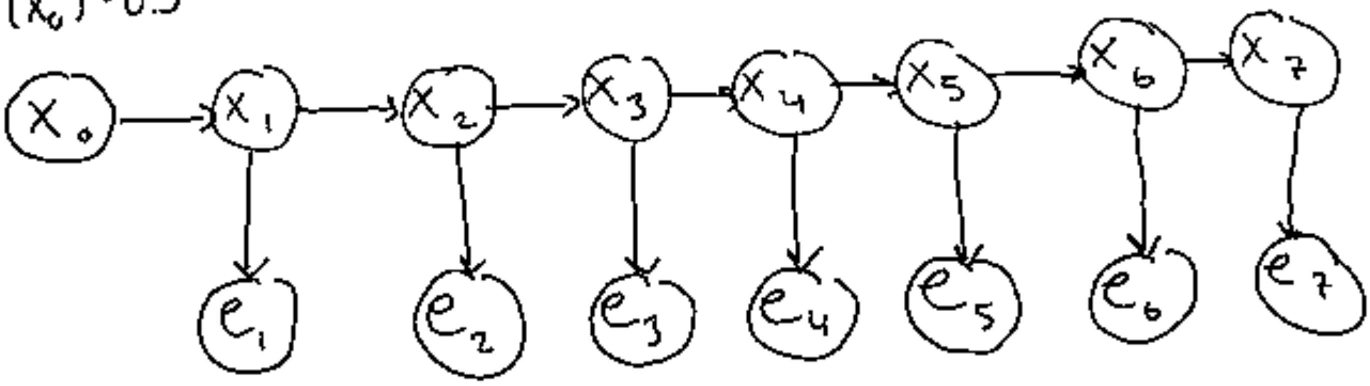
TDT4171 Assignment 2

Hanna Waage Hjelmeland, MTTK

1. a) X_t: fish nearby on day t

E_t: set of observations of birds nearby

$P(x_0) = 0.5$



X_{t-1}	$P(X_t)$	$P(\overline{X_t})$
T	0.8	0.2
F	0.3	0.7

X_t	$P(E_t)$	$P(\overline{E_t})$
T	0.75	0.25
F	0.2	0.8

Task 1

1b) Define functions

```
import numpy as np

def normalize(array):
    return array*(1/np.sum(array))

# True system
p0 = np.array([0.5, 0.5])
T = np.array([[0.8, 0.3], [0.2, 0.7]])
O_true = np.diag([0.75, 0.2])
O_false = np.diag([0.25, 0.8])
O = np.array([O_true, O_false])
evidence = np.array([1, 1, 0, 1, 0, 1])

def forward(e, current):
    return normalize(O[e] @ T.T @ current)



if __name__ == '__main__':
    current = p0
    for ind, e in enumerate(evidence):
        ind = ind + 1
        current = forward(e, current)
        print(f"P(X{(ind)} | e1:{ind}) = {current}")
```

```
P(X1 | e1:1) = [0.23809524 0.76190476]
P(X2 | e1:2) = [0.15050167 0.84949833]
P(X3 | e1:3) = [0.62983746 0.37016254]
P(X4 | e1:4) = [0.2872692 0.7127308]
P(X5 | e1:5) = [0.70471478 0.29528522]
P(X6 | e1:6) = [0.31764116 0.68235884]
```

This operation is called filtering. This is the act of figuring out what we think of the situation at time = t, given all our observations up until that point.

1c)

```
def predict(last_val, i):
    if i == evidence.shape[0]:
        return last_val
    res = normalize(T.T @ predict(last_val, i-1))
    print(f"P(X{(i)} | e1:{evidence.shape[0]}) = {res}")
    return res

if __name__ == '__main__':
    # b)
    current = p0
    for ind, e in enumerate(evidence):
        current = forward(e, current)

    # c)
    predict(current, 30)
```

```
P(X7 | e1:6) = [0.40536923 0.59463077]
P(X8 | e1:6) = [0.45177184 0.54822816]
P(X9 | e1:6) = [0.47565106 0.52434894]
P(X10 | e1:6) = [0.48776595 0.51223405]
P(X11 | e1:6) = [0.49386797 0.50613203]
P(X12 | e1:6) = [0.49693022 0.50306978]
P(X13 | e1:6) = [0.49846417 0.50153583]
P(X14 | e1:6) = [0.49923185 0.50076815]
P(X15 | e1:6) = [0.49961586 0.50038414]
P(X16 | e1:6) = [0.49980792 0.50019208]
P(X17 | e1:6) = [0.49990396 0.50009604]
P(X18 | e1:6) = [0.49995198 0.50004802]
P(X19 | e1:6) = [0.49997599 0.50002401]
P(X20 | e1:6) = [0.49998799 0.50001201]
P(X21 | e1:6) = [0.499994 0.500006]
P(X22 | e1:6) = [0.499997 0.500003]
P(X23 | e1:6) = [0.4999985 0.5000015]
P(X24 | e1:6) = [0.49999925 0.50000075]
P(X25 | e1:6) = [0.49999962 0.50000038]
P(X26 | e1:6) = [0.49999981 0.50000019]
P(X27 | e1:6) = [0.49999991 0.50000009]
P(X28 | e1:6) = [0.49999995 0.50000005]
P(X29 | e1:6) = [0.49999998 0.50000002]
P(X30 | e1:6) = [0.49999999 0.50000001]
```

