**A: $a_{ij} = P(X_{t+1} = j \mid X_t = i)$**

| $X_t \mid X_{t+1}$ | A | B | H | S |
|---|---|---|---|---|
| A | 0.6 | 0.1 | 0.1 | 0.2 |
| B | 0.0 | 0.3 | 0.2 | 0.5 |
| H | 0.8 | 0.1 | 0.0 | 0.1 |
| S | 0.2 | 0.0 | 0.1 | 0.7 |

**B: $b_{ik} = P(O_t = k \mid X_t = i)$**

| $X_t \mid O_t$ | p | e | b | l |
|---|---|---|---|---|
| A | 0.6 | 0.2 | 0.1 | 0.1 |
| B | 0.1 | 0.4 | 0.1 | 0.4 |
| H | 0.0 | 0.0 | 0.7 | 0.3 |
| S | 0.0 | 0.0 | 0.1 | 0.9 |

**$P(X_t = i)$ :**

| A | B | H | S |
|---|---|---|---|
| 0.4 | 0.2 | 0.1 | 0.3 |

**Find:**

**$P(O_t \mid A, B, P(X_t)) = ?$**

---

**Compute:**

0.24 + 0.02 + 0 + 0 =

$$P(O_t = p) = \underbrace{0.4}_{P(A)} \times \underbrace{0.6}_{P(p|A)} + \underbrace{0.2}_{P(B)} \times \underbrace{0.1}_{P(p|B)} + \underbrace{0.1}_{P(H)} \times \underbrace{0.0}_{P(p|H)} + \underbrace{0.3}_{P(S)} \times \underbrace{0.0}_{P(p|S)} = \underline{0.26}$$

$$P(O_t = e) = \underline{0.4} \times 0.2 + \underline{0.2} \times 0.4 + 0.1 \times \underline{0} + 0.3 \times \underline{0} = \underline{0.16}$$

$$P(O_t = b) = 0.4 \times \underline{0.1} + \underline{0.2} \times \underline{0.1} + 0.1 \times \underline{0.7} + 0.3 \times 0.1 = \underline{0.16}$$

$$P(O_t = l) = \underline{0.4} \times \underline{0.1} + \underline{0.2} \times \underline{0.4} + \underline{0.1} \times \underline{0.3} + \underline{0.3} \times \underline{0.9} = \underline{0.42}$$

$P(O_t) =$

| |
|---|
| 0.26 |
| 0.16 |
| 0.16 |
| 0.42 |

most likely $O_t$ = l

$\pi = P(X_1 = i)$ :

| A | B | H | S |
|---|---|---|---|
| 0.5 | 0.0 | 0.0 | 0.5 |

**observations / emissions:** $o_{1:4} = \{ l , p , p , b \}$

**Find:**

$$P( o_{1:4} \mid A, B, \pi ) = ?$$

**Element-wise product:**

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \odot \begin{bmatrix} d \\ e \\ f \end{bmatrix} = \begin{bmatrix} ad \\ be \\ cf \end{bmatrix}$$

**Compute:**

$O_1 =$ l

$$\alpha_1(i) = \underbrace{\begin{bmatrix} 0.5 \\ 0 \\ 0 \\ 0.5 \end{bmatrix}}^{\pi = P(X_1 = S_i)} \odot \overbrace{\begin{bmatrix} 0.1 \\ 0.4 \\ 0.3 \\ 0.9 \end{bmatrix}}^{P(l \mid S_i)} = \begin{bmatrix} 0.05 \\ 0 \\ 0 \\ 0.45 \end{bmatrix}$$

$O_2 =$ p

$$\alpha_2(i) = \begin{bmatrix} \overset{\alpha_1(A)}{0.05} \times \overset{P(A|A)}{0.6} + \overset{\alpha_1(B)}{0.0} \times \overset{P(A|B)}{0.0} + \overset{\alpha_1(H)}{0.0} \times \overset{P(A|H)}{0.8} + \overset{\alpha_1(S)}{0.45} \times \overset{P(A|S)}{0.2} \\ 0.05 \times 0.1 + 0 \times 0.3 + 0.0 \times 0.1 + 0.45 \times 0 \\ 0.05 \times 0.1 + 0 \times 0.2 + 0.0 \times 0 + 0.45 \times 0.1 \\ 0.05 \times 0.2 + 0 \times 0.5 + 0.0 \times 0.1 + 0.45 \times 0.7 \end{bmatrix} \odot \begin{bmatrix} \overset{P(p|A)}{0.6} \\ 0.1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \overset{\alpha_2(A)}{0.072} \\ 0.0005 \\ 0 \\ 0 \end{bmatrix}$$

