RL - Assignment 2:

1. $p(S_1 | S_0, a_2)$:

Probability of transitioning from current state to state S_1 by taking action a_2.

2. Policy: A policy is a mapping from states to probabilities of selecting each possible action.

If the agent is following policy π at time t, then $\pi(a|s)$ is the probability that $A_t = a$ if $S_t = s$

Optimal policy: Best possible policy among all policies.

For any MDP, there exists a policy π^* such that $V_{\pi^*}(s) \geq V_{\pi}(s)$

$$q_{\pi^*}(s,a) \geq q_{\pi}(s,a)$$

3. $V_{\pi}(s) > V_{\pi^*}(s)$ - False.

4.