This operation is called predicting. This means peeking into the future and predicting the most likely outcome, based on the observations we have availiable. We see that the distribution converges towards [0.5, 0.5] as t increases. This makes sense, as the further we get from our observations, the probability looks more and more like the prior probability of the model.

1d)

```
def backward(e, current):
    return T @ O[e] @ current

def smooth(k):

    t = evidence.shape[0]
    current_b = np.ones(2)
    for i in range(t-k):
        current_b = backward(evidence[t-1-i], current_b)

    current_f = p0
    for i in range(k):
        current_f = forward(evidence[i], current_f)


    return normalize(current_f * current_b)
```

```
if __name__ == '__main__':

    t = evidence.shape[0]
    for i in range(t):
        ind = i + 1
        print(f"P(X{(ind)} | e1:{t}) = {smooth(ind)}")
```

```
P(X1 | e1:6) = [0.24260158 0.75739842]
P(X2 | e1:6) = [0.28458721 0.71541279]
P(X3 | e1:6) = [0.62182416 0.37817584]
P(X4 | e1:6) = [0.44452214 0.55547786]
P(X5 | e1:6) = [0.63254888 0.36745112]
P(X6 | e1:6) = [0.31764116 0.68235884]
```

This operation is called smoothing, and it gives us an updated view on how the situation looked a while ago, based on all the information we have collected. Thus, we "smooth" the information about the observations out, making them influence more than just the one parent node. We see that this gives more even resulting distributions.

1e)

```
def forward_mle(e, current):
    return np.diag(O[e] * (T.T @ current).max(0))
```

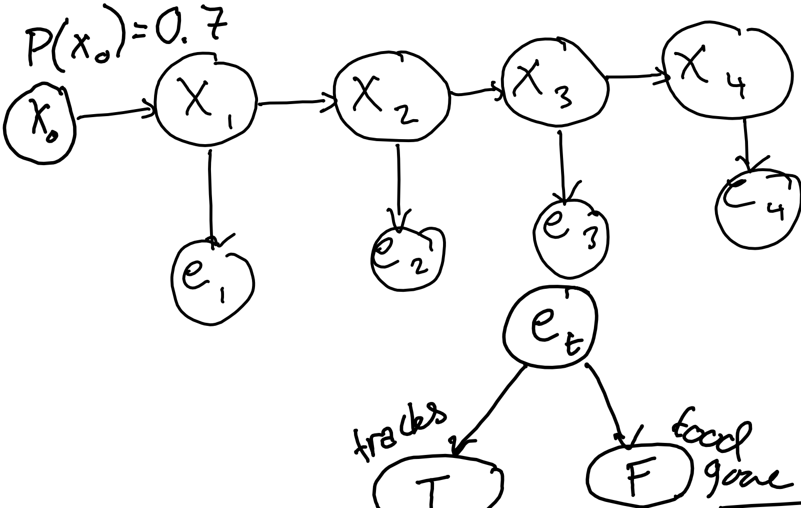
```
if __name__ == '__main__':
    current = forward(evidence[0], p0)
    t = evidence.shape[0]
    path = np.zeros(t)
    for i, e in enumerate(evidence):
        ind = i + 1
        current = forward_mle(e, current)
        p = not np.argmax(current)
        path[i] = p
        print(f"argmax P(x1, x2, ..., X{(ind)} | e1:{ind}) = {path[0:ind]}")
```

```
argmax P(x1, x2, ..., X1 | e1:1) = [0.]
argmax P(x1, x2, ..., X2 | e1:2) = [0. 0.]
argmax P(x1, x2, ..., X3 | e1:3) = [0. 0. 1.]
argmax P(x1, x2, ..., X4 | e1:4) = [0. 0. 1. 0.]
argmax P(x1, x2, ..., X5 | e1:5) = [0. 0. 1. 0. 1.]
argmax P(x1, x2, ..., X6 | e1:6) = [0. 0. 1. 0. 1. 0.]
```

This is the concept of Most Likely Explanation (MLE). This provides the sequence that is most likely given the sequence of observations. We see that in this case, the most likely explanation follows the observations of the birds, as predicted.

2.

X_t: animals nearby on day t.



X_{t-1}	$P(X_t)$
T	0.8
F	0.3

X_t	$P(T)$	$P(\overline{T})$
T	0.7	0.3
F	0.2	0.8

X_t	$P(F)$	$P(\overline{F})$
T	0.3	0.7
F	0.1	0.9

$$P(e_t=\{T,F\}) = P(T \mid X_t)\cdot P(F \mid X_t) = \langle 0.21, 0.02\rangle$$

$$P(e_t=\{T,\overline{F}\}) = \langle 0.49, 0.18\rangle$$

$$P(e_t=\{\overline{T},F\}) = \langle 0.09, 0.08\rangle$$

$$P(e_t=\{\overline{T},\overline{F}\}) = \langle 0.21, 0.72\rangle$$

b) $e_1 = \{T, F\}$

$P(x_1 | e_1)$

$= \alpha P(e_1 | x_1) \cdot \sum_{x_0} P(x_1 | x_0) \cdot P(x_0)$

$= \alpha \cdot \langle 0.21, 0.02 \rangle$

$\cdot (\langle 0.8, 0.2 \rangle \cdot 0.7$ $\leftarrow$ $x_0 = \text{true}$

$+ \langle 0.3, 0.7 \rangle \cdot 0.3)$ $\leftarrow$ $x_0 = \text{false}$

$= \alpha \cdot \langle 0.1365, 0.0007 \rangle$

$= \langle 0.995, 0.005 \rangle$

$$e_2 = \{\overline{T}, F\} \Rightarrow P(e_2|X_2) = \langle 0.09, 0.08\rangle$$

$$\Rightarrow P(X_2|e_{1:2}) = \alpha \cdot P(e_2|X_2)P(X_2|e_1)$$

$$= \alpha \cdot P(e_2|X_2) \cdot \sum_{x_1} P(X_2|x_1) \cdot P(x_1|e_1)$$

$$= \alpha \cdot P(e_2|X_2) \cdot$$

$$(\langle 0.8, 0.2\rangle \cdot 0.995 \leftarrow x_1 \text{ true}$$

$$+(\langle 0.3, 0.7\rangle \cdot 0.005 \leftarrow x_1 \text{ false}$$

$$= \alpha \cdot \langle 0.09, 0.08\rangle \cdot \langle 0.798, 0.203\rangle$$

$$= \alpha \cdot \langle 0.0718, 0.0162\rangle$$

$$= \underline{\underline{\langle 0.816, 0.184\rangle}}$$

$$e_3 = \{\overline{T}, \overline{F}\} \Rightarrow P(e_3|X_3) = \langle 0.21, 0.72\rangle$$

$$\Rightarrow P(X_3|e_{1:3}) = \alpha P(e_3|X_3) \sum_{x_2} P(X_3|X_2) \cdot P(X_2|e_2)$$

= ... same procedure ...

$$= \alpha \cdot \langle 0.21, 0.72\rangle\langle 0.708, 0.292\rangle$$

$$= \underline{\underline{\langle 0.4142, 0.5858\rangle}}$$

$e_4 = \{T, \bar{F}\} \Rightarrow P(e_4 | X_4) = \langle 0.49, 0.18 \rangle$

$P(X_4 | e_{1:4}) = \alpha P(e_4 | X_4) \sum_{x_3} P(X_4 | x_3) \cdot P(x_3 | e_3)$

$= \ldots = \alpha \cdot \langle 0.49, 0.18 \rangle \langle 0.507, 0.493 \rangle$

$= \underline{\underline{\langle 0.737, 0.263 \rangle}}$

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

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

$= \underline{\underline{\langle 0.669, 0.332 \rangle}}$

$P(X_6 | e_{1:4}) = \sum_{x_5} P(X_5 | x_4) \cdot P(x_4 | e_{1:4}) = \ldots$

$= \underline{\underline{\langle 0.6343, 0.3658 \rangle}}$

$P(X_7 | e_{1:4}) = \sum_{x_6} P(X_6 | x_5) \cdot P(x_5 | e_{1:4}) = \ldots$

$= \underline{\underline{\langle 0.617, 0.383 \rangle}}$

$P(X_8 | e_{1:4}) = \sum_{x_7} P(X_7 | x_6) \cdot P(x_6 | e_{1:4}) = \underline{\underline{\langle 0.609, 0.391 \rangle}}$

