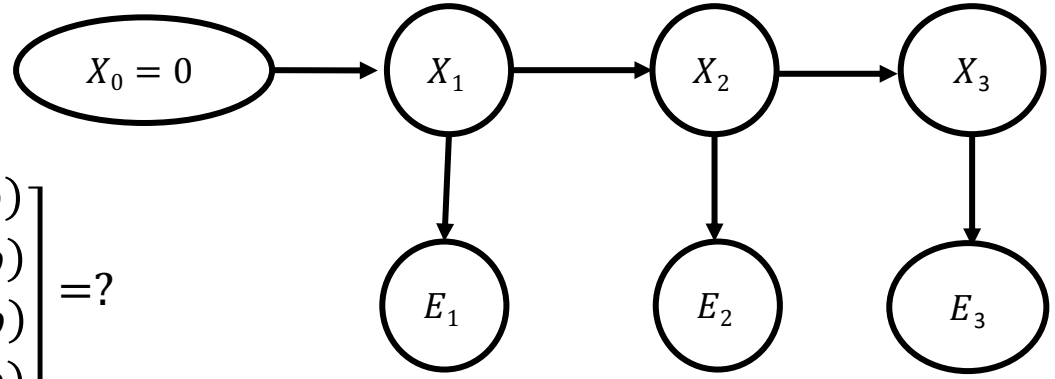
$\boldsymbol{P}(X_2|E_1, E_2) = \boldsymbol{f}_{1:2} = \alpha \boldsymbol{O}_2 \boldsymbol{T}^T \boldsymbol{f}_{1:1}$

$$\boldsymbol{O}_2 = \begin{bmatrix} 0.2 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.4 \end{bmatrix} \quad \boldsymbol{T} = \begin{bmatrix} 0.3 & 0.2 & 0.1 & 0.1 \\ 0.25 & 0.1 & 0.05 & 0.15 \\ 0.05 & 0.4 & 0.15 & 0.2 \\ 0.5 & 0.2 & 0.2 & 0.05 \end{bmatrix} \quad \boldsymbol{f}_{1:1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\boldsymbol{f}_{1:2} = \alpha \begin{bmatrix} 0.2 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.4 \end{bmatrix} \begin{bmatrix} 0.3 & 0.2 & 0.1 & 0.1 \\ 0.25 & 0.1 & 0.05 & 0.15 \\ 0.05 & 0.4 & 0.15 & 0.2 \\ 0.5 & 0.2 & 0.2 & 0.05 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$= \alpha \begin{bmatrix} 0.06 \\ 0.1 \\ 0.01 \\ 0.04 \end{bmatrix} = \frac{1}{0.06+0.1+0.01+0.04} \begin{bmatrix} 0.06 \\ 0.1 \\ 0.01 \\ 0.04 \end{bmatrix} = \frac{1}{0.21} \begin{bmatrix} 0.06 \\ 0.1 \\ 0.01 \\ 0.04 \end{bmatrix} = \begin{bmatrix} 0.285 \\ 0.476 \\ 0.0476 \\ 0.1904 \end{bmatrix}$$

c. Calculate $\boldsymbol{P}(X_3|E_1, E_2, E_3)$



- $\boldsymbol{P}(X_3|E_1 = d, E_2 = p, E_3 = o) = \begin{bmatrix} P(X_3 = S|E_1 = d, E_2 = p, E_3 = o) \\ P(X_3 = A|E_1 = d, E_2 = p, E_3 = o) \\ P(X_3 = B|E_1 = d, E_2 = p, E_3 = o) \\ P(X_3 = C|E_1 = d, E_2 = p, E_3 = o) \end{bmatrix} = ?$

$\boldsymbol{P}(X_3|E_1, E_2, , E_3) = \boldsymbol{f}_{1:3} = \alpha \boldsymbol{O}_3 \boldsymbol{T}^T \boldsymbol{f}_{1:2}$

$$\boldsymbol{O}_3 = \begin{bmatrix} 0.2 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 0.3 \end{bmatrix} \quad \boldsymbol{T} = \begin{bmatrix} 0.3 & 0.2 & 0.1 & 0.1 \\ 0.25 & 0.1 & 0.05 & 0.15 \\ 0.05 & 0.4 & 0.15 & 0.2 \\ 0.5 & 0.2 & 0.2 & 0.05 \end{bmatrix} \quad \boldsymbol{f}_{1:2} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\boldsymbol{f}_{1:3} = \alpha \begin{bmatrix} 0.2 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 0.3 \end{bmatrix} \begin{bmatrix} 0.3 & 0.2 & 0.1 & 0.1 \\ 0.25 & 0.1 & 0.05 & 0.15 \\ 0.05 & 0.4 & 0.15 & 0.2 \\ 0.5 & 0.2 & 0.2 & 0.05 \end{bmatrix}^T \begin{bmatrix} 0.285 \\ 0.476 \\ 0.0476 \\ 0.1904 \end{bmatrix} = \begin{bmatrix} 0.2637 \\ 0.2824 \\ 0.298 \\ 0.1557 \end{bmatrix}$$

Check Q7BayesianFiltering.jpynb