function Q-LEARNING-AGENT(e) **returns** an action

static: Q, a table of action values
N, a table of state-action frequencies
a, the last action taken
i, the previous state visited
r, the reward received in state i

j $\leftarrow$ STATE[e]
if i is non-null **then**
$N[a,i] \leftarrow N[a,i] + 1$
$Q[a,i] \leftarrow Q[a,i] + \alpha(r + \max_{a'} Q[a',j] - Q[a,i])$
if TERMINAL?[e] **then**
$i \leftarrow$ null
else
$i \leftarrow j$
$r \leftarrow$ REWARD[e]
$a \leftarrow \arg\max_{a'} f(Q[a',j], N[a',j])$
return a