

PPO + invalid action masking + half self-play / half bots + encoder-decoder

