Week Four – Prisoners Dilemma

This is like the snowdrift in some circumstances. Two individuals who can cooperate, C or defect; D

$$\begin{array}{c} \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} 3 & 0 \\ 5 & 1 \end{pmatrix} \end{array}$$

Clearly if both cooperate they get more than if they both defect, however if one defects, then he will do better & once that happens, the cooperator will obviously defect. Indeed, D is a strict Nash equilibrium. Clearly "Natural selection" favors defection since one defector will be more fit & can eventually invade & take over. In experimental game theory, people are more likely to cooperate than would be predicted by the PD. Rational players will always defect, so one of the big questions in theoretical & evolutionary biology is how does cooperation evolve into a population.

Consider the general matrix

R = reward, T = temptation,
P = punishment, S = sucker!

Prisoners Dilemma

$$\begin{array}{c} \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} \quad T > R > P > S,\ R > \frac{(T+P)}{2}$$

prevents alternation of CDCD from having bigger payoff than CCC

In reality, the game is not played once; rather it is played many times. So that if you defect, maybe you will be known as a jerk so that others will never cooperate. Consider the strategy GRIM → this cooperates on the first move & then cooperates as long as opponent doesn't defect. Once opponent defects, there is no forgiveness & you always D. Consider GRIM vs ALLD, & suppose you play exactly m rounds. Here is the payoff:

Suckered once & then always defect

$$\begin{array}{c} \\ \text{GRIM} \\ \text{ALLD} \end{array}\begin{array}{c} \begin{array}{cc} \text{GRIM} & \text{ALLD} \end{array} \\ \begin{pmatrix} mR & S+(m-1)P \\ T+(m-1)P & mP \end{pmatrix} \end{array}$$

Cash in on first move then always play against defector

If $mR > T + (m-1)P$, Then GRIM is a S.N.E. Note That ALLD is also S.N.E since $mP > S + (m-1)P$ as $P > S$. GRIM is S.N.E if $m > \frac{T-P}{R-P}$

for our model $m > 2$. This seems to be a way to stabilize C.

Since ALLD is also S.N.E., need a way to establish C in a world of D

Even worse - if we know There are <u>exactly</u> m rounds Then why not defect on The last round. Lets call This GRIM*

$$
\begin{array}{c|cc}
 & \text{GRIM} & \text{GRIM*} \\
\text{GRIM} & mR & (m-1)R + S \\
\text{GRIM*} & (m-1)R + T & (m-1)R + P
\end{array}
$$

suckered last round

both D last round

small gain last round

Since $T > R$, GRIM is NOT a N.E. & since $P > S$, GRIM* is a S.N.E. so will take over. But Then once every one defects on last round, why not defect on penultimate round, etc until once again ALLD!

Humans do not use This strategy. Perhaps because you never know when you might run into some one again

Lets do repeated P.D. with variable # rounds. After each round, prob w of another round. Expected # rounds is $\bar{m} = 1/(1-w)$ (Prove This! $P(1) = 1-w$

$P(2) = w(1-w)$, $P(3)$ $w^2(1-w)$, $P(n) = w^{n-1}(1-w)$

$\bar{m} = 1-w + 2(w(1-w)) + 3(w^2(1-w)) + \cdots = (1-w)[1 + 2w + 3w^2 + +) = (1-w) \cdot \frac{d}{dw}[w + w^2 + w^3 + \cdots] = (1-w)\left[\frac{d}{dw}\left[\frac{1}{1-w} - 1\right]\right)$

$[\underbrace{1 + w + w^2 + \cdots}_{\frac{1}{1-w}} - 1] = \frac{1}{1-w}$

Thus, we have:

$$\begin{array}{cc} \text{GRIM} & \text{ALLD} \end{array}$$
$$\begin{pmatrix} \bar{m}R & S+(\bar{m}-1)P \\ T+(\bar{m}-1)P & \bar{m}P \end{pmatrix}$$

+ GRIM is E.S. if $\bar{m} > \frac{T-P}{R-P}$, as before, but now can't just defect on last move. Still have ALLD as SNE as well. Is GRIM the best strategy? What if someone defects only once to see if he can get away with it? No reconciliation → ALLD, so maybe there is a better strategy.

