RBE595 - Week 4 Assignment

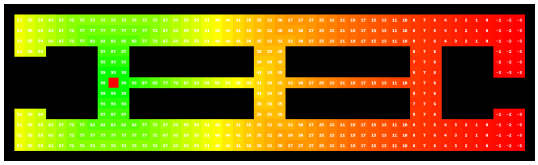
Keith Chester

Due date: February 6, 2023

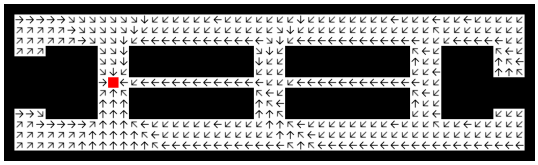
In this assignment, we prepared a simple gridworld with a robot that followed either a stochastic or a deterministic policy. In the deterministic polciy, we chose the highest possible score; if multiple tied this score, then an equal probability of selecting any of those higher scores. For the stochastic policy, we had an 80% chance to select the highest value and a 20% chance to select one of the adjacent directions within a 45 degree turn. If there were ties, then some combination of these tiles were selected per the probabilities dictated within the rules.

Here is a table that demonstrates the number of iterations and time it took to process each method of policy iteration, value iteration, and generalized policy iteration for both the deterministic and stochastic agents:

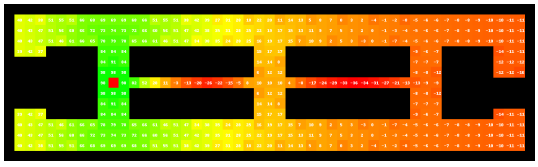
Method		Iterations	Seconds
Policy Iteration			
	Deterministic	31	5.21
	Stochastic	29	29.56
Value Iteration			
	Deterministic	11	0.36
	Stochastic	10	0.45
Generalized Policy Iteration			
	Deterministic	20	0.65
	Stochastic	53	2.00



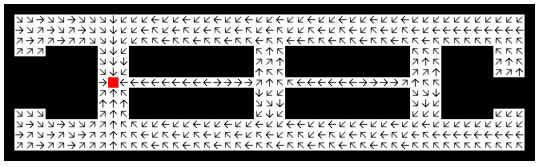
(a) Deterministic Policy Iteration Value Map



(b) Deterministic Policy Iteration Policy Map

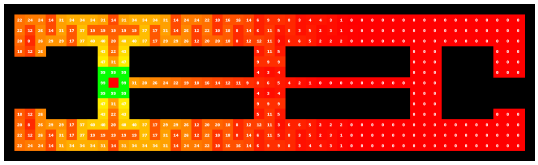


(c) Stochastic Policy Iteration Value Map

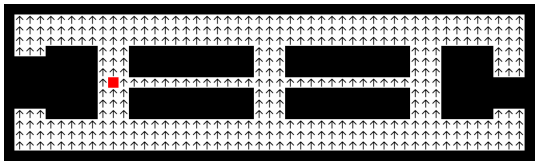


(d) Stochastic Policy Iteration Value Map

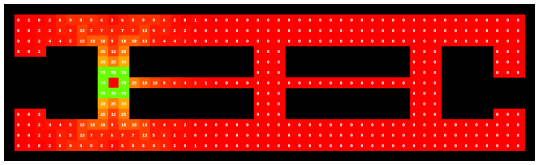
Figure 1: Policy and value maps generated with Policy Iteration



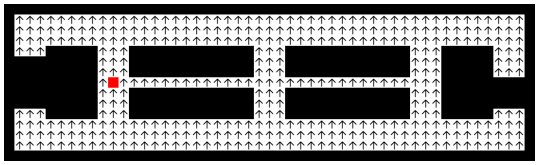
(a) Deterministic Value Iteration Value Map



(b) Deterministic Value Iteration Policy Map

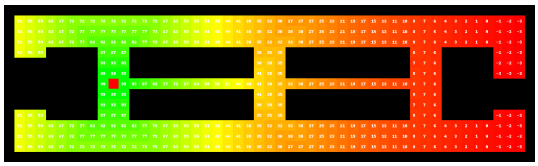


(c) Stochastic Value Iteration Value Map

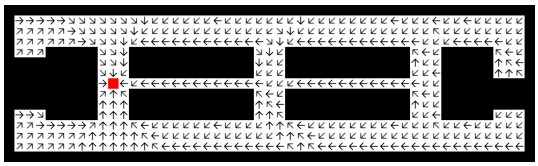


(d) Stochastic Value Iteration Value Map

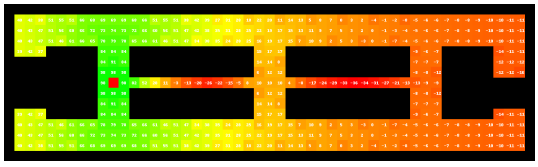
Figure 2: Policy and value maps generated with Value Iteration



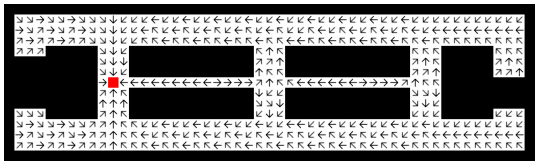
(a) Deterministic Generalized Policy Iteration Value Map



(b) Deterministic Generalized Policy Iteration Policy Map



(c) Stochastic Generalized Policy Iteration Value Map



(d) Stochastic Generalized Policy Iteration Value Map

Figure 3: Policy and value maps generated with Generalized Policy Iteration