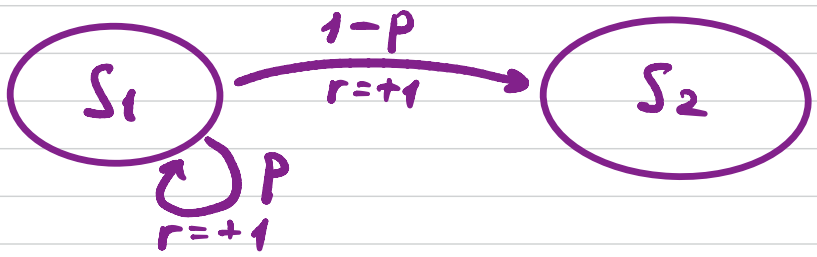
Q1. 5.5:



Have T = episode duration = 10 (steps); $\gamma = 1$, $G_T = 10$.

Find first-visit and every-visit estimates for S_1.

Solution: Looking for $V_\pi(S_1) = E_\pi[G_t | S_t = S_1]$.

As $G_{10} = 10$, have the following transitions:

$$\underbrace{S_1 \xrightarrow{+1} S_1 \xrightarrow{+1} \dots \xrightarrow{+1} S_1}_{9 \text{ times}} \xrightarrow{+1} S_2$$

According to MC approach, $V(S) = E_\pi[G_t | S_t = S] \approx \frac{G_{t_1} + G_{t_2} + \dots + G_{t_n}}{n}$, where n is # of episodes, G_{t_i} = returns for state S for episode # i.

Thus, for the first-visit, we have $V(S_1) \approx \frac{1}{1} = 1$.

For, every-visit, we have $V(S_1) \approx \frac{1 + \gamma \cdot 1 + \dots + \gamma \cdot 1}{1} = \frac{1 \cdot 9}{1} = 9$