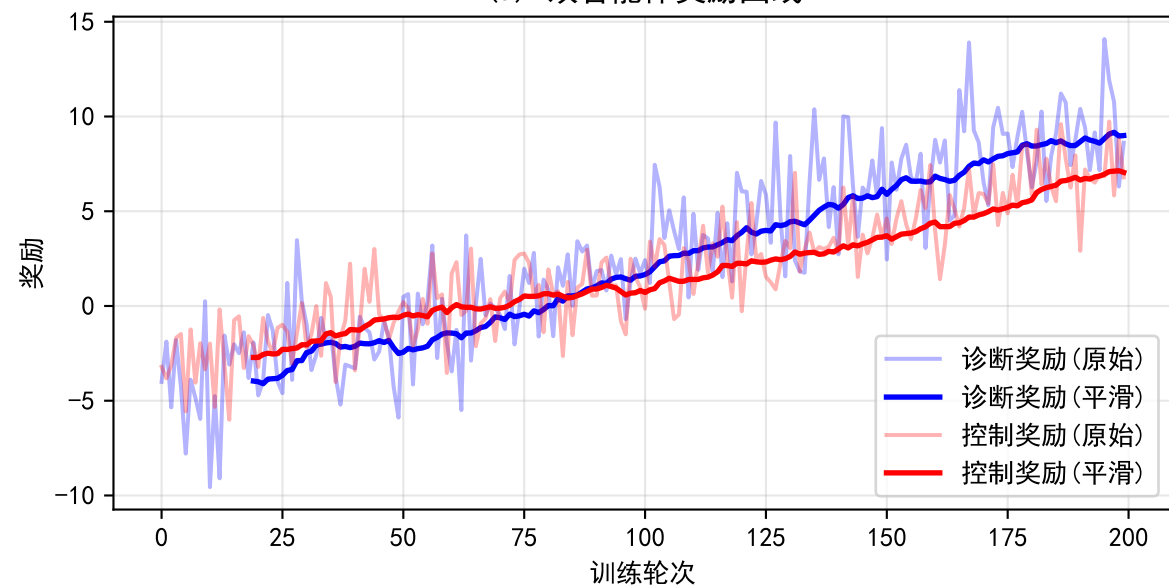
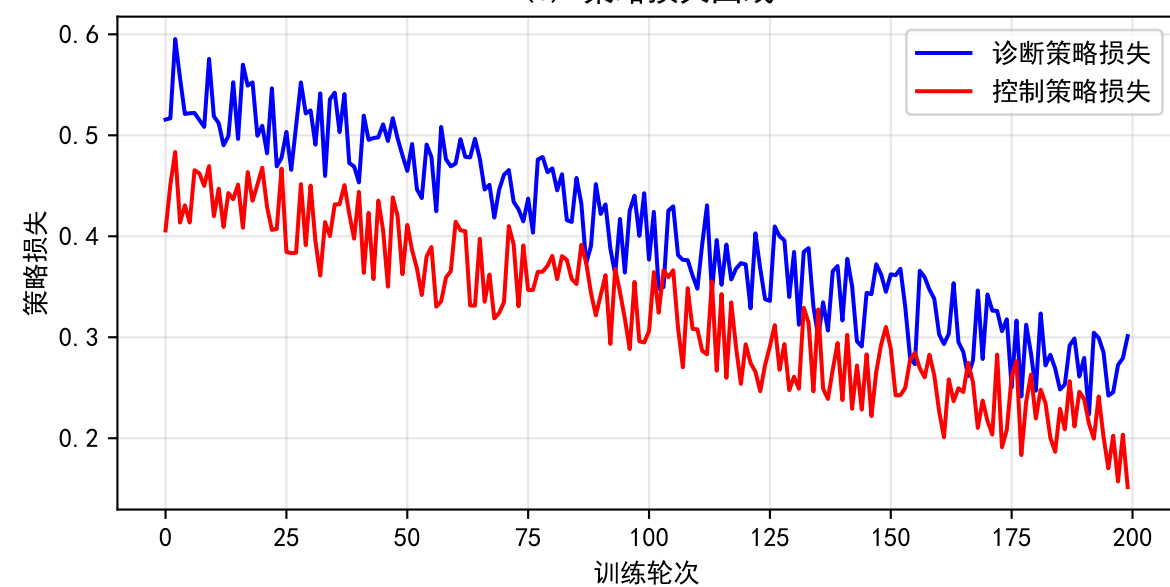


