Game Theory - Prisoner's Dilemma

Ambrose Law

March 17, 2017

1 Introduction

The Iterated Prisoner's Dilemma is a famous example of a finitely repeated normal form game, originally framed by Merrill Flood and Melvin Dresher in 1950 [1]. 2 players are given 2 strategies to choose from; Cooperate or Defect (C and D respectively). They play a finite amount of rounds and their end utilities are compared. In this case, the aim is to maximise a player's utility, and the game's utility bi-matrix looks like:

$$\begin{pmatrix} (3,3) & (0,5) \\ (5,0) & (1,1) \end{pmatrix}$$

where player 1 is the row player, and player 2 is the column player. Row and column 1 are strategy C, and row and column 2 are strategy D.

I will be running a tournament of different strategies. This will utilise the code provided by a community on GITHUB running a Tournament with submitted strategy profiles, similar to what Robert Axelrod did in the 1980's.

2 Game Theory

To look at how well a strategy performs, we first look at the theory that was considered when the profile was designed.

2.1 Nash Equilibria

In a normal form game, such as the Prisonor's Dilemma, we can use rationality to see what strategy we should use. If we look at the best responses for each player, we can deduce that there is no reason to move from playing D. This is our Nash Equilibrium for pure and mixed strategies.

Also, by the Theorem of a sequence of stage Nash profiles, we can see that the Nash Equilibrium is subgame perfect. So, at any period, our stage game's Nash Equilibrium is D.

However, it might be beneficial to not play D at every stage game, and this is because of 'Reputation'. If my opponent and I chose C every round, our payoff at the end would be higher than if we played D each time. Then again, my payoff would be maximised if my opponent played C every turn and I played D every turn. This isn't likely to happen, so is there a combination of C and D which maximises our payoff? From here, our strategies are now sentences that dictate what we do at each stage game to resemble this combination, e.g. "Alternate between C and D".

3 The Code

Python has a package called 'Axelrod'. It has all the required code to run a tournament between different strategies. To decide if, due to reputation, the Nash equilibrium is the best strategy, I will put the 'Defector' strategy (Always plays D) along with 3 other strategies, into a tournament. The other strategies shall be: Random, that plays a mixed strategy of $(0.5, 0.5)$, Tit For Tat (TFT) which starts with C and then copies the opponment's last move, and Slow Tit For Two Tat (STFTT) which plays C twice and then copies the opponent's last move if it was played twice. Running the code:

```
players = [axl.SlowTitForTwoTats(), axl.Defector(), axl.Random(), axl.TitForTat()]
tournament = axl.Tournament(players)
```

set up the tournaments with the named strategies. The results were:

```
['Tit For Tat', 'Random: 0.5', 'Slow Tit For Two Tats', 'Defector']
```

So we do indeed see that, in this case, Defector was not the best strategy. In fact, looking at the scores and result in Figure 1, Defector performs very poorly in comparison. The 2 Tit for Tat strategies punish Defector by playing D against it, scoring only 1. But, against each other and Random, they can consistently score more than 1.

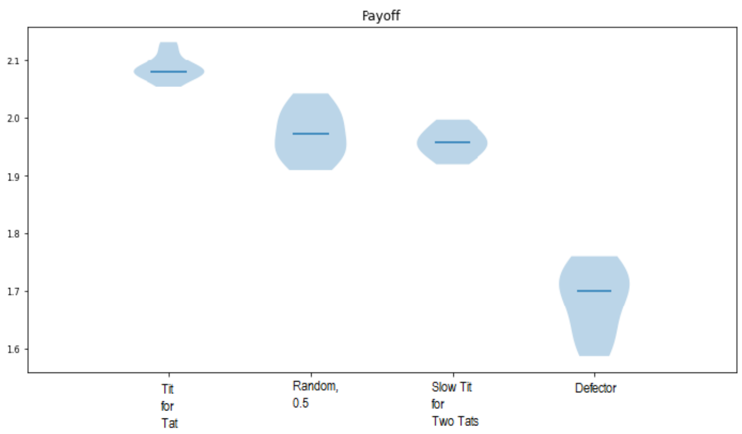

Figure 1: Score of Tournament

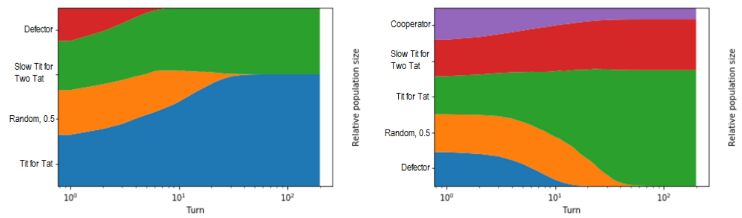

Figure 2: Strategy Population Evolution over time

4 Evolutionary Stable Strategies

To further this point, we can consider the evolution of strategies in a population. This means that, given a infinite population, where each 'person' is a strategy, does one strategy 'over take' the others? We would say a strategy is an Evolutionary Stable Strategy (ESS) when; in a population consisting of the one strategy, given some amount of the population start to play a different strategy, the payoffs of the original strategy are greater than the payoffs of the newly introduced strategy.

By running the code

```
eco = axl.Ecosystem(results)
eco.reproduce(200)
```

we put all of our strategies into a population, and then get them to play stage games for 200 rounds. When a strategy loses, they then take up the other strategy.

By looking at Figure 2: Left, it is clear that TFT is the best strategy in this population. So TFT is an ESS for this population. Furthermore, Defector is 'phased out' fairly fast, which further defends our point that the Nash Equilibrium isn't necessarily the best strategy in an iterated game.

Though, if we added the strategy 'Cooperator', who only plays C, and run the tournament again, the results are:

```
['Defector', 'Random: 0.5', 'Tit For Tat', 'Slow Tit For Two Tats', 'Cooperator'].
```

So, in this tournament, Defector, and so the Nash Equilibrium, is the winning strategy. However, running the ecosystem again, but now involving Cooperator, we get Figure 2: Right. Again, Defector is out played, causing it to be phased out.

5 Conclusion

We have seen that in some iterative games, it is not always the best strategy to play the Nash Equilibria, and is, instead, better to consider Reputation. For the Prisoner's Dilemma, one's reputation means that other players are more likely to cooperate with you, as long as, in return, you do the same. In an infinite population, cooperating leads to those strategies being copied and over take the population. However, this all can depend on your strategy profiles and space. There could be a set of strategies which has a corresponding population where the Defector can come out on top.

References

[1] Wikipedia. Prisoner's dilemma wikipedia.