Question 3 —

To show that Q_n is an exponential recency weighted average without initial bias.

Given, say $\beta_n \triangleq \alpha / \bar{o}_n$

where $\bar{o}_n \triangleq \bar{o}_{n-1} + \alpha(1 - \bar{o}_{n-1})$ for $n \geq 0$

where $\bar{o}_0 \triangleq 0$.

Method 1.

Solving for $\bar{o}_1$

$\bar{o}_1 = \bar{o}_0 + \alpha(1 - \bar{o}_0)$

$= 0 + \alpha$

$\boxed{\bar{o}_1 = \alpha}$

Solving for $\bar{o}_2$

$\bar{o}_2 = \bar{o}_1 + \alpha(1 - \bar{o}_1)$

$= \alpha + \alpha(1 - \alpha)$

$= 2\alpha - \alpha^2$

$\boxed{\bar{o}_2 = \alpha(2 - \alpha)}$

We know $Q_{n+1} = \alpha R_n + (1 - \alpha) Q_n$

Substituting α with $\beta_n \triangleq \alpha / \bar{o}_n$

Thus Q_1 doesn't exist (divide by 0 error).

FRIDAY

22

MARCH

2019

Notes

Appointment

SATURDAY

23

MARCH

2019

Putting $n=1$

$$Q_2 = \frac{\alpha}{\bar{O}_1} R_1 + \left(1 - \frac{\alpha}{O_1}\right) Q_1$$

$$= \frac{\alpha}{\alpha} R_1 + \left(1 - \frac{\alpha}{\alpha}\right) Q_1$$

$$= R_1 + 0$$

$$\therefore \boxed{Q_2 = R_1} \quad - 1$$

$$Q_3 = \frac{\alpha}{\bar{O}_2} R_2 + \left(1 - \frac{\alpha}{\bar{O}_2}\right) Q_2$$

$$= \frac{\alpha}{\alpha(2-\alpha)} R_2 + \left(1 - \frac{\alpha}{\alpha(2-\alpha)}\right) R_1$$

$$= \frac{1}{2-\alpha} \cdot R_2 + \left(1 - \frac{1}{2-\alpha}\right) R_1$$

$$= \frac{R_2}{2-\alpha} + \left(\frac{2-\alpha+1}{2-\alpha}\right) R_1$$

$$= \frac{1}{2-\alpha}\left[R_2 + (3-\alpha) R_1\right] \quad - 2$$

Thus further calculations will show that Q_n only depends on previous 2 rewards and value of constant step size.
Due to no term of Q_1 (shown again next), Q_n is an exponential recency weighted average without initial bias.

Method 2

SUNDAY
24
MARCH
2019

We know, $Q_{n+1} = Q_n + \alpha [R_n - Q_n]$

$= \alpha R_n + (1-\alpha) Q_n$ — (1)

Substituting α with $\beta_n \triangleq \frac{\alpha}{\bar{O}_n}$ in equation (1)

$$Q_{n+1} = \frac{\alpha}{\bar{O}_n} R_n + \left(1 - \frac{\alpha}{\bar{O}_n}\right) Q_n$$

$$= \frac{\alpha}{\bar{O}_n} R_n + \left(\frac{\bar{O}_n - \alpha}{\bar{O}_n}\right) Q_n$$

$$= \frac{\alpha}{\bar{O}_n} R_n + \left[\frac{\bar{O}_{n-1} + \alpha(1 - \bar{O}_{n-1}) - \alpha}{\bar{O}_n}\right] Q_n$$

$$= \alpha \cdot \frac{R_n}{\bar{O}_n} + \frac{\bar{O}_{n-1}}{\bar{O}_n} (1-\alpha) (Q_n) \quad \text{— (2)}$$

Substituting for Q_n in equation (2)

$$Q_{n+1} = \alpha \cdot \frac{R_n}{\bar{O}_n} + \frac{\bar{O}_{n-1}}{\bar{O}_n} (1-\alpha) \left[\frac{\alpha}{\bar{O}_{n-1}} R_{n-1} + \left(1 - \frac{\alpha}{\bar{O}_{n-1}}\right) Q_{n-1}\right]$$

$$= \alpha \frac{R_n}{\bar{O}_n} + \frac{\alpha(1-\alpha)}{\bar{O}_n} R_{n-1} + \frac{\bar{O}_{n-1}}{\bar{O}_n} (1-\alpha)(1-\alpha) \frac{\bar{O}_{n-2}}{\bar{O}_{n-1}} \cdot Q_{n-1}$$

$$= \alpha \cdot \frac{R_n}{\bar{O}_n} + \frac{\alpha(1-\alpha)}{\bar{O}_n} R_{n-1} + (1-\alpha)^2 \cdot \frac{\bar{O}_{n-2}}{\bar{O}_n} Q_{n-1}$$

$$Q_{n+1} = (1-\alpha)^n \frac{\bar{O}_0}{\bar{O}_2} Q_1 + \sum_{i=1}^{n} \alpha (1-\alpha)^{n-i} \frac{R_i}{\bar{O}_i} \quad \text{— (3)}$$

Notes	Appointment

MONDAY

25

MARCH

2019

Thus from equation (3) where,

$$Q_{n+1} = (1-\alpha)^n \frac{\bar{O}_0}{\bar{O}_2} \cdot Q_1 + \sum_{i=1}^{n} \alpha(1-\alpha)^{n-i} \frac{R_i}{\bar{O}_i}$$

Since $\bar{O}_0 \overset{\Delta}{=} 0$

$$\therefore Q_{n+1} = \sum_{i=1}^{n} \alpha(1-\alpha)^{n-i} \frac{R_i}{\bar{O}_i} \quad -(4)$$

From eq. (4), we can infer that Q_{n+1} doesn't depend on initial bias due to absence of term Q_1 and is an exponential recency average weigented due to term $\alpha(1-\alpha)^{n-i}$ (as explainted in Sutton and Barto).

Thus we can clearly say that Q_n is free from any initial bias.