Exercise 2

Problem 1 – Hidden Markov Models

a) *Formulate the information given above as a hidden Markov model, and provide the complete probability tables for the model.*

We know that the state variable X_t denotes whether there is fish nearby on day *t*.
Let e_t (evidence variables) denote whether there are birds nearby on day *t*.
Observe that X_t is the parent of nodes X_{t+1} and e_t for all $t \neq 0$. As we have a Bayesian Network with a single (hidden) state variable and an observation variable for each state variable, this is a Hidden Markov Model.

We have the following probability tables, from the assignment

$P(X_0)$	0.5
$P(\neg X_0)$	0.5

$P(X_t \mid X_{t-1})$	0.8
$P(\neg X_t \mid X_{t-1})$	0.2
$P(X_t \mid \neg X_{t-1})$	0.3
$P(\neg X_t \mid \neg X_{t-1})$	0.7

$P(e_t \mid X_t)$	0.75
$P(\neg e_t \mid X_t)$	0.25
$P(e_t \mid \neg X_t)$	0.2
$P(\neg e_t \mid \neg X_t)$	0.8

b) *Compute $P(X_t \mid e_{1:t})$, for t = 1, ..., 6. What kind of operation is this (filtering, prediction, smoothing, likelihood of the evidence, or most likely sequence)? Describe in words what kind of information this operation provides us.*

Probability distribution, P(X1 | e{1:1}):

X_1(No fish)	0.1791
X_1(Fish)	0.8209

Probability distribution, P(X2 | e{1:2}):

X_2(No fish)	0.0980
X_2(Fish)	0.9020

Probability distribution, P(X3 | e{1:3}):

X_3(No fish)	0.5148
X_3(Fish)	0.4852

Probability distribution, P(X4 | e{1:4}):

X_4(No fish)	0.1835
X_4(Fish)	0.8165

Probability distribution, P(X5 | e{1:5}):

X_5(No fish)	0.5687
X_5(Fish)	0.4313

Probability distribution, P(X6 | e{1:6}):

X_6(No fish)	0.2003
X_6(Fish)	0.7997

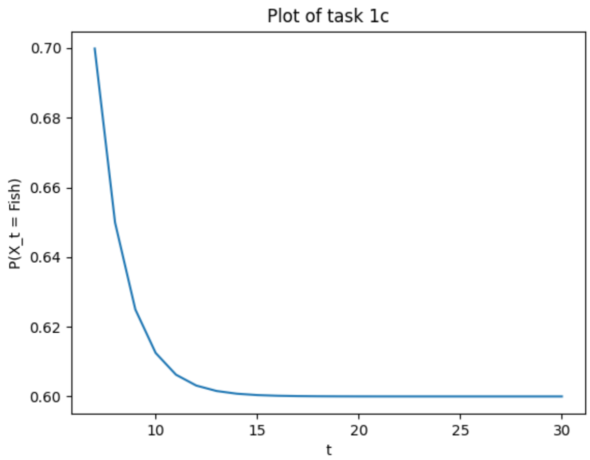
This is a filtering operation, as it finds the distribution over the "current" hidden variable, given evidence up to and including current day.

This information provides us probabilities of the current status of the lake. It tells us how likely it is that the lake contains fish at the current day, given all the observations made until now. As an example, given the observations made on day 1-6, the chance of the lake containing fish is 79.97 %.

