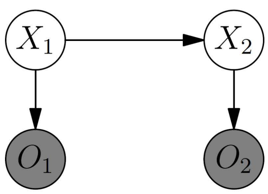
ARTIFICIAL INTELLIGENCE (E016330)
GHENT UNIVERSITY
AY 2024/2025
Professor Aleksandra Pizurica
Asst: Niels Vyncke, Nicolas Vercheval
E-mail: ai@lists.ugent.be

Solutions: Probabilistic reasoning over time

1. Consider the following simple Hidden Markov Model with state variables X_t and observation (evidence) variables O_t, which are shaded below.



X_1	$\mathrm{P}(X_1)$
0	0.3
1	0.7

X_t	X_{t+1}	$\mathrm{P}(X_{t+1}\|X_t)$
0	0	0.4
0	1	0.6
1	0	0.8
1	1	0.2

X_t	O_t	$\mathrm{P}(O_t\|X_t)$
0	A	0.9
0	B	0.1
1	A	0.5
1	B	0.5

Suppose that $O_1 = A$ and $O_2 = B$ is observed.

(a) Use the *Forward Algorithm* to compute the probability distribution

$$\mathbf{P}(X_2 \mid O_1 = A,\ O_2 = B).$$

Show your work.

(b) Use the Viterbi algorithm to compute the maximum probability sequence $X_1,\ X_2$. Show your work.

Solution:

(a) In theory slides for *Temporal Probability Models*, we have seen that a *filtered estimate* is given by:

$$\mathbf{P}(X_{t+1} \mid \mathbf{e}_{1:t+1}) = \alpha\,\mathbf{P}(e_{t+1} \mid X_{t+1}) \underbrace{\sum_{x_t} \mathbf{P}(X_{t+1} \mid x_t) P(x_t \mid \mathbf{e}_{1:t})}_{\mathbf{P}(X_{t+1}|\mathbf{e}_{1:t})}.$$

and we denoted this procedure as the forward algorithm:

$$\mathbf{f}_{1:t+1} = \alpha\,\textsc{Forward}(\mathbf{f}_{1:t}, e_{t+1}), \quad \text{where} \ \ \mathbf{f}_{1:t} = \mathbf{P}(X_t|\mathbf{e}_{1:t})$$

Now, it easily follows that:

$$\mathbf{P}(X_2 \mid O_1 = A, O_2 = B) = \alpha\, \mathbf{P}(O_2 = B \mid X_2) \sum_{x_1} \mathbf{P}(X_2 \mid x_1) P(x_1 \mid O_1 = A)$$

$$= \alpha\langle 0.5, 0.1\rangle \left[\langle 0.2, 0.8\rangle \frac{P(O_1 = A \mid x_1 = 1)P(x_1 = 1)}{P(O_1 = A)} + \langle 0.6, 0.4\rangle \frac{P(O_1 = A \mid x_1 = 0)P(x_1 = 0)}{P(O_1 = A)}\right]$$

$$= \underbrace{\frac{\alpha}{P(O_1 = A)}}_{:=\alpha'} \langle 0.5, 0.1\rangle \left[\langle 0.2, 0.8\rangle \cdot 0.5 \cdot 0.7 + \langle 0.6, 0.4\rangle \cdot 0.9 \cdot 0.3\right]$$

$$= \alpha'\langle 0.5, 0.1\rangle \left[\langle 0.07, 0.28\rangle + \langle 0.162, 0.108\rangle\right]$$

$$= \alpha'\langle 0.116, 0.0388\rangle = \langle 0.75, 0.25\rangle.$$

(b) We know that the *Viterbi algorithm* is defined by the recurrence relation:

$$\begin{aligned}
\mathbf{m}_{1:t+1} &= \max_{x_1 \dots x_t} \mathbf{P}(x_1, \dots, x_t, X_{t+1} \mid \mathbf{e}_{1:t+1}) \\
&= \alpha\, \mathbf{P}(e_{t+1} \mid X_{t+1}) \max_{x_t}\Big(\mathbf{P}(X_{t+1} \mid x_t) \max_{x_1 \dots x_{t-1}} \mathbf{P}(x_1, \dots, x_{t-1}, X_t \mid \mathbf{e}_{1:t})\Big)
\end{aligned}$$

