

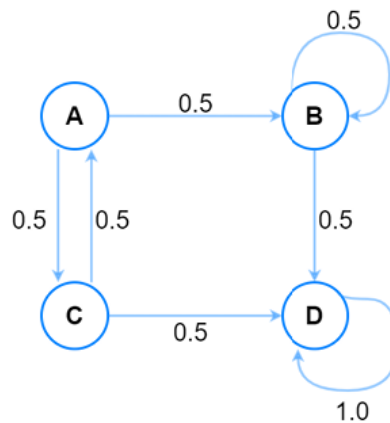
人工智能原理

作业 7

注意：

- 1) 请在网络学堂提交电子版；
- 2) 请在 5 月 30 日晚 23:59:59 前提交作业，不接受补交；

1. 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，乙获得的分数，则序列 $\{X_1, X_2, \dots\}$ 为一个马尔可夫过程。
 - a) 请给出状态空间和状态转移矩阵。
 - b) 问在乙积 1 分的情况下，再赛两局可以结束比赛的概率是多少？
2. 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。



- a) 考虑折现因子 $\gamma=0.2$ 时，这四种状态的状态价值；
 - b) 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；
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，那么行动价值 $q_\pi(4, \text{left})$ 、 $q_\pi(7, \text{right})$ 是多少？