

$$Q_{\theta}(s_{t+1}, a_{t+1})$$

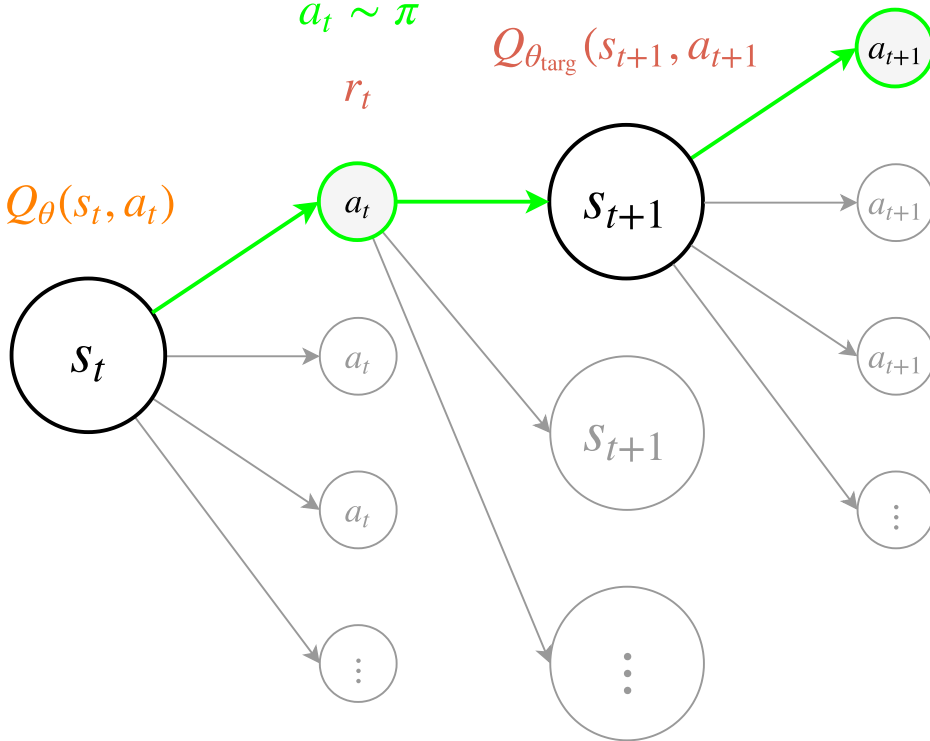
$$a_{t+1} \sim \pi$$

$$a_t \sim \pi$$

$$r_t$$

$$Q_{\theta_{\text{targ}}}(s_{t+1}, a_{t+1})$$

$$Q_{\theta}(s_t, a_t)$$



$$L(\theta) = (r_t + \gamma Q_{\theta_{\text{targ}}}(s_{t+1}, \arg \max_{a_{t+1}} Q_{\theta}(s_{t+1}, a_{t+1})) - Q_{\theta}(s_t, a_t))^2$$