The first step is filtering:

$$\begin{aligned}
\mathbf{m}_{1:1} = \langle M^1_{1:1}, M^2_{1:1}\rangle &= \mathbf{P}(X_1 | O_1 = A) = \alpha\, \mathbf{P}(A|X_1)\mathbf{P}(X_1) = \alpha\langle 0.5, 0.9\rangle\langle 0.7, 0.3\rangle \\
&= \alpha\langle 0.35, 0.27\rangle = \langle 0.565, 0.435\rangle
\end{aligned}$$

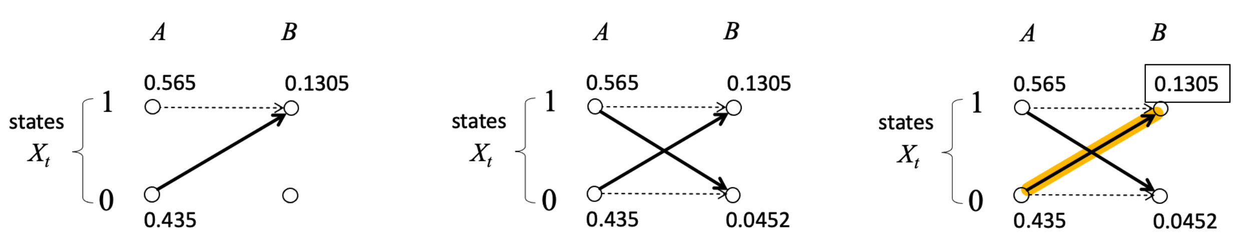
Then we evaluate the two possible ways to reach state 1 at time instant 2:

$$\begin{aligned}
M^1_{1:2} &= \alpha P(O_2 = B|X_2 = 1) \max\{P(X_2 = 1|X_1 = 1)M^1_{1:1}, P(X_2 = 1|X_1 = 0)M^2_{1:1}\} \\
&= 0.5 \max\{\underbrace{0.2 \times 0.565}_{0.113}, \underbrace{0.6 \times 0.435}_{\mathbf{0.261}}\}\alpha = 0.1305\alpha
\end{aligned}$$

The second term was larger, i.e., the transition from the state $X_1 = 0$ was chosen and this is indicated in the trellis diagram (left subfigure). Similarly,

$$\begin{aligned}
M^2_{1:2} &= \alpha P(O_2 = B|X_2 = 0) \max\{P(X_2 = 0|X_1 = 1)M^1_{1:1}, P(X_2 = 0|X_1 = 0)M^2_{1:1}\} \\
&= 0.1 \max\{\underbrace{0.8 \times 0.565}_{\mathbf{0.452}}, \underbrace{0.4 \times 0.435}_{0.174}\}\alpha = 0.0452\alpha
\end{aligned}$$

The transition from $X_1 = 1$ was chosen (we denote this in the trellis diagram, middle subfigure) It is not necessary to calculate the normalized values with actual α (although that will help when the sequence is longer in order not to have too small numbers). Obviously, the state 1 is most probable at the end, so tracing back we obtain the most likely sequence: $X_1 = 0$, $X_2 = 1$.



2. **(Old exam question)** Suppose you returned from holidays and due to the current Covid-19 regulations have to stay in quarantine for two weeks. You are staying in a good ventilated, but windowless basement room in your home. To make time go by faster, you want to develop a simple weather forecast system that only gives prognosis for average daily weather (W) characterized as being either *sunny* (s), *rainy* (r) or *foggy* (f), i.e., $w \in \{s, rf\}$.

   You will make the system more sophisticated by making use of some indirect evidence that you get by observing the caretaker that visits you each morning. In particular you pay attention to whether the caretaker caries an umbrella or not. You assume that the probability that the caretaker carries an umbrella is 0.1 if the weather is sunny, 0.8 if the weather is rainy and 0.3 if it is foggy.

   Based on some statistical data that you could find for Ghent region in Belgium, you set probabilities of tomorrow's weather based on today's weather as in the table below.

		Tomorrow's weather		
Today's weather		*Sunny*	*Rainy*	*Foggy*
	Sunny	0.8	0.05	0.15
	Rainy	0.2	0.6	0.2
	Foggy	0.2	0.3	0.5

   (a) Draw the corresponding state transition diagram that represents the transition model with the state transition probabilities.

   (b) Specify the sensor model with an appropriate table.

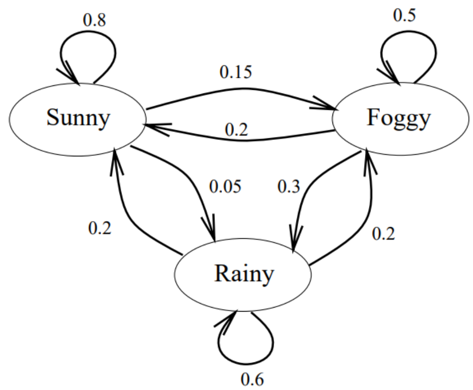
   (c) Suppose first you discard the information that you can obtain by observing the caretaker. What is the probability that it will be rainy two days from now given that today is foggy?

   (d) You have no idea anymore about what the weather was before you landed in the basement and it is your second day there. What is the probability that on this day 2 of your quarantine the weather outside is sunny provided that the caretaker didn't carry the umbrella on day 1 and caries it on day 2?

   (e) On the first 3 days your umbrella observations are: $\{no\_umbrella, umbrella, umbrella\}$. Find the most probable weather-sequence using the Viterbi algorithm.

Solution:

(a) We obtain the following diagram:



(b) The only piece of evidence you have is whether the person who comes into the room is carrying an umbrella or not.

Weather	Probability of umbrella
Sunny	0.1
Rainy	0.8
Foggy	0.3

(c) There are three ways to get from foggy today to rainy two days from now:

$$\{foggy,\ foggy,\ rainy\},\quad \{foggy,\ rainy,\ rainy\}\quad \text{and}\quad \{foggy,\ sunny,\ rainy\}.$$

Therefore we have to sum over these paths. Let us use the shorter notation $\{r, s, f\}$ and denote by W_n the weather on day n. We obtain:

$$\begin{aligned}
P(W_3 = r \mid W_1 = f) &= P(W_3 = r \mid W_1 = f, W_2 = f)P(W_2 = f \mid W_1 = f) \\
&+ P(W_3 = r \mid W_1 = f, W_2 = s)P(W_2 = s \mid W_1 = f) \\
&+ P(W_3 = r \mid W_1 = f, W_2 = r)P(W_2 = r \mid W_1 = f) \\
&= P(W_3 = r \mid W_2 = f)P(W_2 = f \mid W_1 = f) \\
&+ P(W_3 = r \mid W_2 = s)P(W_2 = s \mid W_1 = f) \\
&+ P(W_3 = r \mid W_2 = r)P(W_2 = r \mid W_1 = f) \\
&= 0.3 \cdot 0.5 + 0.05 \cdot 0.2 + 0.6 \cdot 0.3 = 0.34.
\end{aligned}$$

(d) Here we have a filtering task:

$$\mathbf{P}(X_{t+1} \mid \mathbf{e}_{1:t+1}) = \alpha\, \mathbf{P}(e_{t+1} \mid X_{t+1}) \sum_{x_t} \mathbf{P}(X_{t+1} \mid x_t) P(x_t \mid \mathbf{e}_{1:t}).$$

By assuming that the weather on day one is equiprobable, we obtain the following:

$$\begin{aligned}
\mathbf{P}(W_2 \mid U_1 = f, U_2 = t) &= \alpha\,\mathbf{P}(U_2 = t \mid W_2) \sum_{w_1} \mathbf{P}(W_2 \mid w_1) P(w_1 \mid U_1 = f) \\
&= \alpha\,\langle 0.1, 0.8, 0.3\rangle [\langle 0.8, 0.05, 0.15\rangle\, P(W_1 = s \mid U_1 = f) \\
&+ \langle 0.2, 0.6, 0.2\rangle\, P(W_1 = r \mid U_1 = f) \\
&+ \langle 0.2, 0.3, 0.5\rangle\, P(W_1 = f \mid U_1 = f)] \\
&= \underbrace{\frac{\alpha}{P(U_1 = f)} \frac{1}{3}}_{:=\alpha'} [\langle 0.8, 0.05, 0.15\rangle\, P(U_1 = f \mid W_1 = s)
\end{aligned}$$

