TOURNAMENT

吳海韜, 黃奎鈞, 巫佳樺

June 13, 2018

Introduction

在一場捉對廝殺的競賽中，兩位玩家(p1,p2)在每一局中同時決定合作/背叛：

(p1得分,p2得分)	p2合作	p2背叛
p1合作	(2,2)	(-1,3)
p1背叛	(3,-1)	(0,0)

如此連續進行N局，最後結算雙方的平均得分

Definition(1)

- 「一局」比賽將有N輪，其中雙方無法探知確切輪數、且非常多$N \to \infty$
- 在第n輪（round）中雙方玩家$\{p_1, p_2\}$要做出決策$\{$合作,背叛$\} \mapsto \{1, 0\}$
- 可能agents的集合稱為「池子」
  - 代號：$\{A_1, A_2, ..., A_k\}$
  - 一場捉對比賽抽到對手A_i的機率是$\frac{1}{k}$

Definition(2)

- 本局比賽前n輪的狀態矩陣ϕ_n

$$\phi_n = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \\ r_{31} & r_{32} \end{bmatrix}$$

- 當我們和Agent對局的時候我們是player1（column1）、Agent是player2（column2）每一輪Agent的反應函數記作：

$$A_i(\phi_{n-1}) = r_n = \begin{cases} 1, & \text{cooperate} \\ 0, & \text{betray} \end{cases}$$

- 也就是依據前n-1局(n-1 rows)的資訊來做第n局的決策r_n
  - 假設我們是p_1、agent是p_2來說，agent第n輪的回應就是新矩陣ϕ_n中的r_{n2}值

Agent反應函數(1)

- AgentMu

$$\text{AgentMu}_{\mu}(\phi_{n-1}) = X_{\mu}\ ,\ X_{\mu} \sim Bernoulli(\mu)$$

- AgentQ
  - 透過ϕ_{n-1}去估計對手可能的合作機率
  - 第n局時，定義觀察到前n-1局對手合作的次數是m：

$$m = \sum_{i=1}^{n-1} r_i^{opponent}$$

  - 在完全沒有任何資訊的情況下，我們希望初始機率是$\frac{1}{2}$於是調整估計式q為：

$$\begin{aligned} q &= \frac{(m)+1}{(n-1)+2} \\ &= \frac{m+1}{n+1} \end{aligned} \Rightarrow \text{AgentQ}(\phi_{n-1}) = X_p\ ,\ \begin{cases} X_p \sim Bernoulli(p) \\ p = q = \frac{m+1}{n+1} \end{cases}$$

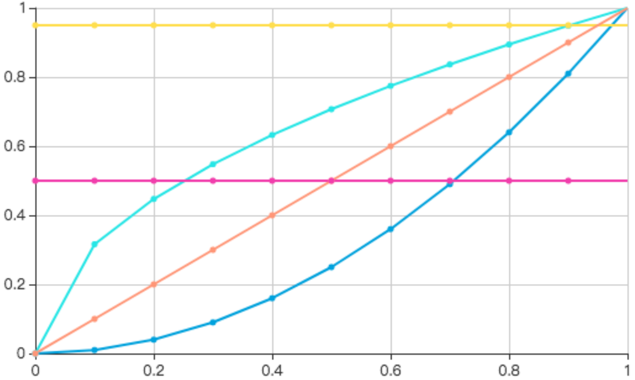
Agent反應函數(2)

- AgentGamma：另加入保守係數γ，改寫為AgentGamma如下：

$$0 < q^{\gamma} < 1 \text{ , } 0 < \gamma < \infty$$

$$\text{AgentGamma}_{\gamma}(\phi_{n-1}) = X_p \text{ , } \begin{cases} X_p \sim Bernoulli(p) \\ p = q^{\gamma} = (\frac{m+1}{n+1})^{\gamma} \text{ , } 0 < \gamma < \infty \end{cases}$$

Agent的行為



運用貝氏定理推算對手類型(1)

$$P(A_i \mid \phi_n) = \frac{\overbrace{P(\phi_n \mid A_i)}^{Likelihood}\, P(A_i)}{\underbrace{P(\phi_n)}_{Evidence}}$$

已知Agent只有k種$\{A_1, A_2, ..., A_k\}$
其中發生對局紀錄ϕ_n的機率$P(\phi_n)$可以由總機率法則：

$$P(A_1 \mid \phi) = \frac{P(\phi \mid A_1)P(A_1)}{P(\phi \mid A_1)P(A_1) + P(\phi \mid A_2)P(A_2) + \cdots + P(\phi \mid A_k)P(A_k)}$$

