人工智能第七次作业

1　第一题

在这道题目中，将体育锻炼作为实际例子，进行建模。

1.1　马尔可夫过程

首先，锻炼的种类包括了羽毛球、健身和游泳，休息的方式包括了洗澡、睡觉，以及在体育锻炼中的间歇划水，由于只在健身过程中产生厌烦情绪，因此划水只会在健身过程中出现。同时由于需要补充能量，在体育锻炼中间会出现间歇休息，并且洗澡活动只在大量的羽毛球和健身锻炼之后出现。

因此得到的马尔可夫过程的状态空间 S 如下：

$$S = \{羽毛球, 健身, 游泳, 洗澡, 睡觉, 划水, 休息\} \tag{1}$$

将状态转移概率图显示图下，在显示的过程中会隐含状态转移矩阵。

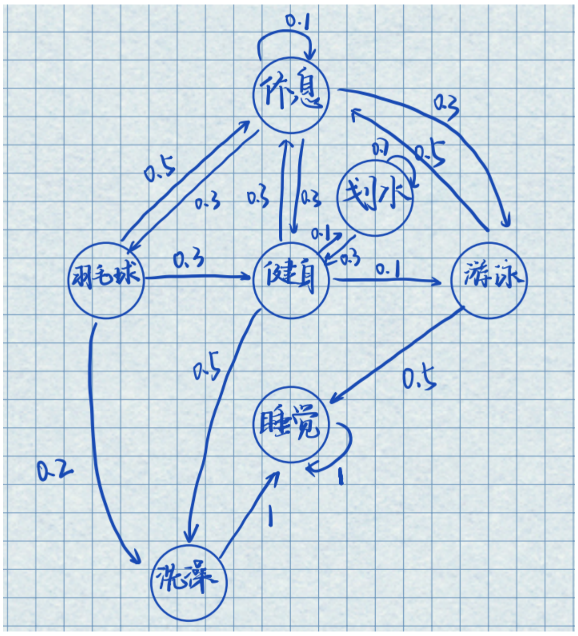

图 1: 状态转移图

因此得到状态转移概率如下：

$$P = \begin{bmatrix} 0 & 0.3 & 0 & 0.2 & 0 & 0 & 0.5 \\ 0 & 0 & 0.1 & 0.5 & 0 & 0.1 & 0.3 \\ 0 & 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0.3 & 0 & 0 & 0 & 0.7 & 0 \\ 0.3 & 0.3 & 0.3 & 0 & 0 & 0 & 0.1 \end{bmatrix} \tag{2}$$