$$\begin{aligned}
&+\langle 0.2, 0.6, 0.2\rangle\, P(U_1 = f \mid W_1 = r) + \langle 0.2, 0.3, 0.5\rangle\, P(U_1 = f \mid W_1 = f)] \\
&= \alpha' \langle 0.1, 0.8, 0.3\rangle [\langle 0.8, 0.05, 0.15\rangle 0.9 + \langle 0.2, 0.6, 0.2\rangle 0.2 + \langle 0.2, 0.3, 0.5\rangle 0.7] \\
&= \alpha' \langle 0.1, 0.8, 0.3\rangle [\langle 0.72, 0.045, 0.135\rangle + \langle 0.04, 0.12, 0.04\rangle + \langle 0.14, 0.21, 0.35\rangle] \\
&= \alpha' \langle 0.1, 0.8, 0.3\rangle\, \langle 0.9, 0.375, 0.525\rangle = \alpha' \langle 0.09, 0.3, 0.1575\rangle \approx \langle 0.164, 0.548, 0.287\rangle
\end{aligned}$$

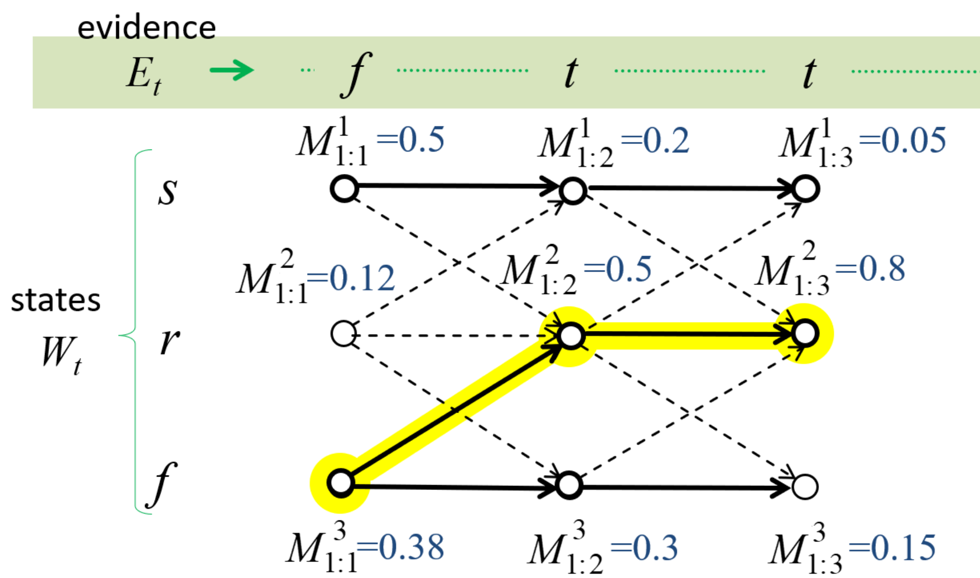
So, we obtain that

$$P(W_2 = s \mid U_1 = f, U_2 = t) = 0.164.$$

(e) The *Viterbi algorithm* is defined by the recurrence relation:

$$\begin{aligned}
\mathbf{m}_{1:t+1} &= \max_{x_1 \ldots x_t} \mathbf{P}(x_1, \ldots, x_t, X_{t+1} \mid \mathbf{e}_{1:t+1}) \\
&= \alpha\,\mathbf{P}(e_{t+1} \mid X_{t+1}) \max_{x_t} \Big(\mathbf{P}(X_{t+1} \mid x_t) \max_{x_1 \ldots x_{t-1}} \mathbf{P}(x_1, \ldots, x_{t-1}, X_t \mid \mathbf{e}_{1:t}) \Big)
\end{aligned}$$

To visualize the algorithm, consider the following state trellis, a graph of states and transitions over time:

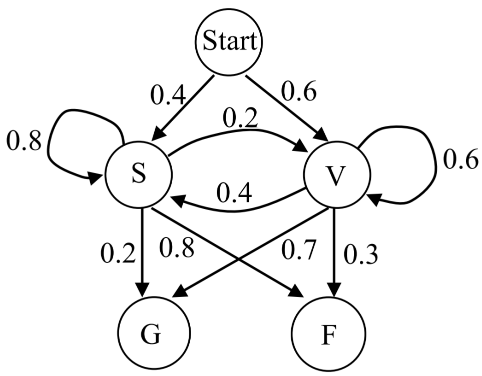


The most likely sequence is: $W_3 = r$, $W_2 = r$ and $W_1 = f$.