Set of all strategies is infinite, naturally. Could use strategy which depends on history, eg

Each round has 4 possible outcomes: CC, CD, DC, DD so there are $2^4 = 16$ strategies that look at last round only. + 0000 → always defect 1000 C if CC else D, 1111 always C; 2^{16} strategies using last 2 moves, 2^{4^m} that use last m moves. Random strategies span 4^m-dim strategy space, $[0,1] \times \cdots \times [0,1]$

Axelrod set up a bunch of tournaments — people submitted strategies + all played each other multiple times (not knowing when they ended). Winner had highest total payoff. TFT won out of 14 entrants:

C on first move, then answer C for C and D for D.

In a second tournament 63 strategies & again TFT won. TFT never tries to get more than opponent, but overall matches TFT won! On average it does better against X than others do against X.

$$\begin{array}{c|cc} & TFT & ALLD \\ \hline TFT & \bar{m}R & S+(\bar{m}-1)P \\ ALLD & T+(\bar{m}-1)P & \bar{m}P \end{array} \qquad \bar{m} > \frac{T-P}{R-P}$$

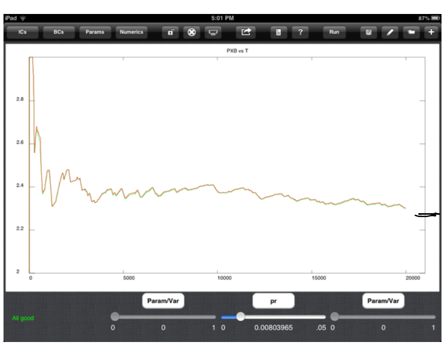
(Like GRIM against ALLD)

In all interactions, there are sometimes mistakes, so that maybe someone meant to push the "C" button but pushed the "D". TFT cannot correct mistakes:

C C C D* C D C D D D D D ← could flip back of course!
C C C C D C D D D D D D

In the end, two TFT* players with small mistakes have same payoff as players going randomly:

$$A(TFT,TFT)_{Noise} = \frac{R+T+P+S}{4}$$



since $R > \frac{T+S}{2}$ & $R > P$, Average $< R$

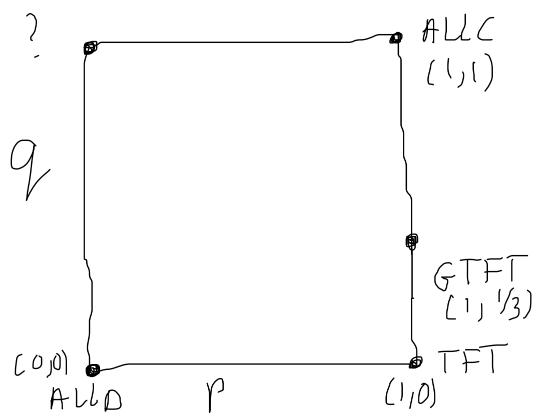
so TFT is weak in presence of noise

$2.25 = \frac{9}{4} = \langle PO \rangle$ TFT + Noise can be invaded by other strat.

Another weakness is that ALLC vs TFT is $\begin{pmatrix} \bar{m}R & \bar{m}R \\ \bar{m}R & \bar{m}R \end{pmatrix}$ neutrally stable so could drift to ALLC which is taken over by ALLD

Reactive Strategies

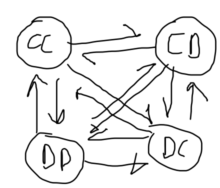
meters p, q. p = prob of cooperating if opponent C and q = prob of C if opponent D. Have short memories but can parametrize TFT $(1,0)$, ALLC $(1,1)$, ALLD $(0,0)$. $S(p,q)$ is a point in unit square. One last corner is the weird $(0,1)$ where C when opponent D + D when opponent C



