T2

1. 状态空间S与状态转移矩阵P如下所示

$$S = \{(0,0),(-1,1),(1,-1),(-2,2),(2,-2)\}$$

$$\mathbf{P} = \begin{bmatrix} r & q & p & 0 & 0 \\ p & r & 0 & q & 0 \\ q & 0 & r & 0 & p \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{P}^T = \begin{bmatrix} r & p & q & 0 & 0 \\ q & r & 0 & 0 & 0 \\ p & 0 & r & 0 & 0 \\ 0 & q & 0 & 1 & 0 \\ 0 & 0 & p & 0 & 1 \end{bmatrix}$$

2.

$$p_2 = (\mathbf{P}^T)^2 \cdot p_0 = [2qr, q^2, r^2 + pq, 0, pr + p]$$
$$p_1 = \mathbf{P}^T \cdot p_0 = [q, 0, r, 0, p]$$

所以，再赛两局之内可以结束比赛的概率为：$sum(p_2[3:]) = pr + p$

刚好在第二局结束比赛的概率为：$sum(p2[3:]) - sum(p_1[3:]) = pr$

T3

$$\mathbf{S} = \{1,2,3,4,5,6,7\}$$
$$\mathbf{A} = \{up, down, left, right\}$$
$$\mathbf{R} = [-1]$$
$$q_\pi(6, up) = E[G_t | S_t = 6, A_t = up]$$
$$q_\pi(6, up) = -1 + \sum p_{ss'}^a v_\pi(s') = -1 + v_\pi(3)$$
$$v_\pi(3) = \frac{1}{4}[(-1 + v_\pi(3)) + (-1 + 0) + (-1 + v_\pi(4)) + (-1 + v_\pi(6))]$$
$$v_\pi(4) = \frac{1}{4} \cdot 4 \cdot (-1 + v_\pi(3))$$
$$v_\pi(6) = \frac{1}{4}[(-1 + v_\pi(3)) * 2 + (-1 + v_\pi(6)) * 2]$$
$$\therefore v_\pi(3) = -7$$
$$\therefore q_\pi(6, up) = -8$$
$$\text{由对称性，} q_\pi(5, down) = q_\pi(3, up) = -1$$

T2

1.

$$S = \{A, B, C\}$$
$$P = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0 & 0 & 1 \end{bmatrix}$$
$$\mathbf{r} = [-1, -1, 0]^T$$
$$\mathbf{v} = (I - \gamma P)^{-1}\mathbf{r} = [-1.33, -1.33, 0]^T$$

2. 如果模型复杂，可以参考数值分析与算法中学到过的迭代法解方程（组）的方法进行数值求解，通过多次迭代，得到模型的近似（数值）解。