

Q-LEARNING($\mathcal{S}, \mathcal{A}, s_0, \gamma, \alpha$)

1 **for** $s \in \mathcal{S}, a \in \mathcal{A}$:

2 $Q[s, a] = 0$

3 $s = s_0$ // Or draw an s randomly from \mathcal{S}

4 **while** True:

5 $a = \text{select_action}(s, Q)$

6 $r, s' = \text{execute}(a)$

7 $Q[s, a] = (1 - \alpha)Q[s, a] + \alpha(r + \gamma \max_{a'} Q[s', a'])$

8 $s = s'$