c) *Compute $P(X_t \mid e_{1:6})$, for t = 7, ..., 30. What kind of operation is this (filtering, prediction, smoothing, likelihood of the evidence, or most likely sequence)? Describe in words what kind of information this operation provides us. What happens to the distribution in Equation (2) as t increases?*

```
Probability distribution, P(X7 | e{1:6}):
+-------------+-------------+
| X_7(No fish) |   0.3001   |
+-------------+-------------+
|  X_7(Fish)   |   0.6999   |
+-------------+-------------+

Probability distribution, P(X8 | e{1:6}):
+-------------+-------------+
| X_8(No fish) |   0.3501   |
+-------------+-------------+
|  X_8(Fish)   |   0.6499   |
+-------------+-------------+

Probability distribution, P(X9 | e{1:6}):
+-------------+-------------+
| X_9(No fish) |   0.3750   |
+-------------+-------------+
|  X_9(Fish)   |   0.6250   |
+-------------+-------------+

Probability distribution, P(X10 | e{1:6}):
+-------------+-------------+
| X_10(No fish) |   0.3875  |
+-------------+-------------+
|  X_10(Fish)   |   0.6125  |
+-------------+-------------+

Probability distribution, P(X11 | e{1:6}):
+-------------+-------------+
| X_11(No fish) |   0.3938  |
+-------------+-------------+
|  X_11(Fish)   |   0.6062  |
+-------------+-------------+

Probability distribution, P(X12 | e{1:6}):
+-------------+-------------+
| X_12(No fish) |   0.3969  |
+-------------+-------------+
|  X_12(Fish)   |   0.6031  |
+-------------+-------------+

Probability distribution, P(X13 | e{1:6}):
+-------------+-------------+
| X_13(No fish) |   0.3984  |
+-------------+-------------+
|  X_13(Fish)   |   0.6016  |
+-------------+-------------+

Probability distribution, P(X14 | e{1:6}):
+-------------+-------------+
| X_14(No fish) |   0.3992  |
+-------------+-------------+
|  X_14(Fish)   |   0.6008  |
+-------------+-------------+

Probability distribution, P(X15 | e{1:6}):
+-------------+-------------+
| X_15(No fish) |   0.3996  |
+-------------+-------------+
|  X_15(Fish)   |   0.6004  |
+-------------+-------------+

Probability distribution, P(X16 | e{1:6}):
+-------------+-------------+
| X_16(No fish) |   0.3998  |
+-------------+-------------+
|  X_16(Fish)   |   0.6002  |
+-------------+-------------+

Probability distribution, P(X17 | e{1:6}):
+-------------+-------------+
| X_17(No fish) |   0.3999  |
+-------------+-------------+
|  X_17(Fish)   |   0.6001  |
+-------------+-------------+

Probability distribution, P(X18 | e{1:6}):
+-------------+-------------+
| X_18(No fish) |   0.4000  |
+-------------+-------------+
|  X_18(Fish)   |   0.6000  |
+-------------+-------------+
```

P(0) = 0.4 and P(1) = 0.6 for all remaining distributions. All distributions are plotted below:



This is an example of a prediction-operation.

This information tells us the chances of a given state (in this case, the probability of the pond containing fish) in the future, based on the evidence we have seen.

As t increases, we can see that the model converges on $P(X_t) = 0.6$. This indicates that as the pond stabilizes on a 60% chance of containing fish. As the time without observations increases, this is the long-term stable probability for the pond.

d) Compute $P(X_t \mid e_{1:6})$, for t = 0, ..., 5. What kind of operation is this (filtering, prediction, smoothing, likelihood of the evidence, or most likely sequence)? Describe in words what kind of information this operation provides us.

```
Probability distribution, P(X0 | e{1:6}):
+--------------+--------------+
| X_0(No fish) |    0.3351    |
+--------------+--------------+
|   X_0(Fish)  |    0.6649    |
+--------------+--------------+

Probability distribution, P(X1 | e{1:6}):
+--------------+--------------+
| X_1(No fish) |    0.1236    |
+--------------+--------------+
|   X_1(Fish)  |    0.8764    |
+--------------+--------------+

Probability distribution, P(X2 | e{1:6}):
+--------------+--------------+
| X_2(No fish) |    0.1342    |
+--------------+--------------+
|   X_2(Fish)  |    0.8658    |
+--------------+--------------+

Probability distribution, P(X3 | e{1:6}):
+--------------+--------------+
| X_3(No fish) |    0.4021    |
+--------------+--------------+
|   X_3(Fish)  |    0.5979    |
+--------------+--------------+

Probability distribution, P(X4 | e{1:6}):
+--------------+--------------+
| X_4(No fish) |    0.2334    |
+--------------+--------------+
|   X_4(Fish)  |    0.7666    |
+--------------+--------------+

Probability distribution, P(X5 | e{1:6}):
+--------------+--------------+
| X_5(No fish) |    0.4292    |
+--------------+--------------+
|   X_5(Fish)  |    0.5708    |
+--------------+--------------+
```

