Multi-Agent System 作业二

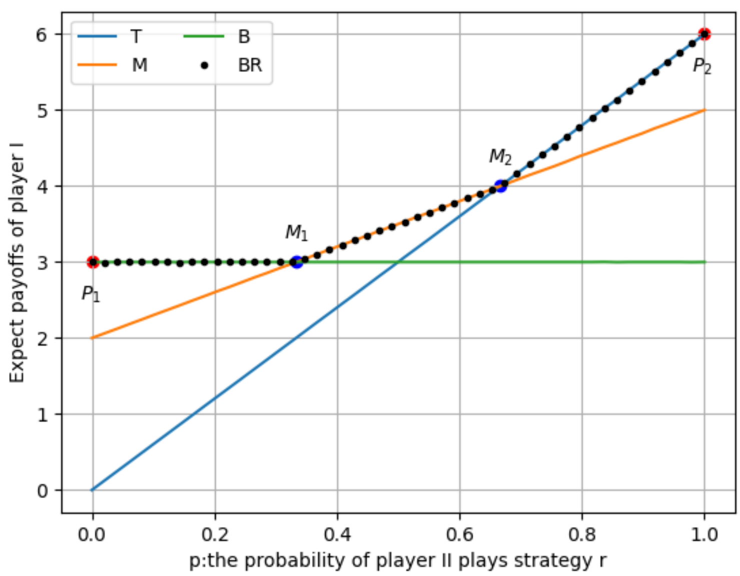
承子杰 202228000243001
数学与系统科学学院，中科院（MASS，CAS）

Problem 1:

问题 1：对图示矩阵博弈，求博弈的全部 Nash 均衡。

I \ II	l	r
T	1 0	0 6
M	0 2	2 5
B	3 3	4 3

Solution:



Nash 均衡只可能在 P_1，P_2，M_1，M_2 处

P_1:当 Player II 选择策略 r 的概率为 0 时，Player I 选择策略 B。可以发现，此时 Player II 的最优策略应该选择 r 的概率为 1，而不是 0。故不是平衡点。

P_2:当 Player II 选择策略 r 的概率为 1 时，Player II 选择策略 T。可以发现，此时 Player I 的最优策略应该选择 r 的概率为 0，而不是 1。故不是平衡点。

M_1: 当 Player II 的混合策略为$Q\left(\frac{2}{3},\frac{1}{3}\right)$,，Player II 选择策略 B 和 M 的混合策略$P(-1,2)$，由于概率取值大于 0，故不是平衡点。

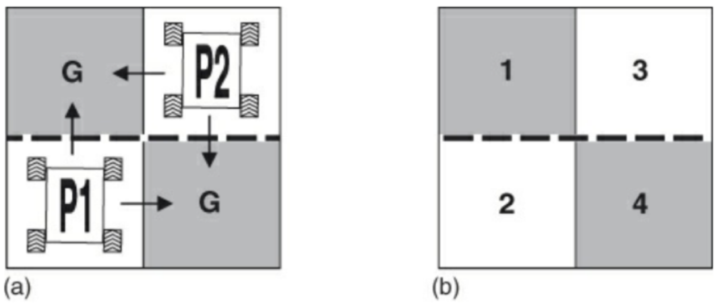
M_2:当 Player II 的混合策略为$Q\left(\frac{1}{3},\frac{2}{3}\right)$, Player I 采取 T 和 M 的混合策略$P(\frac{2}{3},\frac{1}{3})$。则此时，对于 Player II，$E[l,P]=\frac{2}{3},E[r,P]=\frac{2}{3}$。对于 Player I，$E[T,Q]=4,E[M,Q]=4$。故双方均不会发生策略改变，是平衡点。

综上所述：Player I 采取 T 和 M 的混合策略$P\left(\frac{2}{3},\frac{1}{3}\right)$，Player II 采取混合策略$Q\left(\frac{1}{3},\frac{2}{3}\right)$时，达到 Nash 均衡。

Problem 2：

问题 2：针对如下 2*2 的网格博弈，两个玩家 *P1* 和 *P2* 的初始位置分别在左下角和右上角，两个玩家试图在最小步数内到达标为“**G**”的两个目标之一，如图(a)所示。玩家 *P1* 的动作集合为{向上，向右}，玩家 *P2* 的动作集合为{向左，向下}，图(b)给出了游戏中单元格的编号作为状态，如初始状态可表示为$S_1=(2,3)$，第一行单元格和第二行单元格之间的虚线表示障碍，玩家会以 *0.5* 的概率通过障碍，以 *0.5* 的概率保留在原位置。

如果两个玩家移动到同一单元格，奖励为 *0*，返回各自初始位置。
如果有任一玩家到达目标网格，该玩家获得的奖励为 *10*，游戏结束。 令折扣因子为0.9。



1. 在初始状态S_1下，列出两玩家所有联合动作下可能跳转的状态及概率？
2. 此问题的 Nash 均衡策略是什么？给出玩家 *P1* 在均衡策略下的价值函数$V_1(S_1)$。
3. 在初始状态 S_1 下，计算玩家执行联合动作（上,下）后在均衡策略下玩家1的动作-价值函数$Q_1(S_1,上,下)$。
4. 在初始状态 S_1 下，给出玩家 *P2* 在均衡策略下的 Q 表（说明：行表示玩家 *P1* 的动作，列表示玩家 *P2* 的动作）。

Solution：

1. 动作：（上，左），跳转状态：（1，1）概率 p=0.5，（2，1）概率 p=0.5
   动作：（上，下），跳转状态：（1，3）概率 p=0.25，（2，3）概率 p=0.25，（2，4）概率 p=0.25，（1，4）概率 p=0.25
   动作：（右，左），跳转状态（4，1）概率 p=1
   动作：（右，下），跳转状态：（4，4）概率 p=0.5，（4，3）概率 p=0.5

2. Nash 均衡策略为不跨越障碍，两个玩家都到达其相邻的位置，即状态（4，1）。$V_1(S_1) = 10$

3. $Q_1(S_1, 上，下) = 0.25 * 0 + 0.25 * 0.9 * 10 + 0.25 * 10 + 0.25 * 10 = 7.25$

4. Q 表

$Q_2^*(S_1, a_1, a_2)$	上	右
下	7.25	4.5
左	9.5	10