

1. 举例: 一天的状态 $S = \{S_1, S_2, S_3\} = \{\text{学习}, \text{吃饭}, \text{睡觉}\}$

$$\text{转移矩阵 } P = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.5 & 0 & 0.5 \\ 0 & 0.1 & 0.9 \end{bmatrix}$$

(2) 举例: 上课状态 $S = \{S_1, S_2, S_3\} = \{\text{听课}, \text{走神}, \text{休息}\}$

$$P = \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.6 & 0.2 & 0.2 \\ 0.9 & 0 & 0.1 \end{bmatrix}$$

$$r = [-1 \quad 0.1 \quad 0.5]^T \quad \gamma = 0.9$$

(3) 马尔可夫决策举例: 状态 $S = \{S_1, S_2\} = \{\text{出bug}, \text{运行}\}$

行动集合 $A = \{A_1, A_2\} = \{\text{debug}, \text{摸鱼}\}$

状态转移矩阵 $A=A_1$ $P_{SS}^A = \begin{bmatrix} 0.9 & 0.1 \\ 0.8 & 0.2 \end{bmatrix}$

$A=A_2$ $P_{SS}^A = \begin{bmatrix} 0.1 & 0.9 \\ 0.2 & 0.8 \end{bmatrix}$

$$R =$$

	S_1	S_2
A_1	-10	10
A_2	-1	1

2. (1) 状态空间为 $S = \{S_1, S_2, S_3, S_4, S_5\}$ 分别表示甲获得 -2, -1, 0, 1, 2 分

状态转移矩阵 P 为

	S_1	S_2	S_3	S_4	S_5
S_1	1	0	0	0	0
S_2	q	r	p	0	0
S_3	0	q	r	p	0
S_4	0	0	q	r	p
S_5	0	0	0	0	1

(2) 甲积1分的情况下,再赛两局结束比赛 设为事件A

可能的情形为先平局再甲赢. $P(A) = r \cdot p$

$$3. \quad G_1 = b + 0.9 \times b + 0.9^2 \times b + \dots$$

$$= b \times \frac{1 - 0.9^{n+1}}{1 - 0.9}$$

$$= 60.$$

$$G_0 = 2 + r \times G_1 = 2 + 0.9 \times 60 = 56$$