3. **(Old exam question)** Sam has bought an AI agent to be his companion helping him to keep good mood during the exam period. The AI agent can analyse facial expressions and adjust accordingly lighting in the room, music and tell some jokes to stimulate this way different activities like studying or relaxing. During the first day, the agent is just observing Sam's facial expressions and based on average observations in given time intervals infers the type of Sam's activity in those corresponding intervals.

   Suppose that the time interval is set to one hour and that Sam's activity during each time interval is either *Studying* (S) or playing *Video games* (V). His facial expression that the agent observes is either *Grinning* (G) or *Frowning* (F).

   (a) Based on the available information from the given state space diagram, write the tables for the transition and sensor models and any other if available from the problem description. Use D_i to denote the random variable (RV) representing what Sam is doing in hour i and denote by O_i the observation RV in that time slot.



   (b) If in the second hour Sam is *Studying*, what's the probability that in the fourth hour he is playing *Video games*? Assume the model from (a).

   (c) Under the model in (a), what is the probability that Sam is studying in the second hour if the observation sequence for the first two hours is $\{Grinning, Frowning\}$?

   (d) Using the same model, the agent recalculates the probabilities of Sam's activity during hour 2, after observing that his facial expression during hour 3 is *Grinning*. Show this calculation and the updated probabilities of the activities in the specified time slot. Comment on this result.

   (e) Suppose now that Sam's activities also include *ChattingWithFriend* (C) and the set of his observable facial expressions is extended with two others: *Excited* (E) and *Bored* (B). The transition model and the sensor model are given by the two matrices below:

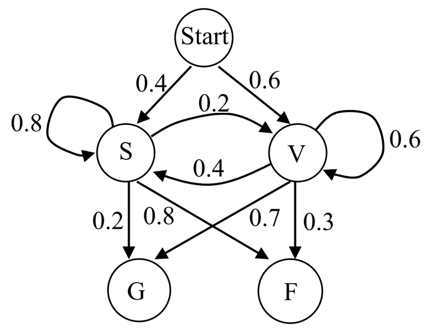
$$\begin{array}{r c} & \begin{array}{ccc} \text{Studying} & \text{VideoGames} & \text{Chatting} \\ S & V & C \end{array} \\ \begin{array}{r} \text{Studying } S \\ \text{VideoGames } V \\ \text{Chatting } C \end{array} & \begin{bmatrix} 0.1 & 0.2 & 0.2 \\ 0.1 & 0.7 & 0.7 \\ 0.8 & 0.1 & 0.1 \end{bmatrix} \end{array} \qquad \begin{array}{r c} & \begin{array}{ccc} \text{Studying} & \text{VideoGames} & \text{Chatting} \\ S & V & C \end{array} \\ \begin{array}{r} \text{Frowning } F \\ \text{Grinning } G \\ \text{Bored } B \\ \text{Excited } E \end{array} & \begin{bmatrix} 0.05 & 0.7 & 0.25 \\ 0.05 & 0.15 & 0.4 \\ 0.25 & 0.05 & 0.05 \\ 0.65 & 0.1 & 0.3 \end{bmatrix} \end{array}$$

For example, the probability that Sam is chatting with a friend in the hour $i+1$ if he was studying during the hour i is $P(C|S) = T_{3,1} = 0.8$ and the probability that he is playing video games if he was studying the hour before is $P(V|S) = T_{2,1} = 0.1$. He is grinning with probability 0.05 while studying and with probability 0.4 while chatting.

The agent observes the following sequence of facial expressions: $\{Grinning, Excited, Frowning\}$. Which sequence of activities does it infer by applying the Viterbi algorithm? Write the resulting sequence and mark it on the trellis diagram!