This is an example of a smoothing-operation.

The operation tells us the most likely state in past variables, given all evidence up to the current date. In this example, it tells us the probability for the pond containing fish for each day in the past, given all the evidence we currently have.

It is different from filtering, as it takes in observations “in the future” (relative to the node it calculates).

e) Compute $\arg\max_{x_1,\ldots,x_{t-1}} P(x_1, \ldots, x_{t-1}, X_t \mid e_{1:t})$*, for t = 1, ..., 6. What kind of operation is this (filtering, prediction, smoothing, likelihood of the evidence, or most likely sequence)? Describe in words what kind of information this operation provides us.*

```
arg max P(X1 | e{1}):
The most likely path is: Fish
With a probability of 0.37500

arg max P(x{1}, ..., x{1}, X2 | e{1:2}):
The most likely path is: Fish -> Fish
With a probability of 0.22500

arg max P(x{1}, ..., x{2}, X3 | e{1:3}):
The most likely path is: Fish -> Fish -> Fish
With a probability of 0.04500

arg max P(x{1}, ..., x{3}, X4 | e{1:4}):
The most likely path is: Fish -> Fish -> Fish -> Fish
With a probability of 0.02700

arg max P(x{1}, ..., x{4}, X5 | e{1:5}):
The most likely path is: Fish -> Fish -> Fish -> Fish -> Fish
With a probability of 0.00540

arg max P(x{1}, ..., x{5}, X6 | e{1:6}):
The most likely path is: Fish -> Fish -> Fish -> Fish -> Fish -> Fish
With a probability of 0.00324
```

The operation done in this task is finding the most likely sequence/path.
This provides us the most likely combination of states leading up to the current state given the evidence we have. In this task, it shows us the most likely sequence of fish/no-fish leading to today, given the evidence we have seen thus far.
The actual calculations above show us that the most likely sequence leading up to all days 1-6 is that the pond has always contained fish.

I realize that I have the wrong answer on this task, but I have worked on this for at least 24 hours now (not straight, but almost), and I simply cannot seem to work it out. I have tried implementing the Viterbi algorithm three times, and none of them worked. My issue arises from the fact that it is unclear to me what is expected from, and how to programmatically calculate, $\max_{x_t}(P(X_{t+1}|x_t) \max_{x_1,\ldots,x_{t-1}} P(x_1, \ldots, x_t|e_{1:t})$. The above are the results from *viterbi()* in *WRONGexercise2.py*

Problem 2 – Dynamic Bayesian Networks

a) *Formulate the information given above as a dynamic Bayesian network and provide the complete probability tables for the model.*

We can see, from the assignment, that we have one state variable – whether there are animals nearby or not. Let X_t represent this state variable on day *t*.
The state variable is only dependent on the state on the previous day, and the evidence variables are only dependent on the current state. This means we have a first-order Markov process.
There are two ways of formulating this – we can either combine the observations to one evidence variable (thus creating a Hidden Markov Model), or by having two evidence variables. I will formulate both, then use whichever makes for the easiest calculations in the following problems.