$O_3 =$ p

$$\alpha_3(i) = \begin{bmatrix} 0.072 \times 0.6 + .0005 \times 0.0 + 0.0 \times 0.8 + 0.0 \times 0.2 \\ 0.072 \times 0.1 + .0005 \times 0.3 + 0.0 \times 0.1 + 0.0 \times 0.0 \\ 0.072 \times 0.1 + .0005 \times 0.2 + 0.0 \times 0 + 0.0 \times 0.1 \\ 0.072 \times 0.2 + .0005 \times 0.5 + 0.0 \times 0.1 + 0 \times 0.7 \end{bmatrix} \odot \begin{bmatrix} 0.6 \\ 0.1 \\ 0.0 \\ 0.0 \end{bmatrix} = \begin{bmatrix} .02592 \\ .000735 \\ 0.0 \\ 0.0 \end{bmatrix}$$

---

$O_4 =$ b

$$\alpha_4(i) = \begin{bmatrix} .02592 \times 0.6 + .000735 \times 0.0 + 0.0 \times 0.8 + 0.0 \times 0.2 \\ .02592 \times 0.1 + .000735 \times 0.3 + 0.0 \times 0.1 + 0.0 \times 0.0 \\ .02592 \times 0.1 + .000735 \times 0.2 + 0.0 \times 0.0 + 0.0 \times 0.1 \\ 0.0259 \times 0.2 + .000735 \times 0.5 + 0.0 \times 0.1 + 0.0 \times 0.7 \end{bmatrix} \odot \begin{bmatrix} .00156 \\ .000281 \\ 0 \\ 0 \end{bmatrix} =$$

$$\alpha_4(i) = \begin{bmatrix} 0.00156 \\ 0.000281 \\ 0.0019173 \\ 0.00055515 \end{bmatrix}$$

$P(o_{1:4} \mid A, B, \pi) = \sum_i \alpha_4(i) = 0.00431$

**A: $a_{ij} = P(X_{t+1} = j \mid X_t = i)$**

| $X_t \mid X_{t+1}$ | A | B | H | S |
|---|---|---|---|---|
| **A** | 0.6 | 0.1 | 0.1 | 0.2 |
| **B** | 0.0 | 0.3 | 0.2 | 0.5 |
| **H** | 0.8 | 0.1 | 0.0 | 0.1 |
| **S** | 0.2 | 0.0 | 0.1 | 0.7 |

**B: $b_{ik} = P(O_t = k \mid X_t = i)$**

| $X_t \mid O_t$ | p | e | b | l |
|---|---|---|---|---|
| **A** | 0.6 | 0.2 | 0.1 | 0.1 |
| **B** | 0.1 | 0.4 | 0.1 | 0.4 |
| **H** | 0.0 | 0.0 | 0.7 | 0.3 |
| **S** | 0.0 | 0.0 | 0.1 | 0.9 |

**Step 1: Initialize $\delta_1(i)$**

$O_1 = \underline{b}$

$$\delta_1(i) = \begin{array}{c} \pi = P(X_1 = i) \\ \begin{bmatrix} 0.5 \\ 0 \\ 0.0 \\ 0.5 \end{bmatrix} \end{array} \odot \begin{array}{c} P(b \mid S_i) \\ \begin{bmatrix} 0.1 \\ 0.1 \\ 0.7 \\ 0.1 \end{bmatrix} \end{array} = \begin{array}{c} \delta_1(i) \\ \begin{bmatrix} 0.05 \\ 0 \\ 0 \\ 0.05 \end{bmatrix} \end{array}$$

