Assignment 4
Tanner Jones

Gridworld

Gridwork is a simple “maze-like” problem. The actor would perform multiple actions (north, south, east, or west) eventually attempting to get to the final “escape” state as quickly as possible. The world was backed by a binary two dimensional array map that described available states and states that would not be accessible by the actor. To demonstrate the full capabilities of the algorithms two maps were created. A smaller map made up of a 10x10 grid to show the basic functionality of the problems. Also a larger map was created. The 50x50 grid allowed a glimpse of what happens when these algorithms had to deal with much more complex tasks. This task can apply in many fields from simple path planners to obstacle avoidance in self driving vehicles. Furthermore this problem is very valid to test our algorithms.

Algorithms

Policy Iteration is the processes of creating policies and finding the utilities under those policies. This iteratively finds if changing the policy could raise the utility until it converges.

Value Iteration evaluates states via the Bellman’s equation until a solution converges. This propagates using states next to the reward until all state’s utility are found.

Q-Learning creates q-values, which are utility values for state action pairs, and an agent will assign new q-values based on a calculated perceived reward. This continues until all true q values are discovered and the algorithm converges on the optimal solution.