FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Master's In Informatics Engineering 2017/2018
Complex Systems

Iterated Prisoner's Dilemma

23.05.2018
Coimbra - Portugal

João André Taborda Barata Portugal Vieira | 2013136370

TABLE OF CONTENTS

I. Introduction

Computer Simulations are a very useful way of understanding and studying the long term behaviour of Complex Systems. In this project, we'll be looking at how it's possible to simulate the problem of the Iterated Prisoner's Dilemma. We will look at previous approaches to this problem, what new ideas we can implement, and how their results compare to the current state of the art.

II. Problem Description

The problem we will be looking at is the Iterated Prisoner's Dilemma. In order to understand it, we must first take a look at the Prison's Dilemma itself:

Two gang thieves are put in separate cells, unable to communicate. They are then given a proposal. They are given the possibility of betraying their fellow gang member by testifying against them. However, the other member is also given the same possibility, and the number of years in prison each one gets depends on the other member's choice. The following table shows the possible values for number of years removed from their sentence depending on the thieves' behaviour:

	Cooperate	Defect
Cooperate	3, 3	0, 5
Defect	5, 0	1, 1

Figure 1 – Prisoner's Dilemma Table [1]

The values in this table can vary, as long as a similar ratio is maintained. We will be using the values in the previous table for this project.

What's important to note about this dilemma is that betraying a partner offers a larger reward than coordinating. However, what makes the question difficult for the thieves is the fact that, if both thieves defect, they both get a smaller payoff than by both coordinating. When people are presented with this choice, they tend to cooperate more often than not, despite it not being necessarily the optimal choice from a rational standpoint.

In order to make the problem more complex, an iterated version of this dilemma was created, which is the one we will be focusing on. Instead of making just one decision, the problem is repeated multiple time, creating a round-based game where each thief is a player who can choose a different strategy based on the opponent's behaviour. The number of rounds for this problem can vary, and many strategies have been developed and compared in tournaments in order to find the most consistent one.

III. Previous Work

In the tournaments that were made to test strategies for the Iterated Prisoner's Dilemma, the "Tit-For-Tat" strategy tended to be the best. This strategy consists of starting by cooperating, and then imitating the opponent's behaviour in the last round. This strategy rewards the opponent for cooperating and punishes them for defecting, creating an incentive for mutual cooperation. Since mutual cooperation gives the most consistent results, this strategy tends to perform best, even if it doesn't get the most points in every match.

Generally strategies work by either reacting to the opponent's last move, or by keeping a series of recent moves in memory. It's important to note that it's been proven that increasing the size of each player's memory does not improve results, however. [2]

Overall, altruistic strategies tend to perform better than selfish strategies. [3] Altruistic strategies refer to strategies that have a tendency to cooperate rather than defect. However, you can't always cooperate, as you will start earning 0 points if your opponent always defects.

IV. Problem Setup

As was mentioned in the introduction to this report, we will be performing a Computer Simulation to analyse this problem and draw conclusions. The setup will be separated into two parts. In the first part, we will simulate different possible strategies and see how they compare when competing with each other in a 2D grid. In the second one, we will try an alternative strategy where players start by cooperating and then defect when they reach a specific k-value turn or their neighbours defect.

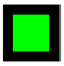
In both cases, we simulate a 20x20 grid over 100 turns, and then repeat the experiment 10 times. Each player's behaviour, when its strategy is dependent on other players, will act based on the majority of its 8 neighbours. For example, in the Tit For Tat strategy, each player will perform the same action as most of its neighbours performed last turn.

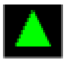
The experiments were designed to provide readability while also conveying the most important information.

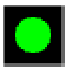
Players are represented by polygons of different shapes. **If they are green, they are cooperating. If they are red, they are defecting.**

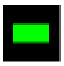
V.A. Part 1 – Simulation Setup

Four strategies were selected for this simulation. They are listed below:

 - Tit For Tat – Repeats the same behaviour its neighbours chose last turn.

 - Pavlov – Repeats last turn's behaviour if it got more than 2 points on average. Otherwise, does the opposite.

 - Random – Randomly cooperates or defects.

