**BITS PILANI, DUBAI CAMPUS**
**DUBAI INTERNATIONAL ACADEMIC CITY, DUBAI**

**FIRST SEMESTER 2023 – 2024**
**COURSE:** BITS F464 (Machine Learning)
**COMPONENT:** Practice Tutorial 5 **DATE:** 22 December 2023

**Q1.**

Calculate the Q-learning table for following listed state, rewards, and actions sequences for two states and two action system.

i. Current state: $S_2$, Reward (R) = +20, Action ($a_1$): $S_2 \rightarrow S_1$
ii. Current state: $S_1$, Reward (R) = -10, Action ($a_2$): $S_1 \rightarrow S_2$

The learning rate ($\alpha$) is 0.5 and a discount factor ($\gamma$) is 0.5 for each step. The initial Q-table is as follows,

| **Q-values** | **S1** | **S2** |
|---|---|---|
| **a1** | -5 | 0 |
| **a2** | -5 | 0 |

Q.B4 ⇒

$$Q(a,s) \leftarrow Q(a,s) + \alpha R(s) + \gamma \max_{a'} [Q(a',s')] - Q(a,s)$$

(New Q-value)

(i) $Q(a_1, S_2) \leftarrow 0 + 0.5(+20 + 0.5 \max_{a'(s'=S_1)} [-5, -5] - 0)$

$Q(a_1, S_2) = 8.75$

| Q | $S_1$ | $S_2$ |
|---|---|---|
| $a_1$ | -5 | 8.75 |
| $a_2$ | -5 | 0 |

(ii) $Q(a_2, S_1) \leftarrow -5 + 0.5(-10 + 0.5 \max_{a'(s'=S_2)} [8.75, 0] - (-5))$

$Q(a_2, S_1) = -5.3125$

| Q | $S_1$ | $S_2$ |
|---|---|---|
| $a_1$ | -5 | 8.75 |
| $a_2$ | -5.3125 | 0 |

**Q2.**

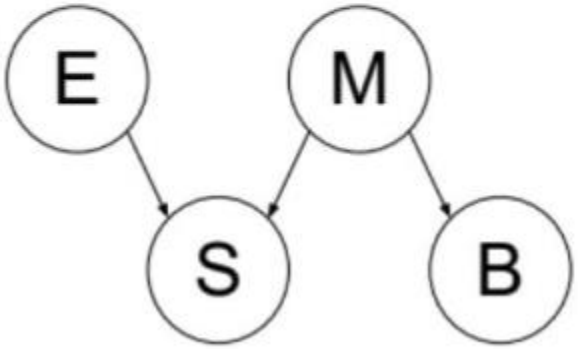
Consider the following Bayesian network and corresponding conditional probability tables (as shown in Figure 3). Calculate the below probabilities considering the given Bayesian network,

a. *P (-e, -s,-m,-b)*

b. *P (+m|+b)*

| $P(E)$ | |
|---|---|
| $+e$ | 0.4 |
| $-e$ | 0.6 |

| $P(S\|E,M)$ | | | |
|---|---|---|---|
| $+e$ | $+m$ | $+s$ | 1.0 |
| $+e$ | $+m$ | $-s$ | 0.0 |
| $+e$ | $-m$ | $+s$ | 0.8 |
| $+e$ | $-m$ | $-s$ | 0.2 |
| $-e$ | $+m$ | $+s$ | 0.3 |
| $-e$ | $+m$ | $-s$ | 0.7 |
| $-e$ | $-m$ | $+s$ | 0.1 |
| $-e$ | $-m$ | $-s$ | 0.9 |




| $P(M)$ | |
|---|---|
| $+m$ | 0.1 |
| $-m$ | 0.9 |

| $P(B\|M)$ | | |
|---|---|---|
| $+m$ | $+b$ | 1.0 |
| $+m$ | $-b$ | 0.0 |
| $-m$ | $+b$ | 0.1 |
| $-m$ | $-b$ | 0.9 |

(a) $P(-e,-s,-m,-b) = P(-e)\,P(-m)\,P(-s|-e,-m)\,P(-b|-m)$

$= (0.6)(0.9)(0.9)(0.9) = 0.4374$

(According to chain rule of conditional probability)

(b) $P(+m|+b) = \frac{P(+b|+m)\,P(+m)}{P(+b)} = \frac{(0.1)(0.1)}{0.19}$

$= 0.5263$

**Q3.**

Suppose Mr. X purchased share A on first day, on next day there is 60% chance of purchasing share A, a 30 % chance of purchasing share B, and a 10% chance of purchasing share C. On a day when Mr. X purchased share B, the next day's he will purchase share A is 40%, a 30% chance of purchasing share B, and a 30% chance of purchasing share C. Lastly, on any specific day when Mr. X purchased share C, there is a 20% chance of purchasing share A, a 50% chance of purchasing share B, and a 30% chance of purchasing share C on a following day. So, by considering Mr. X purchased share A today, calculate the probability that Mr. X will purchase share A two days from now.

Q.3 ⇒ Transition Matrix according to given probabilities is,

$$P = \begin{bmatrix} 0.6 & 0.4 & 0.2 \\ 0.3 & 0.3 & 0.5 \\ 0.1 & 0.3 & 0.3 \end{bmatrix}$$

State vector, $X = \begin{bmatrix} x_A \\ x_B \\ x_C \end{bmatrix}$

For given example ⇒ $x^{(0)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$$x^{(1)} = P x^{(0)} = \begin{bmatrix} 0.6 & 0.4 & 0.2 \\ 0.3 & 0.3 & 0.5 \\ 0.3 & 0.3 & 0.5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.3 \\ 0.1 \end{bmatrix}$$

$$x^{(2)} = P x^{(1)} = \begin{bmatrix} 0.6 & 0.4 & 0.2 \\ 0.3 & 0.3 & 0.5 \\ 0.3 & 0.3 & 0.5 \end{bmatrix} \begin{bmatrix} 0.6 \\ 0.3 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.32 \\ 0.18 \end{bmatrix}$$

∴ Probability of purchasing share A after 2 days is 0.5 or 50%