

机器智能实验 8

博弈论实验- 重复囚徒困境实验报告

实验步骤:

1. 实验场景, 如下所示, 有一个 controller 用于控制整体博弈流程, playerA 和 playerB 是博弈的两名参与者



2. 博弈过程建模

对于参与者 player, 定义它的三种状态: default、ready、end

- default 状态
此状态为初始状态, player 的行为是初始化, 并能够通过 touch 改变 player 博弈中的行为
- ready 状态
此状态根据 default 状态中设置的行为, 为每次博弈作出决定。共有五种行为: 总是不合作、总是合作、Tit For Tat、Tit For Two Tats、随机决定
- end 状态
所有博弈完成后进入此状态, 等待重新初始化并开始新的博弈

参与者的状态由控制器 controller 控制

下图展示了一次博弈的结果, 其中 playerA 采取 TFT 策略, playerB 采取随机策略

```
controller: Round 191, playerA choose C, playerB choose D
controller: Round 192, playerA choose D, playerB choose C
controller: Round 193, playerA choose C, playerB choose C
controller: Round 194, playerA choose C, playerB choose D
controller: Round 195, playerA choose D, playerB choose D
controller: Round 196, playerA choose D, playerB choose C
controller: Round 197, playerA choose C, playerB choose C
controller: Round 198, playerA choose C, playerB choose C
controller: Round 199, playerA choose C, playerB choose C
controller: Round 200, playerA choose C, playerB choose C
controller: Game ends, A score = 461, B score = 461
```

3. 重复的囚徒困境实验

对每两种博弈策略进行循环赛, 每次比赛进行五轮, 每轮迭代 200 次得到得分矩阵, 将五轮分数求均值得到下表

策略	ALL-D	ALL-C	TFT	TF2T	Random	平均得分
ALL-D	200:200	1000:0	204:199	208:198	594.4:101.4	441.28
ALL-C	0:1000	600:600	600:600	600:600	303.6:797.6	420.72
TFT	199:204	600:600	600:600	600:600	447.2:449.2	489.24
TF2T	198:208	600:600	600:600	600:600	379.8:636.8	475.56
Random	101.4:594.4	797.6:303.6	447.2:449.2	379.8:636.8	446.4:445.4	434.48

从表中可知, ALL-C 的博弈策略最差, TFT 的博弈策略最好 (在假定采用各个博弈策略的人数相等的情况下)