Machine, Data and Learning

Assignment 5

Part 1

2018111032
Vivek Pamnani

x = 0.67
y = 2

P(Observation = Red | State = Red)	0.85
P(Observation = Green | State = Green)	0.9

Table : 2

S (set of states):

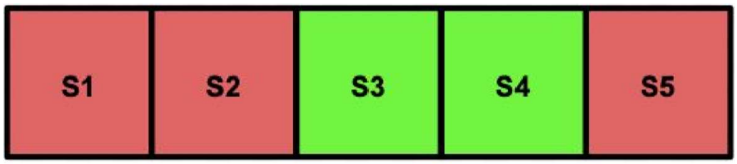

Fig. 1 : State Space

A (set of actions): [Left, Right]

P (table of transition probabilities):

Left	S1	S2	S3	S4	S5
S1	0.67	0.33	0	0	0
S2	0.67	0	0.33	0	0
S3	0	0.67	0	0.33	0
S4	0	0	0.67	0	0.33
S5	0	0	0	0.67	0.33

Right	S1	S2	S3	S4	S5
S1	0.33	0.67	0	0	0
S2	0.33	0	0.67	0	0
S3	0	0.33	0	0.67	0
S4	0	0	0.33	0	0.67
S5	0	0	0	0.33	0.67

Calculation of new belief state:

$$b'(s') = \frac{O(s', a, o) \times \sum T(s, a, s') \cdot b(s)}{Pr(o \mid a, b)}$$

Where,
s: old state,
s': new state,
a: action,
o: observation,
b: old belief state,
b': new belief state,
b(s): probability of being in state 's', given 'b',
b'(s'): probability of being in state 's'', given 'b''

Action 1: Agent took the action Right and observed Red

$$a = Right$$
$$b = [0.333, 0.333, 0, 0, 0.333]$$
$$o = Red$$

b' = new belief state (let)

$$\mathrm{b}'(S1) = O(o, a, S1) \sum_{S} T(S, a, S1) \cdot b(S)$$
$$= 0.85 \cdot [0.33 \cdot 0.333 + 0.33 \cdot 0.333 + 0 + 0 + 0\ ]$$
$$= 0.1868$$
$$\mathrm{b}'(S2) = O(o, a, S2) \sum_{S} T(S, a, S2) \cdot b(S)$$
$$= 0.85 \cdot [0.67 \cdot 0.333 + 0 + 0 + 0 + 0]$$
$$= 0.1896$$
$$\mathrm{b}'(S3) = O(o, a, S3) \sum_{S} T(S, a, S3) \cdot b(S)$$
$$= 0.1 \cdot [0 + 0.67 \cdot 0.333 + 0 + 0 + 0\ ]$$
$$= 0.0223$$
$$\mathrm{b}'(S4) = O(o, a, S4) \sum_{S} T(S, a, S4) \cdot b(S)$$
$$= 0.1 \cdot [0 + 0 + 0 + 0 + 0.33 \cdot 0.333]$$
$$= 0.0110$$
$$\mathrm{b}'(S5) = O(o, a, S5) \sum_{S} T(S, a, S5) \cdot b(S)$$
$$= 0.85 \cdot [0 + 0 + 0 + 0 + 0.67 \cdot 0.333\ ]$$
$$= 0.1896$$

Normalizing denominator $= \sum_{S} b'(S) = 0.5993$

$$\boldsymbol{b' = [0.3117, 0.31637, 0.03721, 0.01835, 0.31637]}$$

Action 2: Agent took the action Left and observed Green

$$a = Left$$
$$b = [0.3117, 0.31637, 0.03721, 0.01835, 0.31637]$$
$$o = Green$$

b' = new belief state (let)

$$\mathrm{b}'(S1) = O(o,a,S1)\sum_{S} T(S,a,S1)\cdot b(S)$$
$$= 0.15\cdot[0.67\cdot 0.3117 + 0.67\cdot 0.31637 + 0 + 0 + 0\,]$$
$$= 0.06312$$
$$\mathrm{b}'(S2) = O(o,a,S2)\sum_{S} T(S,a,S2)\cdot b(S)$$
$$= 0.15\cdot[0.33\cdot 0.3117 + 0 + 0.67\cdot 0.03721 + 0 + 0]$$
$$= 0.01917$$
$$\mathrm{b}'(S3) = O(o,a,S3)\sum_{S} T(S,a,S3)\cdot b(S)$$
$$= 0.9\cdot[0 + 0.33\cdot 0.31637 + 0 + 0.67\cdot 0.01835 + 0\,]$$
$$= 0.10503$$
$$\mathrm{b}'(S4) = O(o,a,S4)\sum_{S} T(S,a,S4)\cdot b(S)$$
$$= 0.9\cdot[0 + 0 + 0.33\cdot 0.03721 + 0 + 0.67\cdot 0.31637]$$
$$= 0.20183$$
$$\mathrm{b}'(S5) = O(o,a,S5)\sum_{S} T(S,a,S5)\cdot b(S)$$
$$= 0.15\cdot[0 + 0 + 0 + 0.33\cdot 0.01835 + 0.33\cdot 0.31637]$$
$$= 0.01657$$

Normalizing denominator $= \sum_{S} b'(S) = 0.40572$

$$\boldsymbol{b' = [0.15558, 0.04725, 0.25887, 0.49746, 0.04084]}$$

Action 3: Agent took the action Left and observed Green

$$a = Left$$
$$b' = [0.15558, 0.04725, 0.25887, 0.49746, 0.04084]$$
$$o = Green$$

b' = new belief state (let)

$$\mathrm{b}'(S1) = O(o, a, S1) \sum_{S} T(S, a, S1) \cdot b(S)$$
$$= 0.15 \cdot [0.67 \cdot 0.15558 + 0.67 \cdot 0.04725 + 0 + 0 + 0\ ]$$
$$= 0.02038$$
$$\mathrm{b}'(S2) = O(o, a, S2) \sum_{S} T(S, a, S2) \cdot b(S)$$
$$= 0.15 \cdot [0.33 \cdot 0.15558 + 0 + 0.67 \cdot 0.25887 + 0 + 0]$$
$$= 0.03372$$
$$\mathrm{b}'(S3) = O(o, a, S3) \sum_{S} T(S, a, S3) \cdot b(S)$$
$$= 0.9 \cdot [0 + 0.33 \cdot 0.04725 + 0 + 0.67 \cdot 0.49746 + 0\ ]$$
$$= 0.314$$
$$\mathrm{b}'(S4) = O(o, a, S4) \sum_{S} T(S, a, S4) \cdot b(S)$$
$$= 0.9 \cdot [0 + 0 + 0.33 \cdot 0.25887 + 0 + 0.67 \cdot 0.04084]$$
$$= 0.10151$$
$$\mathrm{b}'(S5) = O(o, a, S5) \sum_{S} T(S, a, S5) \cdot b(S)$$
$$= 0.15 \cdot [0 + 0 + 0 + 0.33 \cdot 0.49746 + 0.33 \cdot 0.04084]$$
$$= 0.02665$$

Normalizing denominator $= \sum_{S} b'(S) = 0.49626$

$$b' = [0.04107, 0.06795, 0.63273, 0.20455, 0.0537]$$

Final Belief State $= [\mathbf{0.04107}, \mathbf{0.06795}, \mathbf{0.63273}, \mathbf{0.20455}, \mathbf{0.0537}]$