運用貝氏定理推算對手類型(2)

設定每一種agent出現的機率都是$P(A_1) = P(A_2) = ... = P(A_k) = \frac{1}{k}$
整理後得到：

$$P(A_1 \mid \phi) = \frac{P(\phi \mid A_1)}{P(\phi \mid A_1) + P(\phi \mid A_2) + \cdots + P(\phi \mid A_k)}$$

很直覺地給定發生ϕ_n，對手是A_i的機率就是Likelihood的相對比例：

$$\frac{Likelihood(A_i)}{\sum_{j=1}^{k} Likelihood(A_j)}$$

計算Likelihood(AgentMu)

方便起見以下定義p_n為第n輪心中願意合作的機率：

$$
\begin{aligned}
P(\phi_n \mid \mathrm{AgentMu}_\mu) &= \prod_{n=1}^{N} \begin{cases} p_n & \text{, if } r_n \ \textit{cooperate} \\ 1 - p_n & \text{, if } r_n \ \textit{betray} \end{cases} \\
&= \prod_{n=1}^{N} \begin{cases} \mu & \text{, if } r_n \ \textit{cooperate} \\ 1 - \mu & \text{, if } r_n \ \textit{betray} \end{cases} \quad \because p_n = \mu
\end{aligned}
$$

計算Likelihood(AgentGamma)

AgentGamma在第n輪時會觀察對手合作比率q_n
(是ϕ_n的函數：$q_n = q_n(\phi_n)$)
進而依保守係數γ作指數運算得到心中願意合作的機率：

$$p_n = q_n^\gamma$$

$$P(\phi_n \mid \text{AgentGamma}_\gamma) = \prod_{n=1}^{N} \begin{cases} q_n^\gamma & \text{, if } r_n \ cooperate \\ 1 - q_n^\gamma & \text{, if } r_n \ betray \end{cases}$$

判斷對方方法：分組判斷法

- 目標：得分極大化
- 分成兩組：
  - AgentMu
    - 永遠背叛平均得分$\rightarrow 3\mu$
  - AgentGamma
    - 永遠合作平均得分$\rightarrow 2$
- 問題簡化成：

$$P(A_{Trust} \mid \phi_n) = \frac{\sum_{i \in \mathrm{Trust}} Likelihood(A_i)}{\sum_{j=1}^{k} Likelihood(A_j)}$$

Claim

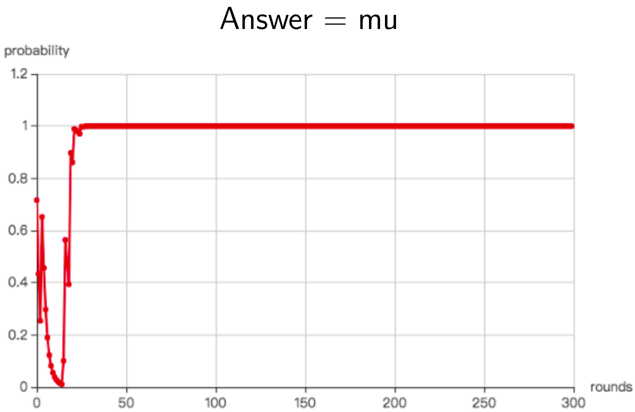
局數趨近無窮多，$P(A_{Trust} \mid \phi_n)$會收斂到真實的對手類型：

$$\lim_{n \to \infty} P(A_{Trust} \mid \phi_n) = \begin{cases} 1 \ , \ A_{Trust} \\ 0 \ , \ \neg A_{Trust} \end{cases}$$

收斂的原因：

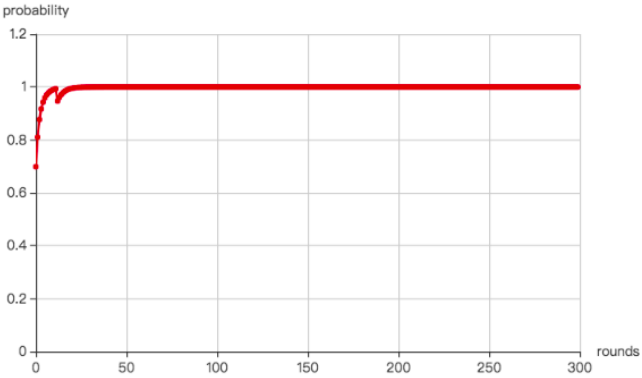
1. 假設*AgentMu*中合作機率最大的$\mu = 1 - \epsilon$
2. 若$\epsilon = 0$, 只要出現一局背叛就能確定$P(A_{\mu=1} \mid \phi_n) = 0$
3. 若$0 < \epsilon < 1$
   - 因為我們全部合作，*AgentGamma*估計的q值會往1收斂
   - 存在n_1使得$n > n_1, q_n > \mu$（對所有*AgentGamma*）
   - 此時所有*AgentGamma*的合作機率都大於*AgentMu*，局數越多越能有效區分

