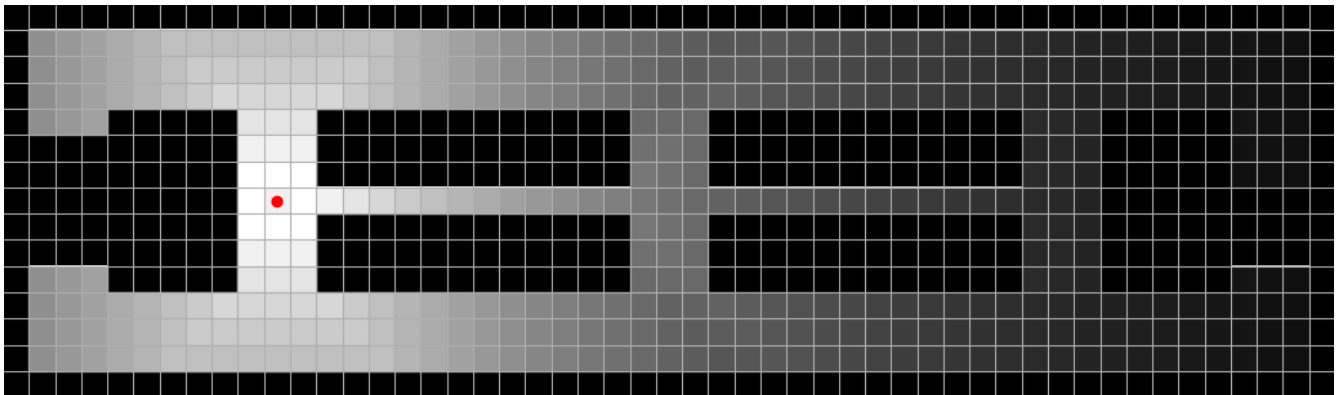
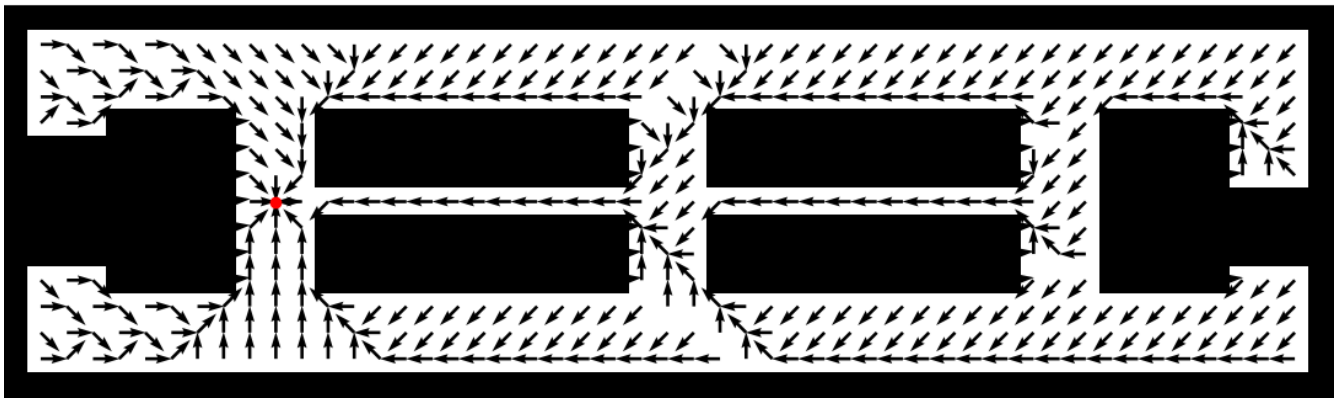


Week 4 Dynamic Programming Exercise

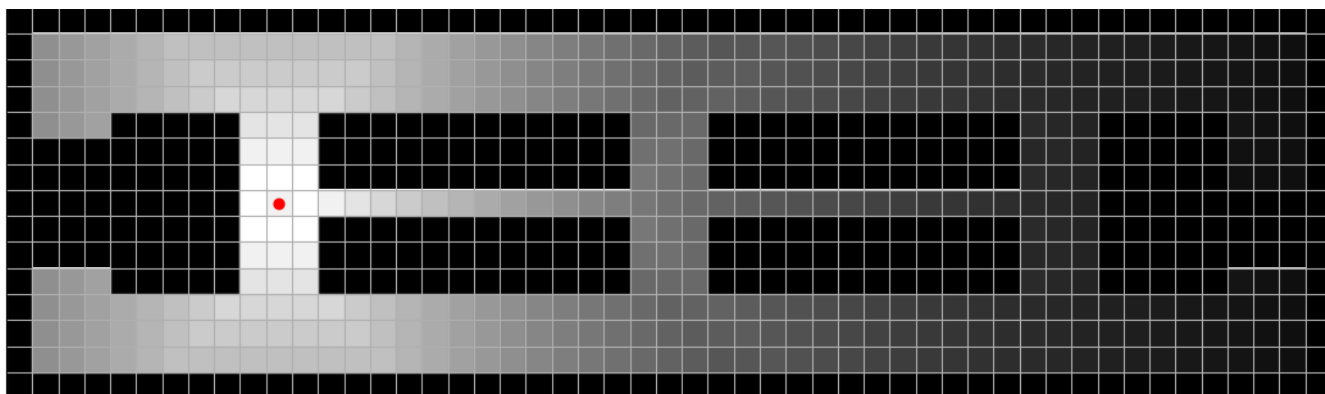
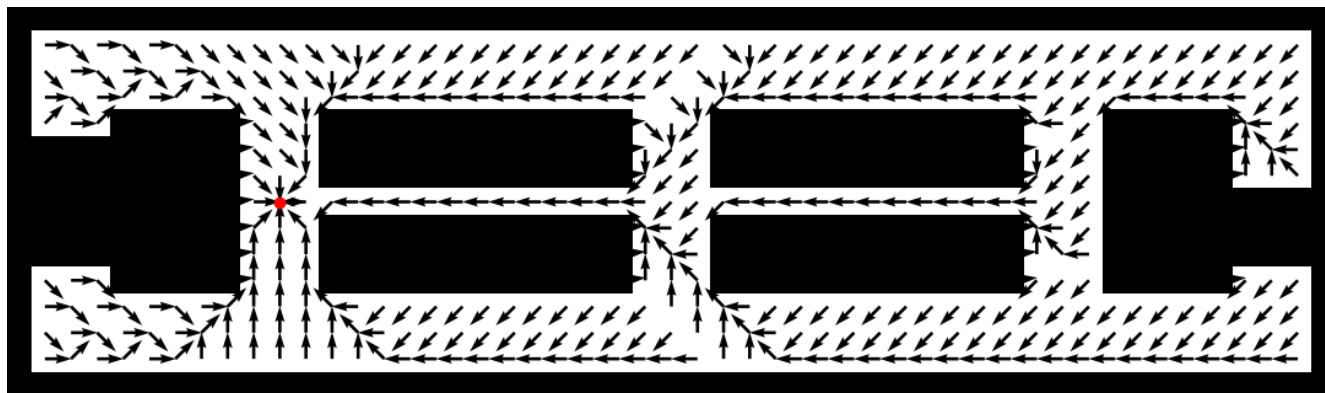
Results

Deterministic

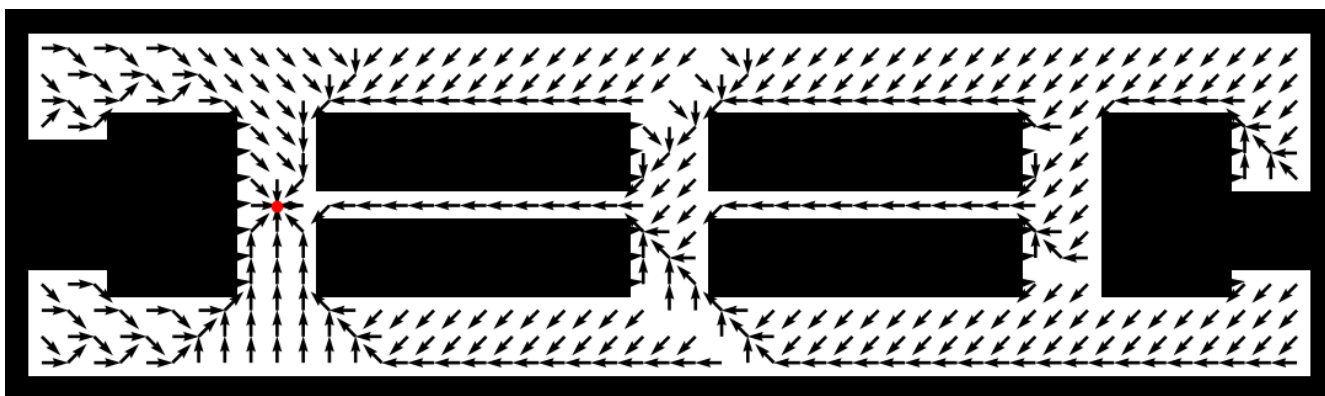
Policy Iteration

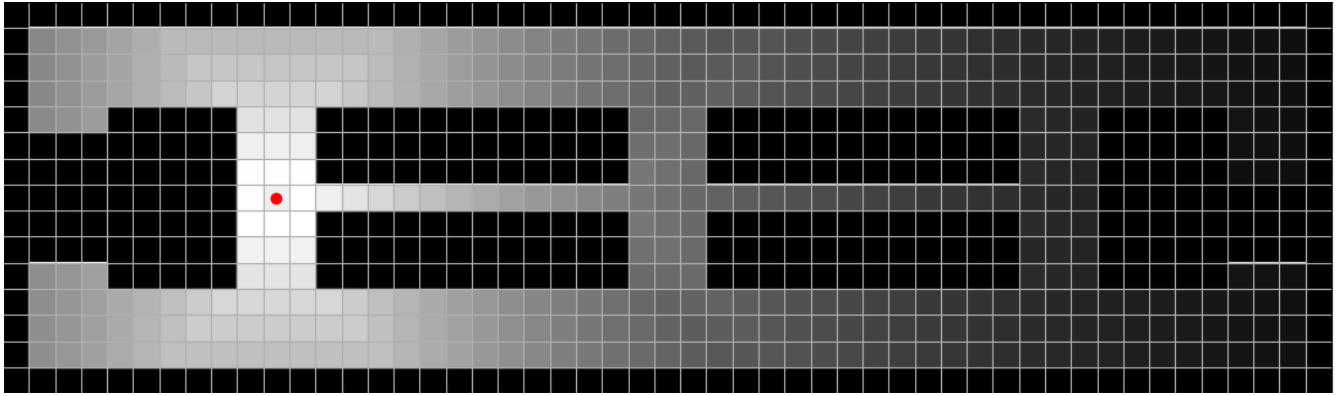


Value Iteration



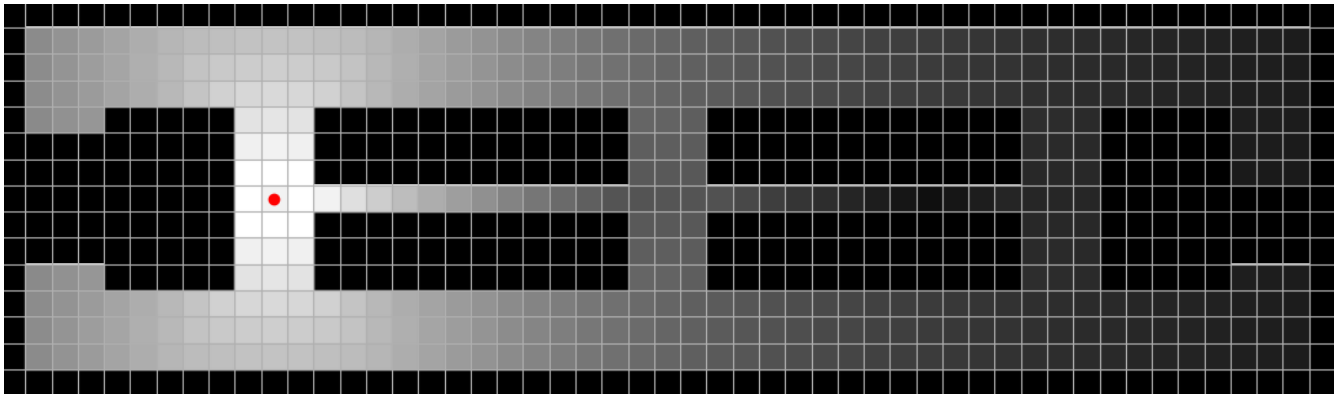
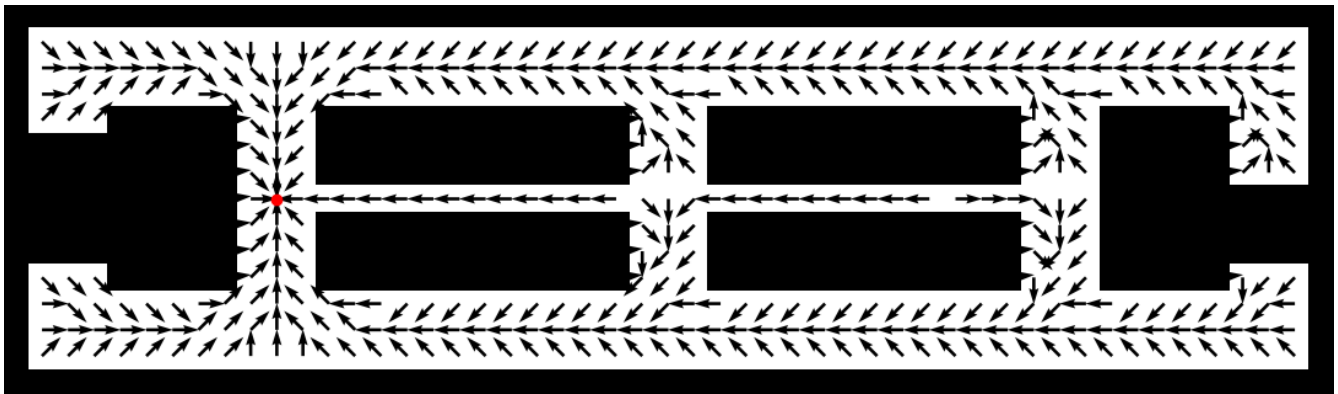
Generalized Policy Iteration



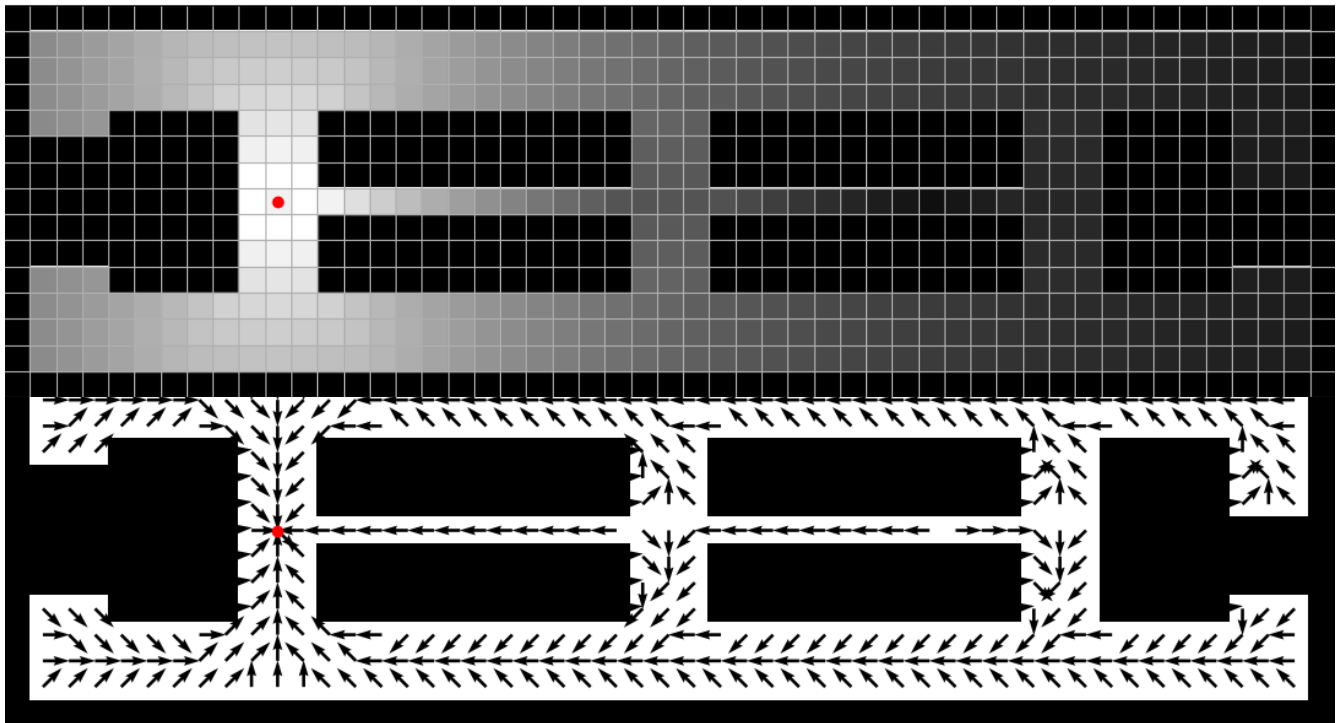


Stochastic

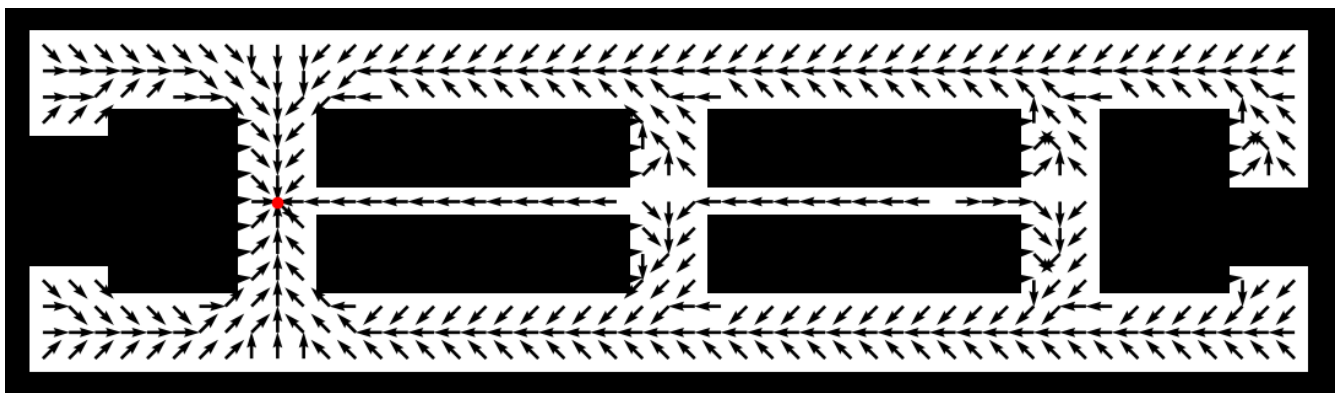
Policy Iteration

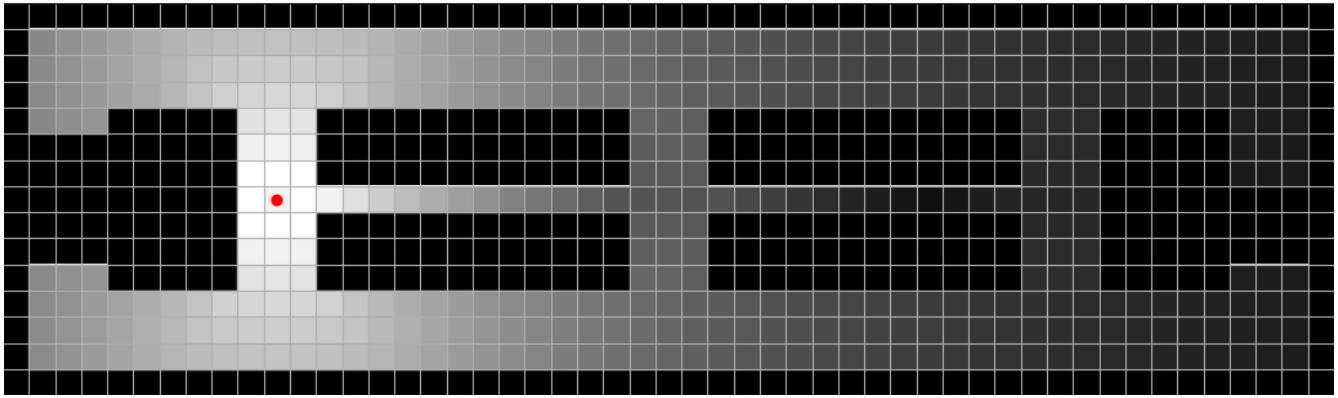


Value Iteration



Generalized Policy Iteration





Discussion of Results

Our plots closely match the plots provided:

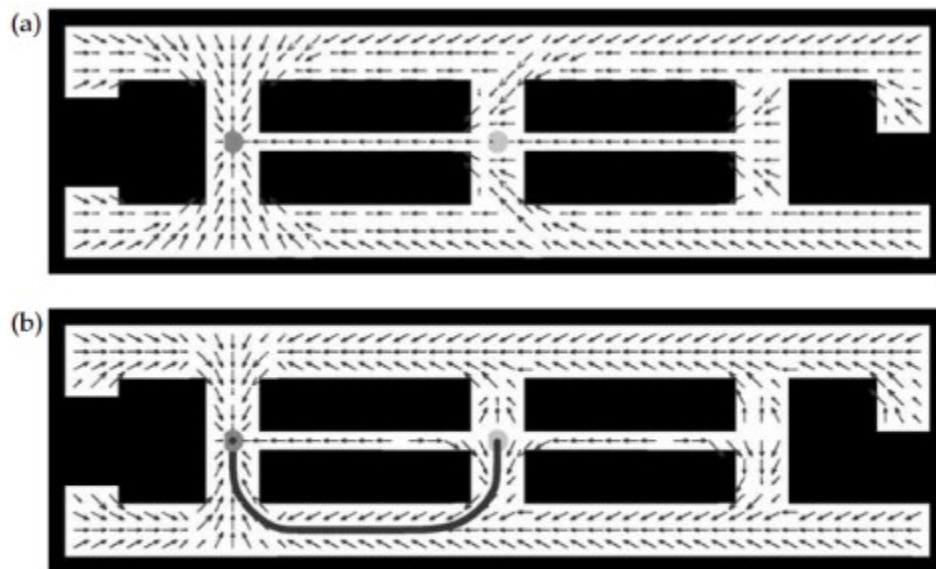
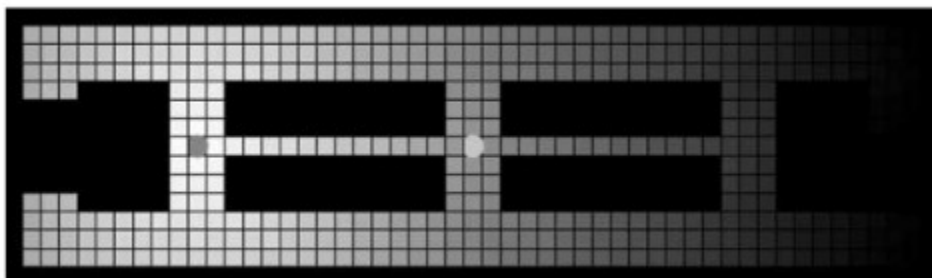


Figure 14.2 The value function and control policy for an MDP with (a) deterministic and (b) nondeterministic action effects. Under the deterministic model, the robot is perfectly fine to navigate through the narrow path; it prefers the longer path when action outcomes are uncertain, to reduce the risk of colliding with a wall. Panel (b) also shows a path.



Our value plots are identical to the provided plots.

Our policy plots are nearly the same. Our stochastic plots are identical, but our deterministic plots vary slightly. Our deterministic model avoids the outer wall more than in the provided plot. For example, in the topmost long corridor, we have two rows of “down and to the right” preferred action, rather than one in the image provided. However, our deterministic and stochastic relationship is the same as in the assignment. The deterministic model the robot prefers traveling through the narrow center passageway, since actions are guaranteed; however, in the stochastic model, the robot avoids the narrow passageway, as there is a 20% chance it will run into the wall and get penalized. Because the overall behavior of our deterministic policy matches that of the assignment, and all the policies make logical sense, the slight differences in preferred actions could result from slight differences in implementation.