Reinforcement Learning formula

Lee Meng Oon

April 15, 2021

Discounted return at time step t:

$$G_t = \sum_{k=0} \gamma^k R_{t+k+1}, \qquad \gamma \in [0, 1] \tag{1}$$

One-step dynamics:

$$p(s', r|s, a) = \mathbb{P}(S_{t+1} = s', R_{t+1} = r|S_t = s, A_t = a) \tag{2}$$