

1 Offline λ -Return Algorithm

(a)

$$\begin{aligned}\delta_0 &= (R_0 + \gamma V(S_1)) - V(S_0) \\ &= (-1 + 1 \times (-1)) - (-3) \\ &= 1 \\ \delta_1 &= (R_1 + \gamma V(S_2)) - V(S_1) \\ &= (-1 + 1 \times (0)) - (-1) \\ &= 0\end{aligned}$$

(b)

with $\lambda = 1$ we recover the MC form of the update equation:

$$V(S_t) \leftarrow V(S_t) + \alpha(G_t^\lambda - V(S_t))$$

then, for states S_t , $t = \{0, 1\}$, we have:

$$\begin{aligned}V(S_0) &\leftarrow -3 + 0.5 \times (-2 - (-3)) \\ &\leftarrow -2.5 \\ V(S_1) &\leftarrow -1 + 0.5 \times (-1 - (-1)) \\ &\leftarrow -1\end{aligned}$$

all other states remain unchanged.

2 Implementation

Note to self: don't forget to multiply (1-done) when calculating TD error