

人工智能原理

作业 6

注意：

- 1) 请在网络学堂提交**电子版**；
- 2) 请在 **2025 年 5 月 21 日晚 23:59:59** 前提交作业，**不接受补交**；
- 3) 4 道题目中选择 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 道需做题目的分数，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：

李 震：lizhen22@mails.tsinghua.edu.cn

李可伊：lky23@mails.tsinghua.edu.cn

王子安：wangza24@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 \mathbf{S} 及状态转移矩阵 \mathbf{P} 。
- 2) 马尔可夫回报过程：写出状态空间 \mathbf{S} 、状态转移矩阵 \mathbf{P} 、状态期望回报 \mathbf{r} 和折现因子 γ 。
- 3) 马尔可夫决策过程：写出状态空间 \mathbf{S} 、行动集合 \mathbf{A} 、状态转移矩阵 \mathbf{P} 、行动期望回报 \mathbf{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。当两人中一人积到 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。

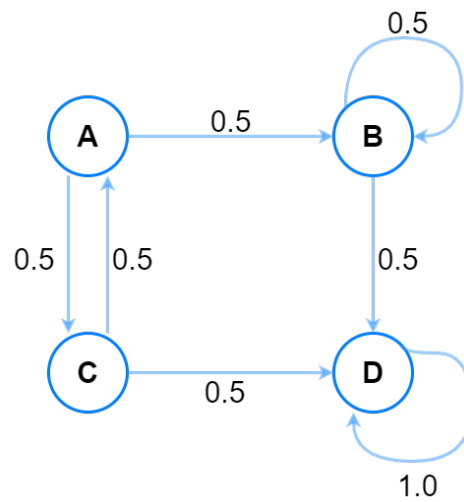
- 1) 请给出状态空间和状态转移矩阵。
- 2) 问在甲积 1 分的情况下，恰好再赛两局可以结束比赛的概率是多少？

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 $\{\text{up、down、left、right}\}$ ，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，那么行动价值 $q_\pi(4, \text{left})$ 、 $q_\pi(7, \text{right})$ 是多少？

4. 如图所示是 A、B、C、D 四种状态及其转移概率，状态期望回报除了 D 为 0 外，其余均为-1.



- (1) 考虑折现因子 $\gamma=0.5$ 时，这四种状态的状态价值。
- (2) 在模型已知且问题规模较小的情况下，该问题可以直接进行求解。若问题规模较大，可以采用什么方法解决？试简单写出求解思路。