We have the following probability tables, from the assignment

$P(X_0)$	0.7

$P(X_t \mid X_{t-1})$	0.8
$P(X_t \mid \neg X_{t-1})$	0.3

Let a_t denote whether there were animal tracks on day *t*, and f_t denote whether the food was gone on day *t*. We then have

$P(a_t \mid X_t)$	0.7
$P(a_t \mid \neg X_t)$	0.2

$P(f_t \mid X_t)$	0.3
$P(f_t \mid \neg X_t)$	0.1

Alternatively, we can let $e_t = (a, f)$ have four states, and represent the combination of tracks and food, where *a* indicates whether there are tracks and *f* represents if the food is gone. Because we are given that animal tracks and food gone are conditionally independent given animals nearby, we can see that

$$P\big((i,j)\big|X_t\big) = P(i|X_t) * P(j|X_t)$$

We can then calculate the probability for all e_t given X_t:

	Calculation	**P**
$P(e_t = (0,0) \mid X_t)$	$= (1 - P(a_t \mid X_t))(1 - P(f_t \mid X_t)) = 0.3 * 0.7$	0.21
$P(e_t = (0,0) \mid \neg X_t)$	$= (1 - P(a_t \mid \neg X_t))(1 - P(f_t \mid \neg X_t)) = 0.8 * 0.9$	0.72
$P(e_t = (0,1) \mid X_t)$	$= (1 - P(a_t \mid X_t))(P(f_t \mid X_t)) = 0.3 * 0.3$	0.09
$P(e_t = (0,1) \mid \neg X_t)$	$= (1 - P(a_t \mid \neg X_t))P(f_t \mid \neg X_t) = 0.8 * 0.1$	0.08
$P(e_t = (1,0) \mid X_t)$	$= P(a_t \mid X_t)(1 - P(f_t \mid X_t)) = 0.7 * 0.7$	0.49
$P(e_t = (1,0) \mid \neg X_t)$	$= P(a_t \mid \neg X_t)(1 - P(f_t \mid \neg X_t)) = 0.2 * 0.9$	0.18
$P(e_t = (1,1) \mid X_t)$	$= P(a_t \mid X_t)P(f_t \mid X_t) = 0.7 * 0.3$	0.21
$P(e_t = (1,1) \mid \neg X_t)$	$= P(a_t \mid \neg X_t)(P(f_t \mid \neg X_t) = 0.2 * 0.1$	0.02

b) *Compute $P(X_t|e_{1:t})$, for t = 1, 2, 3, 4.*

	$\boldsymbol{P(x_t\|e_{1:t})}$	$\boldsymbol{P(\neg x_t\|e_{1:t})}$
$\boldsymbol{t=1}$	0.951	0.049
$\boldsymbol{t=2}$	0.795	0.205
$\boldsymbol{t=3}$	0.4021	0.5979
$\boldsymbol{t=4}$	0.7322	0.2678

Because we have formulated this as a Hidden Markov Model, we can apply known methods for operations such as filtering, which this is an example of. We therefore apply the FORWARD-algorithm to this problem:

$$P(X_t|\, e_{1:t}) = \alpha P(e_t \mid X_t) \sum_{x_{t-1}} P(X_t \mid x_{t-1}) P(x_{t-1}|\, e_{1:t})$$

We first find prior probabilities:

$$P(X_0) = [0.7, 0.3]$$

We then find the rest

$$P(X_1|\, e_{1:1}) = \alpha P(e_1 \mid X_1) \sum_{x_0} P(X_1 \mid x_0) P(x_0)$$

We know $P(e_1 \mid X_t)$ directly from the observation table, $P(X_t \mid x_0)$ from the transition table and $P(x_0)$ from the calculations above. We know from the assignment that $e_1 = (1,1)$

$$= \alpha * P(e_1 = (1,1)|\, X_t) * (P(X_1 \mid x_0)P(x_0) + P(X_1|\neg x_0)P(\neg x_0))$$

$$= \alpha * [0.21, 0.02] * ([0.8, 0.2] * 0.7 + [0.3, 0.7] * 0.3)$$

$$= \alpha * [0.21, 0.02] * [0.65, 0.35]$$

$$= \alpha * [0.1365, 0.007] = [\mathbf{0.951}, \mathbf{0.049}]$$

Prediction from t = 1 to t = 2:

$$P(X_2|e_1) = \sum_{x_1} P(X_2|x_1)\, P(x_1|e_1)$$

$$= [0.8, 0.2] * 0.951 + [0.3, 0.7] * 0.049 = [0.7755, 0.2245]$$

Updating with $e_2 = (0, 1)$:

$$P(X_2|e_{1:2}) = \alpha P(e_2 \mid X_2) \sum_{x_1} P(X_2 \mid x_1) P(x_1|e_1)$$

$$= \alpha * [0.09, 0.08] * ([0.8, 0.2] * 0.951 + [0.3, 0.7] * 0.049)$$

$$= \alpha * [0.09, 0.08] * [0.7755, 0.2245]$$

$$= \alpha * [0.0698, 0.0180] = [\mathbf{0.795}, \mathbf{0.205}]$$

Prediction from t = 2 to t = 3:

$$P(X_3|e_{1:2}) = \sum_{x_2} P(X_3|x_2)\, P(x_2|e_{1:2})$$

$$= [0.8, 0.2] * 0.795 + [0.3, 0.7] * 0.205 = [0.6975, 0.3025]$$

Updating with $e_3 = (0, 0)$:

$$P(X_3|e_{1:3}) = \alpha P(e_3|X_3) * P(X_3|e_{1:2})$$

$$= \alpha * [0.21, 0.72] * [0.6975, 0.3025]$$

$$= \alpha * [0.1465, 0.2178] = [\mathbf{0.4021}, \mathbf{0.5979}]$$

Prediction from t = 3 to t = 4:

$$P(X_4|e_{1:3}) = \sum_{x_3} P(X_4|x_3)\, P(x_3|e_{1:3})$$

$$= [0.8, 0.2] * 0.4021 + [0.3, 0.7] * 0.5979 = [0.5011, 0.4989]$$

Updating with $e_4 = (1, 0)$:

$$P(X_4|e_{1:4}) = \alpha P(e_4|X_4) * P(X_4|e_{1:3})$$

$$= \alpha * [0.49, 0.18] * [0.5011, 0.4989]$$

$$= \alpha * [0.2455, 0.0898] = [\mathbf{0.7322}, \mathbf{0.2678}]$$

c) Compute $P(X_t|e_{1:4})$, for t = 5, 6, 7, 8

	$\boldsymbol{P(x_t|e_{1:4})}$	$\boldsymbol{P(\neg x_t|e_{1:4})}$
$\boldsymbol{t=5}$	0.6661	0.3339
$\boldsymbol{t=6}$	0.6331	0.3669
$\boldsymbol{t=7}$	0.6166	0.3834
$\boldsymbol{t=8}$	0.6083	0.3917

We once again use formulas for Hidden Markov Models to conduct operations, which this prediction is part of. We know $P(X_4|e_{1:4}) = [0.7322, 0.2678]$ from problem (b), so we use the following formula for prediction:

$$P(X_{t+k+1}|e_{1:t}) = \sum_{x_{t+k}} P(X_{t+k+1}|x_{t+k})P(x_{t+k}|e_{1:t})$$

Let t = 4, and k = 0:

$$P(X_5|e_{1:4}) = \sum_{x_4} P(X_5|x_4)P(x_4|e_{1:4})$$

$$= [0.8, 0.2] * 0.7322 + [0.3, 0.7] * 0.2678 = [\mathbf{0.6661}, \mathbf{0.3339}]$$

Let k = 1

$$P(X_6|e_{1:4}) = \sum_{x_5} P(X_6|x_5)P(x_5|e_{1:4})$$

$$= [0.8, 0.2] * 0.6661 + [0.3, 0.7] * 0.3339 = [\mathbf{0.6331}, \mathbf{0.3669}]$$