d)

$$\lim_{t \to \infty} P(X_t | e_{1:4})$$

Let $P_{i,j} = P(x_i | e_{1:4})$ $(x_i = j,\ j = \{1, 0\})$

$$\Rightarrow P(X_t | e_{1:4})$$

$$= \langle 0.8, 0.2 \rangle P_{t,1} + \langle 0.3, 0.7 \rangle P_{t,0}$$

$$= \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} P_{t,1} \\ P_{t,0} \end{bmatrix}$$

$$= \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} P_{t-1,1} \\ P_{t-1,0} \end{bmatrix}$$

$$= \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}^{t-4} \begin{bmatrix} P_{4,1} \\ P_{4,0} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}^{t-4}$$

$$\lim_{t \to \infty} \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}^{t-4} \begin{bmatrix} P_{4,1} \\ P_{4,0} \end{bmatrix} = \lim_{t \to \infty} \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}^{t}$$

$$= \begin{bmatrix} 0.6 & 0.6 \\ 0.4 & 0.4 \end{bmatrix} \Rightarrow \underline{\lim_{t \to \infty} P(X_t | e_{1:4}) = \langle 0.6, 0.4 \rangle} \quad \square$$

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

$$P(X_3 | e_{1:4}) = \alpha \cdot P(X_3 | e_{1:3}) \cdot \underbrace{P(e_{4:4} | X_3)}_{b_{4:4}}$$

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

$$= 0.49 \cdot 1 \cdot \langle 0.8, 0.2 \rangle \leftarrow x_4 = \text{true}$$

$$+ 0.18 \cdot 1 \cdot \langle 0.3, 0.7 \rangle$$

$$= \underline{\langle 0.397, 0.111 \rangle}$$

$$\Rightarrow P(X_3 | e_{1:4}) = \alpha \cdot P(X_3 | e_{1:3}) \cdot b_{4:4}$$

$$= \alpha \cdot \langle 0.414, 0.586 \rangle \cdot \langle 0.397, 0.111 \rangle$$

$$= \underline{\underline{\langle 0.717, 0.283 \rangle}}$$

$$P(X_2 | e_{1:4}) = \alpha \cdot P(X_2 | e_{1:2}) \cdot b_{3:4}$$

$$b_{3:4} = \sum_{x_3} P(e_3 | x_3) b_{4:4}(x_3) P(x_3 | X_2)$$

$$= 0.21 \cdot 0.717 \cdot \langle 0.8, 0.2 \rangle$$

$$+ 0.72 \cdot 0.283 \cdot \langle 0.3, 0.7 \rangle = \langle 0.182, 0.173 \rangle$$

$$\Rightarrow P(X_2 \mid e_{1:4}) = \alpha \cdot \langle 0.816, 0.184 \rangle \cdot b_{34}$$

$$= \underline{\underline{\langle 0.824, 0.177 \rangle}}$$

$$P(X_1 \mid e_{1:4}) = \alpha \cdot P(X_1 \mid e_1) \cdot b_{2:4}$$

$$b_{2:4} = \sum_{x_2} P(e_2 \mid x_2)\, b_{3:4}(x_2)\, P(x_2 \mid X_1)$$

$$= 0.09 \cdot 0.182 \cdot \langle 0.8, 0.2 \rangle$$

$$+ 0.08 \cdot 0.173 \cdot \langle 0.3, 0.7 \rangle = \langle 0.017, 0.013 \rangle$$

$$\Rightarrow P(X_1 \mid e_{1:4}) = \alpha \cdot \langle 0.995, 0.005 \rangle \langle 0.017, 0.013 \rangle$$

$$= \underline{\underline{\langle 0.996, 0.004 \rangle}}$$

$$P(X_0 \mid e_{1:4}) = \alpha\, P(X_0) \cdot b_{1:4}$$

$$b_{1:4} = \sum_{x_1} P(e_1 \mid x_1)\, b_{2:4}(x_1)\, P(x_1 \mid X_0)$$

$$= 0.21 \cdot 0.017 \cdot \langle 0.8, 0.2 \rangle$$

$$+ 0.02 \cdot 0.013 \cdot \langle 0.3, 0.7 \rangle = \langle 0.003, 0.009 \rangle$$

$$\Rightarrow P(X_0 \mid e_{1:4}) = \alpha \cdot \langle 0.7, 0.3 \rangle \cdot b_{1:4} = \underline{\underline{\langle 0.885, 0.115 \rangle}}$$