

CSCI/ROBO 7000/4830: Deep Reinforcement Learning and Robotics Fall 2025 Homework #1: The Wormhole Grid

Steve Gillet

September 23, 2025

Part 1: Policy and Value Analysis [40 Points]

1. Policy Evaluation [15/40 Points]

Consider the following simple, "go-up-and-right" policy, π_{simple} : in every state, the agent attempts to move North. If North is blocked, it tries to move East. If both are blocked, it moves South.

A) Calculate the state-value function, $V^{\pi_{simple}}(s)$, for this policy for all states. Fill in your values on a 4x4 grid.

The state-value function is calculated using the Bellman expectation equation:

$$v^{\pi}(s) = \sum_a \pi(a|s) \sum_{s',r} p(s',r|s,a) [r + \gamma v^{\pi}(s')]$$

Since the environment is deterministic and the policy selects a single action per state, this simplifies to:

$$v^{\pi}(s) = r + \gamma v^{\pi}(s')$$

where s' is the next state and r is the reward (with $\gamma = 0.9$).

The values are as follows (grid with row 0 at the top):

| | | | |
|--|-----|-----|--|
| | | | |
| | N/A | N/A | |
| | | | |
| | | | |

B) Show the setup for your calculations for at least two non-terminal states.

For state (0,0):

$$v(0,0) = -1 + \gamma(-1 + \gamma \cdot 50)$$

For state (3,1):

$$v(3,1) = -1 + \gamma(-1 + \gamma(-1 + \gamma(-1 + \gamma \cdot (-50))))$$