Repeated PD is a Markov Chain on 4 states:

CC	CD	DC	DD
1	2	3	4

Two para



Let there be 2 strategies $S_1: (p_1, q_1)$ and $S_2: (p_2, q_2)$, so for example $CC \to CC$ if both cooperate on next move: $p_1 p_2$

$$M = \begin{array}{c|cccc} & CC & CD & DC & DD \\ \hline CC & p_1 p_2 & p_1(1-p_2) & (1-p_1)p_2 & (1-p_1)(1-p_2) \\ CD & q_1 p_2 & q_1(1-p_2) & (1-q_1)p_2 & (1-q_1)(1-p_2) \\ DC & p_1 q_2 & p_1(1-q_2) & (1-p_1)q_2 & (1-p_1)(1-q_2) \\ DD & q_1 q_2 & q_1(1-q_2) & (1-q_1)q_2 & (1-q_1)(1-q_2) \end{array}$$

m_{ij} = prob state $i \to$ state j

ASIDE: A stochastic matrix is a matrix with nonnegative entries whose row sums are 1. We define a SM, M as regular or irreducible if $\exists\, k > 0$ st all entries of M^k are positive.
This means if we draw $\to$ from node i to node j if $m_{ij} > 0$
then can get from $i \to k$ for all pairs (i, k) in finite steps.

Clearly $M\mathbb{1} = \mathbb{1}$ so 1 is an eigenvalue. Let $\vec{v}$ be the left eigenvector $\vec{v}_1$ eg $M^T\vec{v} = \vec{v}$. There is a Theorem (Perron/Frobenius) That says all entries of $\vec{v}$ are positive. Furthermore $\lambda = 1$ is the principle eigenvalue of M. If λ_j is e.v. then $|\lambda_j| < 1$ for $\lambda_j \neq 1$ + $\lambda = 1$ is simple if M is irreducible END ASIDE

Call $\lambda_1 = 1$ max e.v.

Let $\vec{X}_t = \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \\ x_{4t} \end{pmatrix}$ be prob of state (i) at iterate t

$\vec{X}_{t+1} = M^T \vec{X}_t$. Write $\vec{X}_t = c_1^t \vec{v}_1 + c_2^t \vec{v}_2 + \cdots + c_N^t \vec{v}_N$

$\vec{X}_{t+1} = M^T \vec{X}_t = c_1^t \lambda_1 \vec{v}_1 + \cdots + c_N^t \lambda_N \vec{v}_N$

$\Rightarrow c_j^{t+1} = \lambda_j c_j^t \Longrightarrow c_j^t = \lambda_j^t c_j^0 \Rightarrow c_j^t \to 0$ as $t \to \infty$

Unless $j=1$, so $\vec{X}^t \to \vec{v}_1$ as $t \to \infty$, so stationary density $\to \vec{v}_1$

We can compute $\vec{v}_1$! $r_1 = p_1 - q_1 \quad r_2 = p_2 - q_2$ regular if $|r_1 r_2| \in (0,1)$

$\vec{v}_1 = [s_1 s_2,\ s_1(1-s_2),\ (1-s_1)s_2,\ (1-s_1)(1-s_2)]$

$$s_1 = \frac{q_2 r_1 + q_1}{1 - r_1 r_2}, \quad s_2 = \frac{q_1 r_2 + q_2}{1 - r_1 r_2}$$

Clearly $E(S^1, S^2) = R s_1 s_2 + S s_1(1-s_2) + T(1-s_1)s_2 + P(1-s_1)(1-s_2)$

if $|r_1 r_2| = 1$ or 0, Then get deterministic strategies

EXPERIMENT Generate $n=$ 100 randomly chosen reactive strategies, e.g. choices of p, q. Compute The $n \times n$ payoff matrices using above (100×100) Assume $x_i(0) = \frac{1}{100}$, + do replicator equations on this. Watch the evolution.