人工智能原理-作业6

Author: 夏弘宇 2023011004

2-1. 状态空间 $S=\{-2,-1,0,1,2\}$

状态转移矩阵

$$P=\begin{array}{c c} \begin{array}{ccccc} -2 & -1 & 0 & 1 & 2 \end{array} & \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & r & p & 0 & 0 \\ 0 & q & r & p & 0 \\ 0 & 0 & q & r & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} & \begin{array}{c} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{array} \end{array}$$

2-2 只能先平后赢 $P=pr$

4-1 由题意 $r=(-1\ \ -1\ \ -1\ \ 0)^T$

状态转移矩阵

$$P=\begin{bmatrix} 0 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

贝尔曼期望方程. $v(s)=r+\gamma\sum_{s'\in S}P_{ss'}v(s')$ 即 $\vec{v}=\vec{r}+\gamma P\vec{v}$

解得 $\vec{v}=(I-\gamma P)^{-1}\vec{r}=\begin{bmatrix} 1 & -0.25 & -0.25 & 0 \\ 0 & 0.75 & 0 & -0.25 \\ -0.25 & 0 & 1 & -0.25 \\ 0 & 0 & 0 & 0.5 \end{bmatrix}^{-1}\begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \end{bmatrix}=\begin{bmatrix} -\frac{76}{45} \\ -\frac{4}{3} \\ -\frac{64}{45} \\ 0 \end{bmatrix}$

即四状态的状态价值分别为 $-\frac{76}{45}, -\frac{4}{3}, -\frac{64}{45}, 0$.

4-2 迭代的方法(动态规划)

① 对 $\vec{v}=(0\ \ 0\cdots0)^T$ 设置任意初值如这里的 0.

② 迭代 $\vec{v}^{(k+1)}=\vec{r}+\gamma P\vec{v}^{(k)}$

③ 当相邻两轮 $\vec{v}$ 值均小于某阈值时停止，得到最终的价值函数