 - Reverse Tit For Tat – Does the opposite behaviour as most of its neighbours.

The players with random behaviour are important to serve as a control group. If a strategy underperforms the random choice strategy, it can't be considered effective.

Pavlov strategy's purpose is to try to maximize short term success, by repeating choices that worked and changing unsuccessful choices.

At the end, the simulator prints the total number of points obtained by each player, along with the strategy that it was following. This allows us to draw the statistics that we will use to write conclusions about the system.

V.B. Part 1 – Simulation Results

First, we will take a look at possible initial and final configurations:

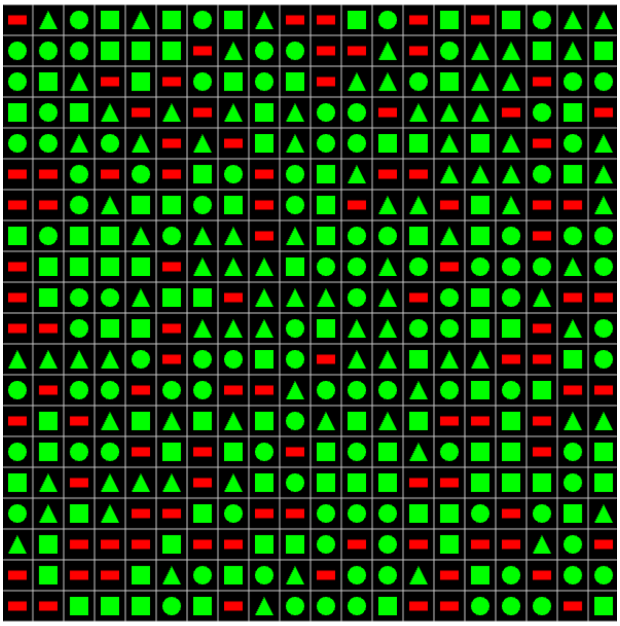


Figure 2 – Initial Configuration for Part 1

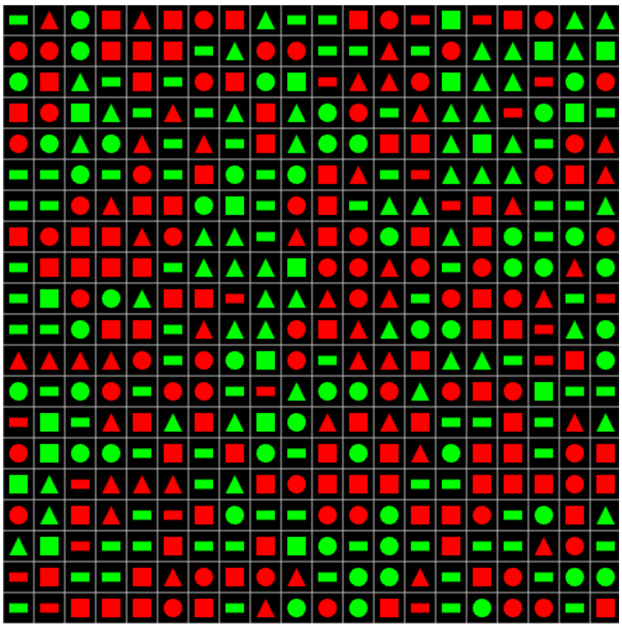


Figure 3 – Final Configuration for Part 1

The experiment was repeated 10 times. On average, these were the results obtained for each type:

	Type				
	Tit For Tat	Pavlov	Random	Reverse Tit For Tat	Grand Total
Average of Points	1107.433735	937.1375	925.4524	735.3376623	929.7746914
	1134.763158	960.9241	962.6528	749.7113402	938.8518519
	1116.039474	932.5455	950.4737	724.2528736	924.9135802
	1110.026316	894.7955	936.0127	760.2716049	921.7006173
	1113.379747	912.9412	916.5904	786.2727273	932.6450617
	1082.593407	890.9242	879.4608	687.7076923	900.3796296
	1131.478873	969.2135	959.8889	742.6413043	938.3641975
	1083.295918	891.1707	924.5467	704.0144928	917.1512346
	1142.30137	955.5417	991.6667	791.4875	965.4753086
	1110.964706	914.8481	916.0976	739.3717949	924.3703704