實驗一：mu∼U(0, 1), gamma∼U(0.2, 1.5)



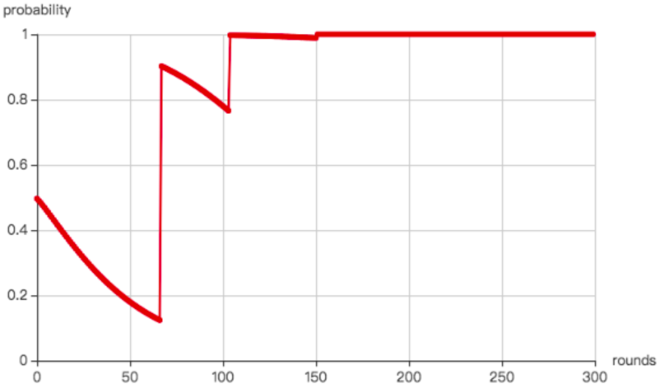
實驗一：mu∼U(0, 1), gamma∼U(0.2, 1.5)

Answer = gamma



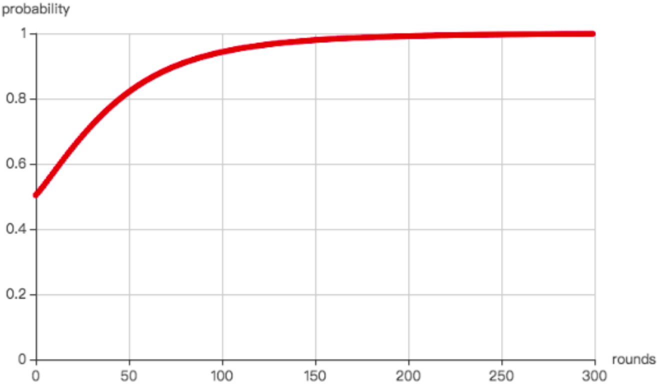
實驗二：mu∼U(0.95, 1), gamma∼U(0.01, 0.1)

Answer = mu



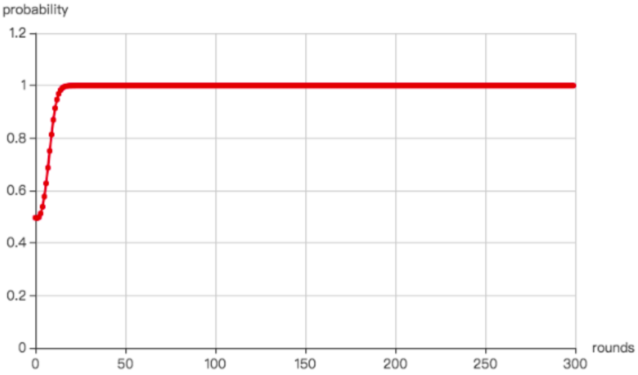
實驗二：mu∼U(0.95, 1), gamma∼U(0.01, 0.1)

Answer = gamma



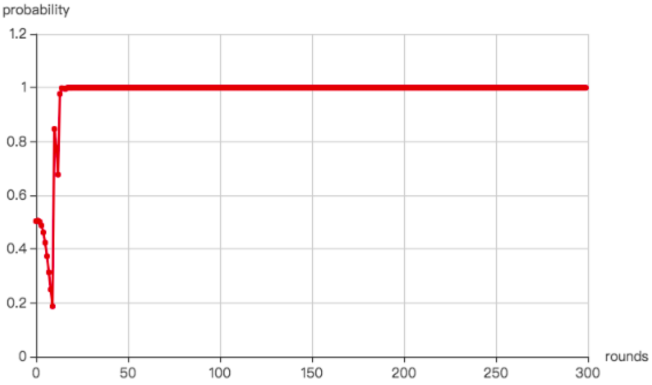
實驗三：mu∼U(0, 0.05), gamma∼U(10, 20)

