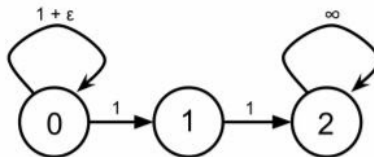


پاسخنامه تمرین ۷ بخش ۱

سوال ۱)

بله امکان دارد که policy تغییر کند. برای مثال اگر ۳ استیت $\{0, 1, 2\}$ را در نظر بگیریم در هر استیت اکشن های قابل انجام توسط یال های خارج شونده مشخص می شوند و استیت مقصد را مشخص می کنند. انتقال استیت ها قطعی است یعنی اگر اکشن انجام شود همواره به رسیدن به حالت بعدی منجر می شود. اگر داشته باشیم $\gamma = 1$ و $\xi \in (0, 1)$ یک ثابت کوچک است.

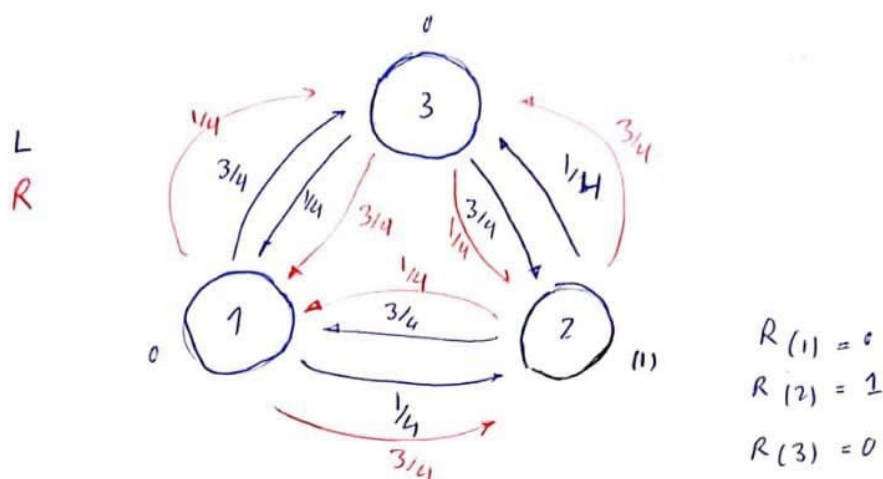


با انجام پیمایش ها جدول زیر را داریم.

| Iteration | $V(0)$ | $V(1)$ | $V(2)$ |
|-----------|--------------|----------|----------|
| 0 | 0 | 0 | 0 |
| 1 | $1 + \xi$ | 1 | ∞ |
| 2 | $2(1 + \xi)$ | ∞ | ∞ |
| 3 | ∞ | ∞ | ∞ |

بهترین policy ها نیز در جدول زیر قابل مشاهده است.

| Iteration | $\pi(0)$ | $\pi(1)$ | $\pi(2)$ |
|-----------|----------|----------|----------|
| 0 | Arb. | Arb. | Arb. |
| 1 | 0 | 2 | 2 |
| 2 | 0 | 2 | 2 |
| 3 | 1 | 2 | 2 |



$$V_{K+1}(s) = \max_a \sum_{s'} T(s, a, s') [R(s, a, s') + \gamma V_K(s')]$$

$$V_0(1) = 0 \quad V_0(2) = 0 \quad V_0(3) = 0$$

$$V_{1(1)} = \begin{cases} R: \frac{3}{4} \times 1 + \frac{1}{4} \times 0 \\ L: \frac{1}{4} \times 1 + \frac{3}{4} \times 0 \end{cases} \rightarrow V_{1(1)} = \frac{3}{4} \rightarrow V_{2(1)} = \begin{cases} R: \frac{3}{4} \times (1+0) + \frac{1}{4} \times [0 + \frac{1}{2} \times \frac{3}{4}] \\ L: \frac{1}{4} \times (1+0) + \frac{3}{4} \times [0 + \frac{1}{2} \times \frac{3}{4}] \end{cases}$$

$$\rightarrow V_{2(1)} = 0.84375 = \frac{27}{32}$$

$$V_{1(2)} = \begin{cases} R: \frac{3}{4} \times 0 + \frac{1}{4} \times 0 \\ L: \frac{1}{4} \times 0 + \frac{3}{4} \times 0 \end{cases} \rightarrow V_{1(2)} = 0 \rightarrow V_{2(2)} = \begin{cases} R: \frac{3}{4} \times (\frac{1}{2} \times \frac{3}{4}) + \frac{1}{4} \times (\frac{1}{2} \times \frac{3}{4}) \\ L: \frac{1}{4} \times (\frac{1}{2} \times \frac{3}{4}) + \frac{3}{4} \times (\frac{1}{2} \times \frac{3}{4}) \end{cases}$$

$$\rightarrow V_{2(2)} = \frac{3}{8}$$

$$V_{1(3)} = \begin{cases} R: \frac{1}{4} \times 1 + \frac{3}{4} \times 0 \\ L: \frac{3}{4} \times 1 + \frac{1}{4} \times 0 \end{cases} \rightarrow V_{1(3)} = \frac{3}{4} \rightarrow V_{2(3)} = \begin{cases} R: \frac{1}{4} \times (1+0) + \frac{3}{4} \times (0 + \frac{1}{2} \times \frac{3}{4}) \\ L: \frac{1}{4} \times (\frac{1}{2} \times \frac{3}{4}) + \frac{3}{4} \times (\frac{1}{2} \times \frac{3}{4}) \end{cases}$$

$$\rightarrow V_{2(3)} = 0.84375 = \frac{27}{32}$$

$$V_{1(1)} = \begin{cases} R: \frac{3}{4} \times (1 + \frac{1}{2} \times 0.25) + \frac{1}{4} \times (0 + \frac{1}{2} \times \frac{27}{32}) \\ L: \frac{1}{4} \times (1 + \frac{1}{2} \times 0.25) + \frac{3}{4} \times (0 + \frac{1}{2} \times \frac{27}{32}) \end{cases} \rightarrow V_{1(1)} = V_{1(3)} = 0.90625 = \frac{29}{32}$$

$$V_{2(2)} = \frac{3}{4} \times (0 + 0.5 \times 0.5) + \frac{1}{4} \times (0 + 0.5 \times 0.5) = 0.25$$