

The link between Value and Policy

$$\pi^*(s) = \underset{a}{\operatorname{argmax}} Q^*(s, a)$$

Finding an optimal value function leads to having an optimal policy.

the state - value function

$$V_{\pi}(s) = E_{\pi}[G_t | S_t = s]$$

calculates the value of a state.

anbar:

the action - value function

$$Q_{\pi}(s, a) = E_{\pi}[G_t | S_t = s, A_t = a]$$

calculates value of state-action pair

the bellman equation

more efficient state - value function

$$V_{\pi}(s) = E_{\pi}[R_{t+1} + \gamma \times V_{\pi}(S_{t+1}) | S_t = s]$$

$$V(S_t) = R_{t+1} + \text{gamma} \times V(S_{t+1})$$