Answer = mu

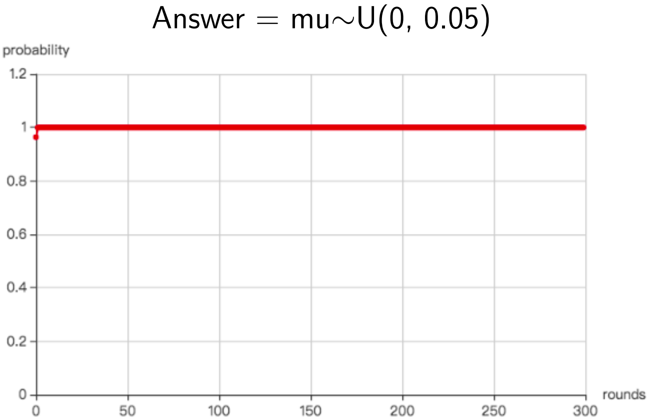


實驗三：mu∼U(0, 0.05), gamma∼U(10, 20)

Answer = gamma

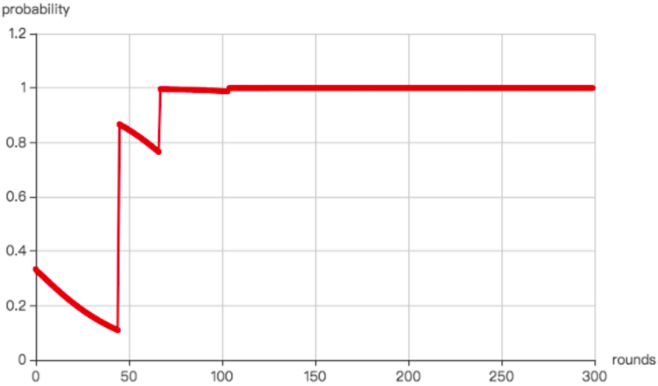


實驗四：極端組合


Answer = mu~U(0, 0.05)
probability
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
250
300
rounds

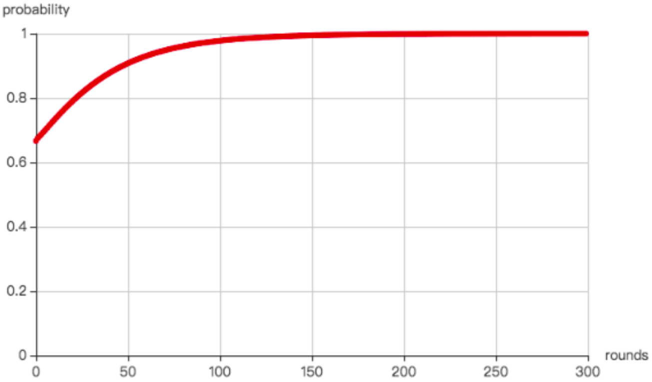
實驗四：極端組合

Answer $=$ mu$\sim$U(0.95, 1)



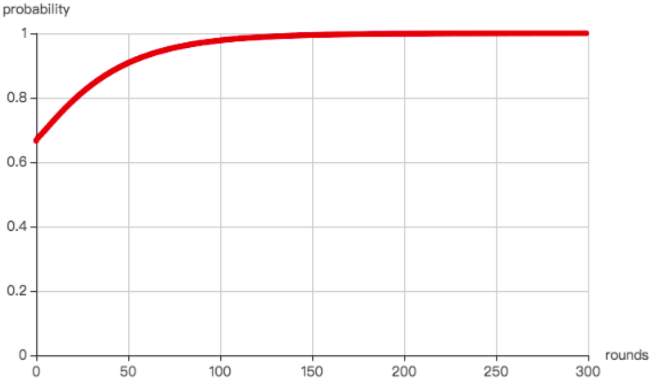
實驗四：極端組合

Answer = gamma∼U(10, 20)



實驗四：極端組合

$$\text{Answer} = \text{gamma} \sim U(0.01,\ 0.1)$$



結論

在只有Mu,Gamma這兩種agent的情境中，最簡潔的一種高分策略是：

- 1. 先全部採取合作 (stage 1)
- 2. 觀察對方是「不值得合作」群組的機率
  - 如果該機率大於某個信心水準α（例如0.99），從此全部採取背叛 (stage 2)

心靈小語

這告訴待人處事抱持著「不主動得罪人」是有益處的， 它讓我們可以找出值得信賴的朋友。（日久見人心）

謝謝大家！