

PPO + invalid action masking + diverse opponents