Solution:

(a) We obtain the following:



X_t	X_{t+1}	$P(X_{t+1}\|X_t)$
S	S	0.8
S	V	0.2
V	S	0.4
V	V	0.6

X_0	$P(X_0)$
S	0.4
V	0.6

X_t	O_t	$P(O_t\|X_t)$
S	G	0.2
S	F	0.8
V	G	0.7
V	F	0.3

(b) Denote the state in hour i by D_i which can be $S = Studying$ and $V = Video\_games$. We obtain the following:

$$\begin{aligned} P(D_4 = V \mid D_2 = S) &= P(D_4 = V \mid D_3 = V, D_2 = S)P(D_3 = V \mid D_2 = S) \\ &+ P(D_4 = V \mid D_3 = S, D_2 = S)P(D_3 = S \mid D_2 = S) \\ &= P(D_4 = V \mid D_3 = V)P(D_3 = V \mid D_2 = S) + P(D_4 = V \mid D_3 = S)P(D_3 = S \mid D_2 = S) \\ &= 0.6 \cdot 0.2 + 0.2 \cdot 0.8 = 0.28 \end{aligned}$$

(c) We are asked to calculate $P(D_2 = S \mid O_1 = G, O_2 = F)$, which is a filtering task. We have the following:

$$\mathbf{P}(D_2 \mid O_1 = G, O_2 = F) = \alpha\,\mathbf{P}(O_2 = F \mid D_2) \sum_{d_1} \mathbf{P}(D_2 \mid d_1)P(d_1 \mid O_1 = G)$$

$$= \alpha\,\langle 0.8, 0.3\rangle\,[\mathbf{P}(D_2 \mid D_1 = S)P(D_1 = S \mid O_1 = G) + \mathbf{P}(D_2 \mid D_1 = V)P(D_1 = V \mid O_1 = G)]$$

$$= \alpha\,\langle 0.8, 0.3\rangle\,(\langle 0.8, 0.2\rangle 0.267 + \langle 0.4, 0.6\rangle 0.733) = \alpha\langle 0.8, 0.3\rangle\langle 0.5, 0.5\rangle \approx \langle 0.72, 0.28\rangle$$

Thus we conclude: $P(D_2 = S \mid O_1 = G, O_2 = F) = 0.72$

Note that we also calculated

$$\begin{aligned} \mathbf{P}(D_1 \mid O_1 = G) &= \alpha\,\mathbf{P}(O_1 = G \mid D_1)\mathbf{P}(D_1) = \alpha\langle 0.2, 0.7\rangle \cdot \langle 0.56, 0.44\rangle \\ &= \alpha\langle 0.112, 0.308\rangle \approx \langle 0.267, 0.733\rangle \end{aligned}$$

and

$$\mathbf{P}(D_1) = \sum_{d_0} \mathbf{P}(D_1 \mid d_0) P(d_0) = \langle 0.8, 0.2 \rangle 0.4 + \langle 0.4, 0.6 \rangle 0.6 = \langle 0.56, 0.44 \rangle$$

(d) Here we have to calculate $\mathbf{P}(D_2 \mid O_1 = G, O_2 = F, O_3 = G)$, which is a smoothing task. We have the following:

$$\begin{aligned} \mathbf{P}(D_2 \mid o_{1:3}) &= \alpha \mathbf{P}(D_2 \mid o_{1:2}) \mathbf{P}(o_3 \mid D_2) = \alpha \langle 0.72, 0.28 \rangle \cdot \langle 0.3, 0.5 \rangle \\ &= \alpha \langle 0.216, 0.14 \rangle \approx \langle 0.6, 0.4 \rangle \end{aligned}$$

We also had to calculate

$$\begin{aligned} \mathbf{P}(O_3 = G \mid D_2) &= \sum_{d_3} P(O_3 = G \mid d_3) \mathbf{P}(d_3 \mid D_2) \\ &= 0.2 \cdot \langle 0.8, 0.4 \rangle + 0.7 \cdot \langle 0.2, 0.6 \rangle = \langle 0.3, 0.5 \rangle \end{aligned}$$

Note also that

$$\begin{aligned} \mathbf{P}(D_2 \mid O_1 = G, O_2 = F) &= \langle 0.72, 0.28 \rangle \\ \mathbf{P}(D_2 \mid O_1 = G, O_2 = F, O_3 = G) &= \langle 0.6, 0.4 \rangle \end{aligned}$$

We have that the smoothed probability of studying is 0.6 which is smaller than the filtered probability of studying which is 0.72. This is because of the persistence in video game playing (and much larger probability of grinning during video games than during studying).

