

序贯决策问题



马尔科夫决策过程MDP(s,A,P,R,gamma)

(S,A,P,R,gamma)

(S,A,P?,R?,gamma?)



基于模型的动态规划方法



无模型的强化学习方法