Prisoners Dilemma 5 Seed Episode Reward Max Config Size 250 Small Medium Large 200 Policy Reward Max 00 01 50 0 1 2 3 4 5 0 1e6 Total steps