

人工智能基础

作业 7

注意：

- 1) 请在网络学堂提交**电子版**；
- 2) 请在**12月8日晚23:59:59**前提交作业，**不接受补交**；
- 3) 4道题目中任选2道解答(多做不加分；4题全做则按题目的解答顺序，只计前2题的分数，如提交作业中题目解答顺序是1、3、2、4,则第2、4题不计分)。
- 4) 如有疑问，请联系助教：

杨鹏帅: ypsl8@mails.tsinghua.edu.cn

鄢启进: yqjl7@mails.tsinghua.edu.cn

崔雪建: cuixjl9@mails.tsinghua.edu.cn

高子靖: gzj21@mails.tsinghua.edu.cn

鲁永浩: yonghao.lu@foxmail.com

牛家赫: njh20@mails.tsinghua.edu.cn

江 澜: jiangl20@mails.tsinghua.edu.cn

尹小旭: yxx21@mails.tsinghua.edu.cn

1. 针对以下三种模型，各举一个生活学习中的实际例子，用课堂上学的描述体系进行建模。

(1) 马尔可夫过程：写出状态空间 S 及状态转移矩阵 P 。

(2) 马尔可夫回报过程：写出状态空间 S 、状态转移矩阵 P 、状态期望回报 r 和折现因子 γ 。

(3) 马尔可夫决策过程：写出状态空间 S 、行动集合 A 、状态转移矩阵 P 、行动期望回报 R 、折现因子 γ 。

2. 甲乙两人进行比赛，设每局比赛甲胜的概率是 p ，乙胜的概率是 q ，平局的概率是 r ，有 $p+q+r=1$ 。设每局比赛后，胜者记‘+1’，负者记‘-1’，平局不记分。当两人中一人积2分时，比赛结束。用 X_n 表示比至第 n 局结束，甲获得的分数，则序列 $\{X_1, X_2, \dots\}$ 为一个马尔可夫过程。

(1) 请给出状态空间和状态转移矩阵。

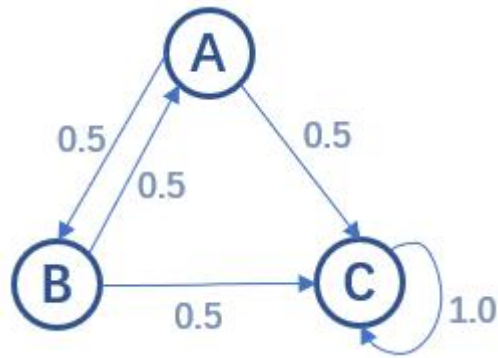
(2) 问在甲积1分的情况下，再赛两局可以结束比赛的概率是多少？

3. 考虑下方一个 3×3 网格图

	1	2
3	4	5
6	7	

非终止状态集合 $S = \{1, 2, \dots, 7\}$ 。每个状态有四种可能的动作 $\{\text{up}, \text{down}, \text{left}, \text{right}\}$ ，对于每次转移 $R_t = -1$ ，每个动作会导致状态转移，但当动作会导致智能体移出网格时，状态保持不变。 $\gamma = 1$ ，若 π 是等概率随机策略，那么 $q_\pi(6, \text{up})$ 、 $q_\pi(5, \text{down})$ 是多少？

4. 如图所示的是 A、B、C 三种状态及其转移概率。状态期望回报除 C 为 0 外，其余均为-1。



- (1) 考虑折现因子 $\gamma = 0.5$ ，求解三种状态的状态价值。
- (2) 在模型已知的情况下，简单情况可直接求解。若直接求解困难，可以采用什么方法解决？试写出求解思路。