Prisoners Dilemma Episode Reward Min Config Size Small 50 Medium Large Obstacle 40 Policy Reward Min 05 05 10 0 0 2 1 3 4 5 1e6 Total steps