Let k = 2

$$P(X_7|e_{1:4}) = \sum_{x_6} P(X_7|x_6)P(x_6|e_{1:4})$$

$$= [0.8, 0.2] * 0.6331 + [0.3, 0.7] * 0.3669 = [\mathbf{0.6166}, \mathbf{0.3834}]$$

Let k = 3

$$P(X_8|e_{1:4}) = \sum_{x_7} P(X_8|x_7)P(x_7|e_{1:4})$$

$$= [0.8, 0.2] * 0.6166 + [0.3, 0.7] * 0.3834 = [\mathbf{0.6083}, \mathbf{0.3917}]$$

d) *By forecasting further and further into the future, you should see that the probability converges towards a fixed point. Verify that* $\lim_{t\to\infty} P(X_t|e_{1:4}) = [0.6, 0.4]$.

The distribution which a Markov chain transforms into when $t \to \infty$ is called its stationary distribution. Let the distribution be called π. Then the stationary distribution has been reached when $\pi_t = \pi_{t+1}$, or $\pi T = \pi$ for a transition matrix T.

We have our transition matrix $T = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$, and are to check that $\pi = [0.6, 0.4]$. We calculate that

$$\begin{aligned} &\pi T \\ &= [0.6, 0.4] * \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \\ &= [0.6 * 0.8 + 0.4 * 0.3, 0.6 * 0.2 + 0.4 * 0.7] \\ &= [0.48 + 0.12, 0.12 + 0.28] \\ &= [0.6, 0.4] = \pi \end{aligned}$$

Thus, the distribution $\pi = [0.6, 0.4]$ is a stationary distribution for the system, and

$$\lim_{t\to\infty} P(X_t|e_{1:4}) = [0.6, 0.4]$$

□

e) Compute $P(X_t|e_{1:4})$*, for t = 0, 1, 2, 3.*

	$\boldsymbol{P(x_t|e_{1:4})}$	$\boldsymbol{P(\neg x_t|e_{1:4})}$
$\boldsymbol{t = 0}$	0.9187	0.0813
$\boldsymbol{t = 1}$	0.9441	0.0559
$\boldsymbol{t = 2}$	0.7240	0.2760
$\boldsymbol{t = 3}$	0.5133	0.4867

We will, once again, use formulas for Hidden Markov Model to do our operations, this time we will use them do to SMOOTHING. The formula for this is as follows:

$$P(X_t|e_{1:k}) = \alpha\, P(X_t|e_{1:t})\, P(e_{t+1:k}|X_t)$$

Insert k = 4, and we get

$$P(X_t|e_{1:4}) = \alpha\, P(X_t|e_{1:t})\, P(e_{t+1:4}|X_t)$$

We know $P(X_t|e_{1:t})$ from task (b) (and the assignment for t = 0):

	$\boldsymbol{P(x_t|e_{1:t})}$	$\boldsymbol{P(\neg x_t|e_{1:t})}$
$\boldsymbol{t = 0}$	0.7	0.3
$\boldsymbol{t = 1}$	0.951	0.049
$\boldsymbol{t = 2}$	0.795	0.205
$\boldsymbol{t = 3}$	0.4021	0.5979
$\boldsymbol{t = 4}$	0.7322	0.2678

To find $P(e_{t+1:4}|X_t)$, we use the BACKWARD algorithm:

$$P(e_{t+1:k}|X_t) = \sum_{x_{t+1}} P(e_{t+1}|x_{t+1})\, P(e_{t+2:k}|x_{t+1})\, P(x_{t+1}|X_t)$$

With k = 4:

$$P(e_{t+1:4}|X_t) = \sum_{x_{t+1}} P(e_{t+1}|x_{t+1})\, P(e_{t+2:4}|x_{t+1})\, P(x_{t+1}|X_t)$$

As this algorithm moves backwards, we will start by finding the probability when t = 3.

