

GridWorld

Christopher Torrington

May 2, 2023

1 Reasonings

1.1 Overall

This problem provides a perfect model of the environment as a Markov Decision Process (MDP), & so I shall use Dynamic Programming algorithms to compute the optimal policies. This is an episodic task as the agent can complete the game once it reaches the terminal states. This allows the problem to be modelled as as finite MDP.