# 双智能体强化学习训练过程

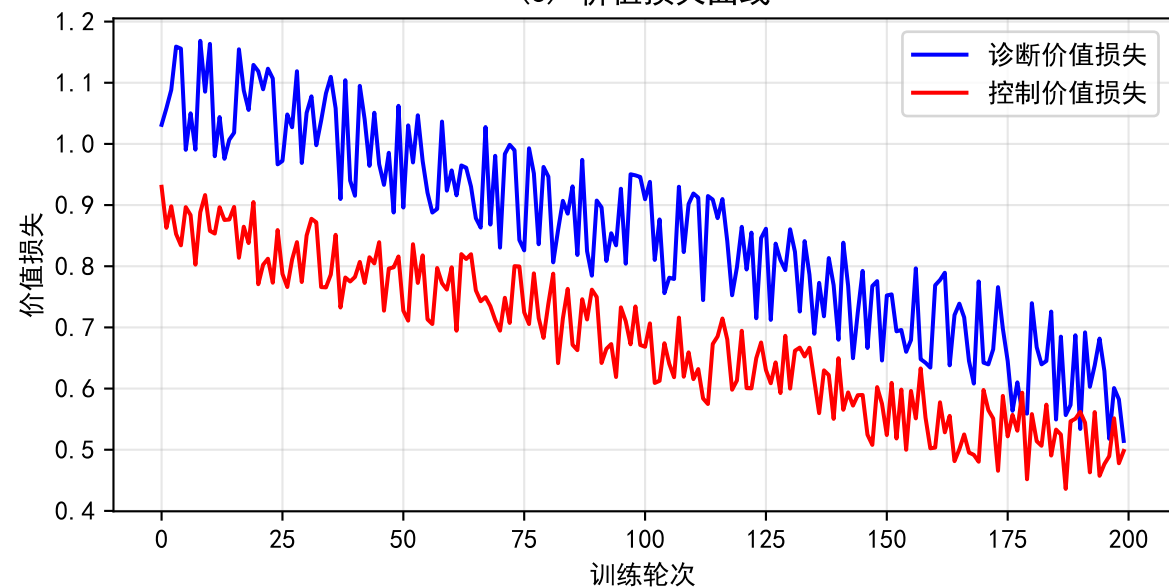
(a) 双智能体奖励曲线



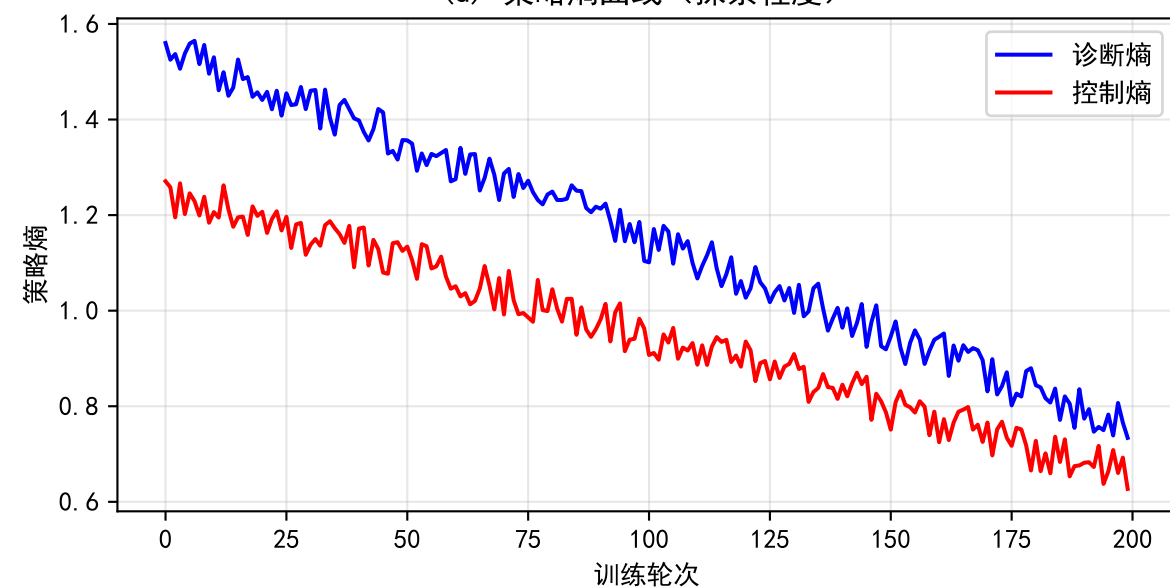
(b) 策略损失曲线



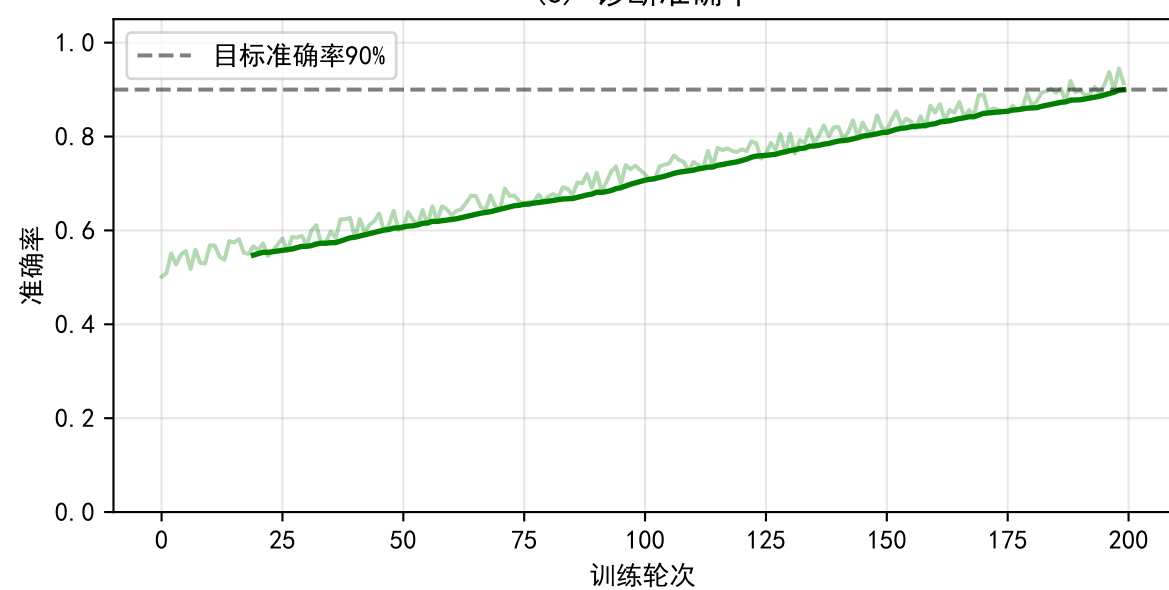
(c) 价值损失曲线



(d) 策略熵曲线 (探索程度)



(e) 诊断准确率



(f) 控制性能 (Pmax维持率)

