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

**for** $s \in \mathcal{S}, a \in \mathcal{A}$ :
    $Q[s, a] = 0$
$s = s_0$ // Or draw an $s$ randomly from $\mathcal{S}$
**while** True:
    $a$ = select_action$(s, Q)$
    $r, s' =$ execute$(a)$
    $Q[s, a] = (1 - \alpha)Q[s, a] + \alpha(r + \gamma \max_{a'} Q[s', a'])$
    $s = s'$