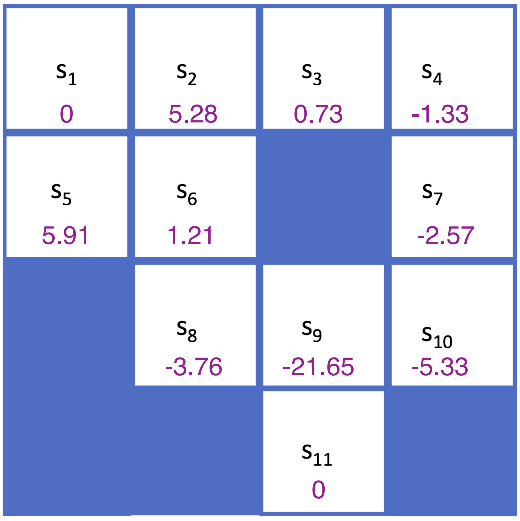
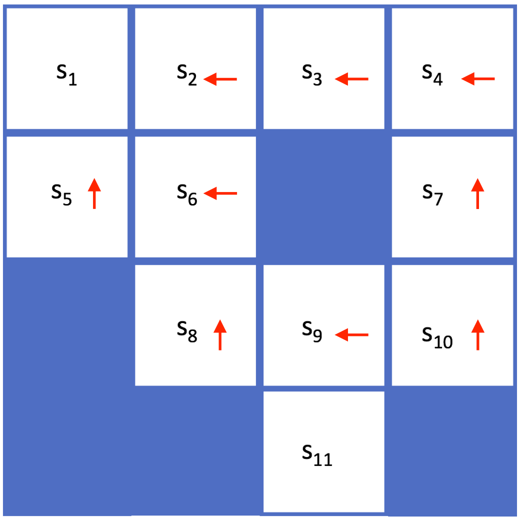
Name: Adrian Löwenstein
CID: 01594572
reward state: s1
$p =$ 0.5
$\gamma =$ 0.65



s_1 0	s_2 5.28	s_3 0.73	s_4 -1.33
s_5 5.91	s_6 1.21	■	s_7 -2.57
■	s_8 -3.76	s_9 -21.65	s_{10} -5.33
■	■	s_{11} 0	■

Figure 1: Optimal value function. Values for each state rounded to 2 decimal places.

.



s_1	s_2 ←	s_3 ←	s_4 ←
s_5 ↑	s_6 ←	■	s_7 ↑
■	s_8 ↑	s_9 ←	s_{10} ↑
■	■	s_{11}	■

Figure 2: Optimal policy. Arrows indicate optimal action direction for each state (deterministic policy), multiple arrows from one state indicate equiprobable choice between indicated directions (stochastic policy).

.