**$\pi = P(X_1 = i)$ :**

| A | B | H | S |
|---|---|---|---|
| 0.5 | 0.0 | 0.0 | 0.5 |

**Observations:**

$O_{1:4} = \{b,p,l,e\}$

**Find:**
**Most likely hidden state sequence:**

$X^*_{1:4}$

**$O_2$ = p**

**$\delta_2(i)$ =**

Get value from

A-matrix B-matrix

Most likely current state

| | | | max probability | argmax state |
|---|---|---|---|---|
| $X_A$: | max ( 0.05 x 0.6 x 0.6 , 0 x 0 x 0.6 , 0 x 0.8 x 0.6 , 0.05 x 0.2 x 0.6 ) | = | 0.018 | ( A ) |
| $X_B$: | max ( 0.05 x 0.1 x 0.1 , 0 x 0.3 x 0.1 , 0 x 0.1 x 0.1 , 0.05 x 0 x 0.1 ) | | 0.0005 | ( A ) |
| $X_H$: | max ( 0.05 x 0.1 x 0 , 0 x 0.2 x 0 , 0 x 0 x 0 , 0.05 x 0.1 x 0 ) | | 0 | ( ∅ ) |
| $X_S$: | max ( 0.05 x 0.2 x 0 , 0 x 0.5 x 0 , 0 x 0.1 x 0 , 0.05 x 0.7 x 0.0 ) | | 0 | ( ∅ ) |

Labels above row $X_A$: $\delta_1(A)$ P(A|A) P(p|A) , $\delta_1(B)$ P(A|B) P(p|A) , $\delta_1(H)$ P(A|H) P(p|A) , $\delta_1(S)$ P(A|S) P(p|A)

Labels above row $X_S$: $\delta_1(A)$ P(S|A) P(p|S) , $\delta_1(B)$ P(S|B) P(p|S) , $\delta_1(H)$ P(S|H) P(p|S) , $\delta_1(S)$ P(S|S) P(p|S)

Most likely Previous state

**$O_3$ = l**

**$\delta_3(i)$ =**

| | | | max. probability | argmax state |
|---|---|---|---|---|
| $X_A$: | max ( 0.018 x 0.6 x 0.1 , .0005 x 0 x 0.1 , 0 x 0.8 x 0.1 , 0 x 0.2 x 0.1 ) | = | 0.00108 | ( A ) |
| $X_B$: | max ( .018 x 0.1 x 0.4 , 0.0005 x 0.3 x 0.4 , 0 x 0.1 x 0.4 , 0 x 0 x 0.4 ) | | 0.00072 | ( A ) |
| $X_H$: | max ( .018 x 0.1 x 0.3 , .0005 x 0.2 x 0.3 , 0 x 0 x 0.3 , 0 x 0.1 x 0.3 ) | | 0.00054 | ( A ) |
| $X_S$: | max ( .018 x 0.2 x 0.9 , .0005 x 0.5 x 0.4 , 0 x 0.1 x 0.9 , 0 x 0.7 x 0.9 ) | | 0.00324 | ( A ) |

**$O_4$ = e**

**$\delta_4(i)$ =**

| | | = | max. probability | argmax state |
|---|---|---|---|---|
| $X_A$: | max ( 0.00108 x 0.6 x 0.2 , 0.00072 x 0 x 0.2 , 0.00054 x 0.8 x 0.2 , 0.00324 x 0.2 x 0.2 ) | | $1.296 \times 10^{-4}$ | ( A , S ) |
| $X_B$: | max ( 0.00108 x 0.1 x 0.4 , 0.00072 x 0.3 x 0.4 , 0.00054 x 0.1 x 0.4 , 0.00324 x 0 x 0.4 ) | | $8.64 \times 10^{-5}$ | ( B ) |
| $X_H$: | max ( 0.00108 x 0.1 x 0 , 0.00072 x 0.2 x 0 , 0.00054 x 0 x 0.0 , 0.00324 x 0.1 x 0 ) | | 0 | ( ∅ ) |
| $X_S$: | max ( 0.00108 x 0.2 x 0 , 0.00072 x 0.5 x 0 , 0.00054 x 0.1 x 0 , 0.00324 x 0.7 x 0 ) | | 0 | ( ∅ ) |

States and deltas over time:

*most likely previous state

| $\delta_1(i)$ | state* | $\delta_2(i)$ | state* | $\delta_3(i)$ | state* | $\delta_4(i)$ | state* |
|---|---|---|---|---|---|---|---|
| 0.05 | – | 0.018 | A | 0.00108 | A | $1.296 \times 10^{-4}$ | A , S |
| 0 | – | 0.0005 | A | 0.00072 | A | $8.64 \times 10^{-5}$ | B |
| 0 | – | 0 | – | 0.00054 | A | 0 | – |
| 0.05 | – | 0 | – | 0.00324 | A | 0 | – |

Backtracking gives two answers: **$X^*_{1:4}$ = {** A , A , A , A **}** **and** **$X^*_{1:4}$ = {** A , A , S , A **}**