1.2 马尔可夫回报过程

在马尔可夫回报过程中，需要折现因子和状态期望回报。
因此在如下图所示的状态转移图中包含了状态期望回报的数值如下。

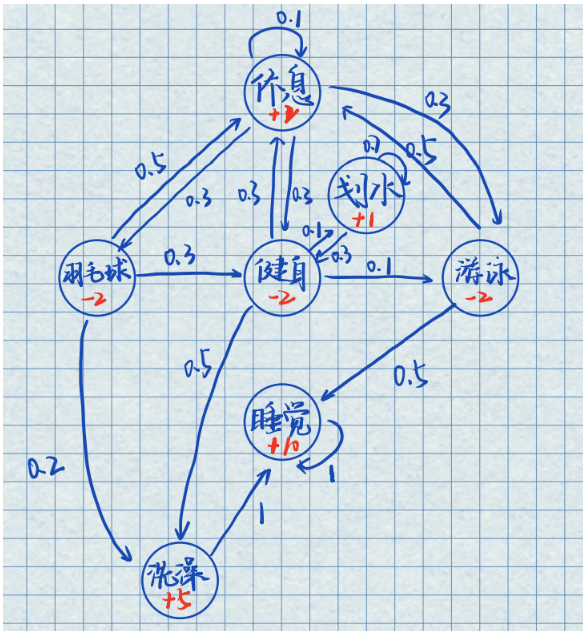

图 2: 包含状态期望回报的状态转移图

设置在这样的状态转移图下的折现因子为 $\gamma = 0.1$，得到相应的状态期望回报显示如下：

$$r = \{-2, -2, -2, +5, +10, +1, +2\} \tag{3}$$

马尔可夫回报过程中的状态空间 S 和状态转移矩阵 P 和马尔可夫过程中的是保持一样的。

1.3 马尔可夫决策过程

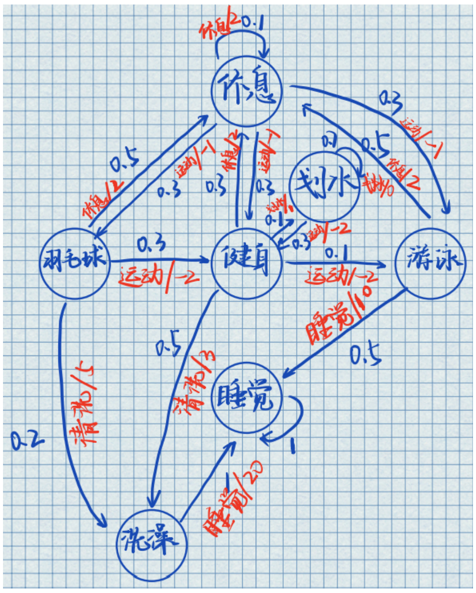

图 3: 马尔可夫决策过程的状态转移图

状态空间 S 和马尔可夫过程中的相同，也就是公式 (1)，行动空间 A 如下：

$$A = \{运动, 休息, 清洗, 划水, 睡觉\} \tag{4}$$

状态转移矩阵 P 是一个三维的矩阵 $P = (p^a_{ss'})_{nxnxm}$。由于使用图像来表示一个三维矩阵过于复杂，这里使用数据的形式来表示，如下，羽毛球是状态 1，健身是状态 2，游泳是状态 3，洗澡是状态 4，睡觉是状态 5，划水是状态 6，休息是状态 7。运动是行动 1，休息是行动 2，清洗是行动 3，划水是行动 4，睡觉是行动 5。

同时设置在状态睡觉的时候如果重复进行睡觉行动的话，期望回报数值为 1。对于状态羽毛球如下：

$$\begin{aligned} p^1_{12} &= 1 \\ p^3_{14} &= 1 \\ p^2_{17} &= 1 \end{aligned} \tag{5}$$

对于状态健身如下：

$$\begin{aligned} p^3_{24} &= 1 \\ p^1_{23} &= 1 \\ p^4_{26} &= 1 \\ p^2_{27} &= 1 \end{aligned} \tag{6}$$

对于状态游泳如下：

$$\begin{aligned} p^5_{35} &= 1 \\ p^2_{37} &= 1 \end{aligned} \tag{7}$$

对于状态洗澡如下：

$$p^5_{45} = 1 \tag{8}$$

对于状态睡觉如下:

$$p_{55}^{5} = 1 \tag{9}$$

对于状态划水如下:

$$\begin{aligned} p_{66}^{4} &= 1 \\ p_{62}^{1} &= 1 \end{aligned} \tag{10}$$

对于状态休息如下:

$$\begin{aligned} p_{71}^{1} &= \frac{1}{3} \\ p_{72}^{1} &= \frac{1}{3} \\ p_{73}^{1} &= \frac{1}{3} \\ p_{77}^{5} &= 1 \end{aligned} \tag{11}$$

在这个三维矩阵中的其他的所有元素都是 0。

折现因子设置为 0.1。

得到的行动期望回报矩阵 R 如下所示:

$$R = \begin{bmatrix} -2 & 2 & 5 & 0 & 0 \\ -2 & 2 & 3 & 1 & 0 \\ 0 & 2 & 0 & 0 & 10 \\ 0 & 0 & 0 & 0 & 20 \\ 0 & 0 & 0 & 0 & 1 \\ -2 & 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 & 0 \end{bmatrix} \tag{12}$$

2 第二题

2.1 状态空间和状态转移矩阵

状态空间 $S = \{-2, -1, 0, 1, 2\}$,状态转移矩阵如下,不妨将状态空间映射到 $S^{'} = \{1, 2, 3, 4, 5\}$,其中 $S[i] - -S^{'}[i]$。

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & r & p & 0 & 0 \\ 0 & q & r & p & 0 \\ 0 & 0 & q & r & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{13}$$

2.2 在甲积 1 分的情况下，再赛两局可以结束比赛的概率

甲积 1 分的话，可以先平局一局，再获胜一局即得到甲胜利，乙如果想要胜利的话，至少需要比赛 3 局。

即 $p = rp$。