NAME: ....
CSCI S-89c Deep Reinforcement Learning
Part I of Assignment 3

Please consider a Markov Decision Process with $\mathcal{S} = \{s^L, s^R\}$.

Given a particular state $s \in \mathcal{S}$, the agent is allowed to either stay there or try switching to another state. The change of the state, given the agent's intention is to switch, however, happens in 70% of cases only - the states are "sticky." Let's denote an intention to stay by 0 and an attempt to change the state by 1, i.e. $\mathcal{A}(s^L) = \{0, 1\}$ and $\mathcal{A}(s^R) = \{0, 1\}$. Further, assume that the agent receives reward c_1 on the entrance to s^L and reward c_2 on the entrance to s^R.

(a) Please write down the transition probabilities $p(s', r|s, a)$. You need to explicitly specify all possible cases.

(b) Assuming policy $\pi(a|s)$ is to always take action 1 (i.e. try to switch states), find $v_\pi(s^L)$ and $v_\pi(s^R)$.
    Hint: Write down the Bellman equation for $v_\pi(s)$ and solve it.

SOLUTION: