Exercise 4.1: Policy Evaluation using 2 Array

Bookmark this page

Policy Evaluation calculates the value function for a policy, given the policy and the full definition of the associated Markov Decision Process. The full definition of an MDP is the set of states, the set of available actions for each state, the set of rewards, the discount factor, and the state/reward transition function.

In this exercise, you will implement the algorithm for Iterative Policy Evaluation using the 2 array approach. In the 2 array approach, one array holds the value estimates for each state computed on the previous iteration, and one array holds the value estimates for the states computing in the current iteration.

Make sure that you have:

1. Completed the setup requirements as described in the Set Up Lab Environments section
2. Downloaded and extracted the Lab files

Now, run jupyter notebook and open the “Ex4.1A Policy Evaluation 2 Arrays.ipynb” notebook.

Examine the notebook and implement the algorithm in the appropriate location.

Make sure you don’t change the function signature for the primary function you are implementing, and the call to the tester code that verifies its correctness.

When you finish your implementation of the function, execute the code cell and verify that the code passes. If it does, save the printed “passcode” value for when you later submit your results on the course webpage below. If it doesn’t pass, correct your code and try again.

Lab Question

1/1 point (graded)

What is the "pass-code" you received upon successfully implementing the policy evaluation with 2 arrays?

9986-145

$9986 - 145$

Submit You have used 1 of 5 attempts

Save Show Answer

Correct (1/1 point)

Previous Next