

CS 3600: Markov Decision Process problem

Given the following Markov Decision Problem, use Value Iteration to find the optimal policy.

There are four states, s_1 , s_2 , s_3 , and s_4 , arranged in a grid. State s_4 is a sink (absorbing) state. The immediate rewards are given above. The agent can move UP, DOWN, LEFT, RIGHT and the transition model is such that there is an 80% chance of a correct move, and a 10% chance of moving to either side in error (e.g., if performing UP, there is a 10% chance of performing LEFT instead and a 10% chance of performing RIGHT instead).

Let the initial utility values for states are shown below.

- $U_0(s_1) = 0.1$
- $U_0(s_2) = 0.1$
- $U_0(s_3) = 0.1$
- $U_0(s_4) = 1.0$
- $\gamma = 0.5$

s_1 $r = -0.04$	s_4 $r = 1$
s_2 $r = -0.04$	s_3 $r = -0.04$

