Stag Hunt 5 Seed Self-Play Episode Reward Max Config Size Small Medium 120 Large Obstacle 100 Policy Reward Max 80 60 40 20 0 1 2 3 5 0 4 1e6 Total steps