(e) We know that the *Viterbi algorithm* is defined by the recurrence relation:

$$\begin{aligned} \mathbf{m}_{1:t+1} &= \max_{x_1 \ldots x_t} \mathbf{P}(x_1, \ldots, x_t, X_{t+1} \mid \mathbf{e}_{1:t+1}) \\ &= \alpha\, \mathbf{P}(e_{t+1} \mid X_{t+1}) \max_{x_t} \Big(\mathbf{P}(X_{t+1} \mid x_t) \max_{x_1 \ldots x_{t-1}} \mathbf{P}(x_1, \ldots, x_{t-1}, X_t \mid \mathbf{e}_{1:t}) \Big) \end{aligned}$$

The first step is filtering:

$$\begin{aligned} \mathbf{m}_{1:1} &= \langle M^1_{1:1}, M^2_{1:1}, M^3_{1:1} \rangle = \mathbf{P}(X_1 | O_1 = G) = \alpha\, \mathbf{P}(G | X_1) \mathbf{P}(X_1) \\ &= \alpha \langle 0.05, 0.15, 0.4 \rangle \left\langle \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right\rangle = \underbrace{\frac{\alpha}{3}}_{\alpha'} \langle 0.05, 0.15, 0.4 \rangle = \langle 0.08, 0.25, 0.67 \rangle \end{aligned}$$

Then we evaluate the possible ways to reach state *Studying* at time instant 2:

$$\begin{aligned} M^1_{1:2} &= \alpha P(O_2 = E | X_2 = S) \max\{ P(S|S) M^1_{1:1}, P(S|V) M^2_{1:1}, P(S|C) M^3_{1:1} \} \\ &= 0.65 \max\{ \underbrace{0.1 \times 0.08}_{0.008}, \underbrace{0.2 \times 0.25}_{0.05}, \underbrace{0.2 \times 0.67}_{\mathbf{0.1345}} \} \alpha = 0.0871 \alpha \end{aligned}$$

The third term is the largest, i.e., the transition from the state $X_1 = C$ was chosen and this is indicated in the trellis diagram. Similarly, to reach state $VideoGaming$ at time instant 2:

$$\begin{aligned} M^2_{1:2} &= \alpha P(O_2 = E|X_2 = V) \max\{P(V|S)M^1_{1:1}, P(V|V)M^2_{1:1}, P(V|C)M^3_{1:1}\} \\ &= 0.1 \max\{\underbrace{0.1 \times 0.08}_{0.008}, \underbrace{0.7 \times 0.25}_{0.175}, \underbrace{0.7 \times 0.67}_{\mathbf{0.469}}\}\alpha = 0.0469\alpha \end{aligned}$$

The transition from $X_1 = C$ was chosen (we denote this in the trellis diagram). Finally, we obtain for state $Chatting$ at time instant 2:

$$\begin{aligned} M^3_{1:2} &= \alpha P(O_2 = E|X_2 = C) \max\{P(C|S)M^1_{1:1}, P(C|V)M^2_{1:1}, P(C|C)M^3_{1:1}\} \\ &= 0.3 \max\{\underbrace{0.8 \times 0.08}_{0.064}, \underbrace{0.1 \times 0.25}_{0.025}, \underbrace{0.1 \times 0.67}_{\mathbf{0.067}}\}\alpha = 0.02\alpha \end{aligned}$$

Now we will calculate the normalized values with actual α as this will help when the sequence is longer in order not to have too small numbers. Thus, for $\alpha = 6.493$ we have that