Observe that $e_4 = (1, 0)$:

$$P(e_{4:4}|X_3) = \sum_{x_4} P(e_4|x_4)\, P(e_{5:4}|x_4)\, P(x_4|X_3)$$

$$= P(e_4|x_4)\, P(e_{5:4}|x_4)\, P(x_4|X_3) + P(e_4|\neg x_4)\, P(e_{5:4}|\neg x_4)\, P(\neg x_4|X_3)$$

$$= \ 0.49 * 1 * [0.8, 0.3] + 0.18 * 1 * [0.2, 0.7]$$

$$= [0.428, 0.273]$$

Thus:

$$P(X_3|e_{1:4}) = \alpha\, P(X_3|e_{1:3})\, P(e_{4:4}|X_3)$$
$$= \alpha * [0.4021, 0.5979] * [0.428, 0.273]$$
$$= \alpha * [0.1721, 0.1632] = [\mathbf{0.5133}, \mathbf{0.4867}]$$

Iterate the BACKWARD algorithm the whole way, until t = 0:

t = 2, $e_3 = (0, 0)$

$$P(e_{3:4}|X_2) = \sum_{x_3} P(e_3|x_3)\, P(e_{4:4}|x_3)\, P(x_3|X_2)$$
$$= P(e_3|x_3)\, P(e_{4:4}|x_3)\, P(x_3|X_2) + P(e_3|\neg x_3)\, P(e_{4:4}|\neg x_3)\, P(\neg x_3|X_2)$$
$$= 0.21 * 0.428 * [0.8, 0.3] + 0.72 * 0.273 * [0.2, 0.7]$$
$$= [0.1112, 0.1646]$$

t = 1, $e_2 = (0, 1)$

$$P(e_{2:4}|X_1) = \sum_{x_2} P(e_2|x_2)\, P(e_{3:4}|x_2)\, P(x_2|X_1)$$
$$= P(e_2|x_2)\, P(e_{3:4}|x_2)\, P(x_2|X_1) + P(e_2|\neg x_2)\, P(e_{3:4}|\neg x_2)\, P(\neg x_2|X_1)$$
$$= 0.09 * 0.1112 * [0.8, 0.3] + 0.08 * 0.1646 * [0.2, 0.7]$$
$$= [0.01064, 0.01222]$$

t = 0, $e_1 = (1,1)$

$$P(e_{1:4}|X_0) = \sum_{x_1} P(e_1|x_1)\, P(e_{2:4}|x_1)\, P(x_1|X_0)$$
$$= P(e_1|x_1)\, P(e_{2:4}|x_1)\, P(x_1|X_0) + P(e_1|\neg x_1)\, P(e_{2:4}|\neg x_1)\, P(\neg x_1|X_0)$$
$$= 0.21 * 0.01064 * [0.8, 0.3] + 0.02 * 0.01222 * [0.2, 0.7]$$
$$= [0.001836, 0.0008414]$$

We can then find

$$P(X_2|e_{1:4}) = \alpha\, P(X_2|e_{1:2})\, P(e_{3:4}|X_2)$$
$$= \alpha * [0.795, 0.205] * [0.1112, 0.1646]$$
$$= \alpha * [0.0884, 0.0337] = [\mathbf{0.7240}, \mathbf{0.2760}]$$

$$P(X_1|e_{1:4}) = \alpha\, P(X_1|e_{1:1})\, P(e_{2:4}|X_1)$$
$$= \alpha * [0.951, 0.049] * [0.01064, 0.01222]$$
$$= \alpha * [0.01012, 0.0005988] = [\mathbf{0.9441}, \mathbf{0.0559}]$$

$P(X_0|e_{1:4}) = \alpha\ P(X_0|)\ P(e_{1:4}|X_0)$

$= \alpha * [0.7, 0.3] * [0.001836, 0.0008414]$

$= \alpha * [0.002852, 0.0002524] = [\mathbf{0.9187}, \mathbf{0.0813}]$