$$\mathbf{m}_{1:2} = \langle 0.565, 0.3, 0.135 \rangle.$$

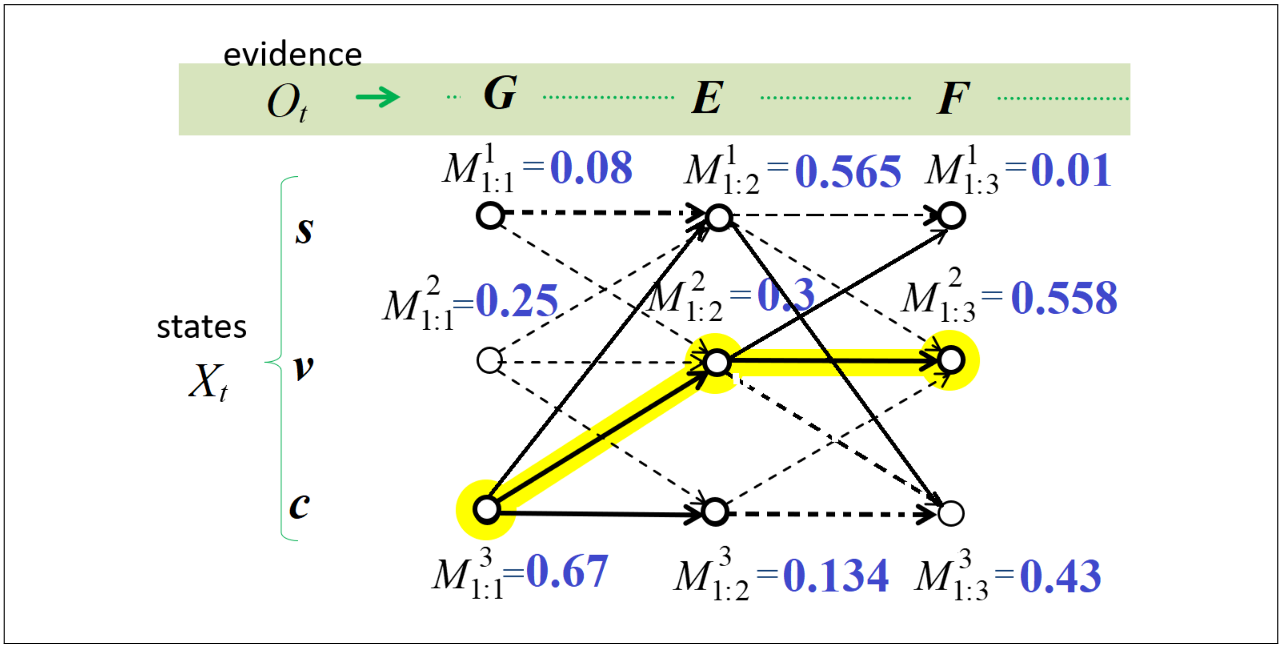
Similarly, as we did before, we do now at time instant 3. We obtain the following:

$$\begin{aligned} M^1_{1:3} &= \alpha P(O_3 = F|X_2 = S) \max\{P(S|S)M^1_{1:2}, P(S|V)M^2_{1:2}, P(S|C)M^3_{1:2}\} \\ &= 0.05 \max\{\underbrace{0.1 \times 0.565}_{0.0565}, \underbrace{0.2 \times 0.3}_{\mathbf{0.06}}, \underbrace{0.2 \times 0.135}_{0.027}\}\alpha = 0.003\alpha \\ M^2_{1:3} &= \alpha P(O_3 = F|X_2 = V) \max\{P(V|S)M^1_{1:2}, P(V|V)M^2_{1:2}, P(V|C)M^3_{1:2}\} \\ &= 0.7 \max\{\underbrace{0.1 \times 0.565}_{0.0565}, \underbrace{0.7 \times 0.3}_{\mathbf{0.21}}, \underbrace{0.7 \times 0.135}_{0.0945}\}\alpha = 0.147\alpha \\ M^3_{1:3} &= \alpha P(O_3 = F|X_2 = C) \max\{P(C|S)M^1_{1:2}, P(C|V)M^2_{1:2}, P(C|C)M^3_{1:2}\} \\ &= 0.25 \max\{\underbrace{0.8 \times 0.565}_{\mathbf{0.452}}, \underbrace{0.1 \times 0.3}_{0.03}, \underbrace{0.1 \times 0.135}_{0.0135}\}\alpha = 0.113\alpha \end{aligned}$$

By normalizing the obtained values, where $\alpha = 3.8$, we have that

$$\mathbf{m}_{1:3} = \langle 0.01, 0.558, 0.43 \rangle.$$

Obviously, the state $VideoGaming$ is most probable at the end, so tracing back we obtain the most likely sequence: $X_1 = Chatting$, $X_2 = VideoGaming$, $X_3 = VideoGaming$.


evidence
O_t
G
E
F
states
X_t
s
v
c
$M^1_{1:1}=0.08$
$M^1_{1:2}=0.565$
$M^1_{1:3}=0.01$
$M^2_{1:1}=0.25$
$M^2_{1:2}=0.3$
$M^2_{1:3}=0.558$
$M^3_{1:1}=0.67$
$M^3_{1:2}=0.134$
$M^3_{1:3}=0.43$