Figure 4 – Part 1 Results

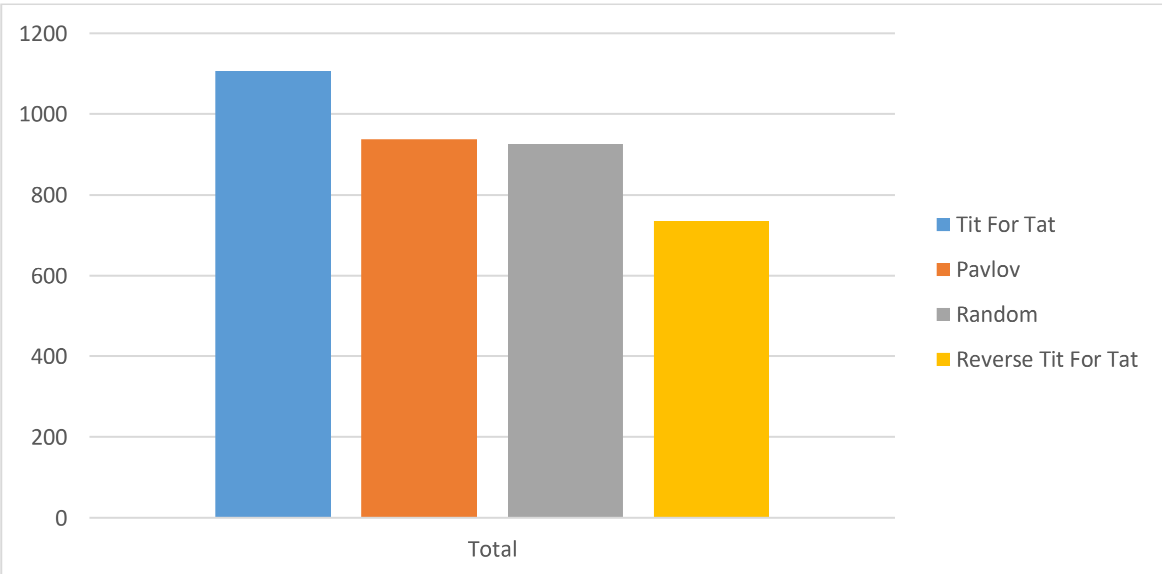


Figure 5 – Part 1 Result Histogram

As expected, the Tit For Tat strategy performed best of all of them. Reverse Tit For Tat was outperformed by the Random strategy, which shows it's not a very consistent strategy, even if it won some rounds. Overall, the results were not too surprising, and seem to corroborate previous research.

VI.A. PART 2 – SIMULATION SETUP

For this part, the strategy employed consists of having every agent cooperate, until they reach a certain k-value turn, when they defect. This k-value is randomly chosen for each agent at the beginning. If most of the agent's neighbours defect, then the agent also defects, and then keeps defecting until the end.

We will compare this strategy to the previous ones, and also try to identify the best possible value for k.

VI.B. PART 2 – SIMULATION RESULTS

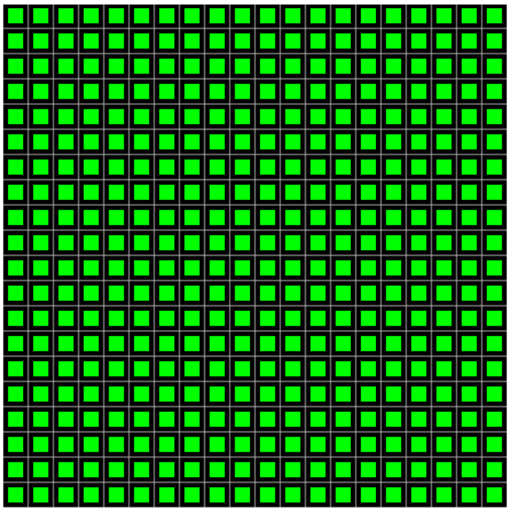


Figure 6 – Initial Configuration for Part 2

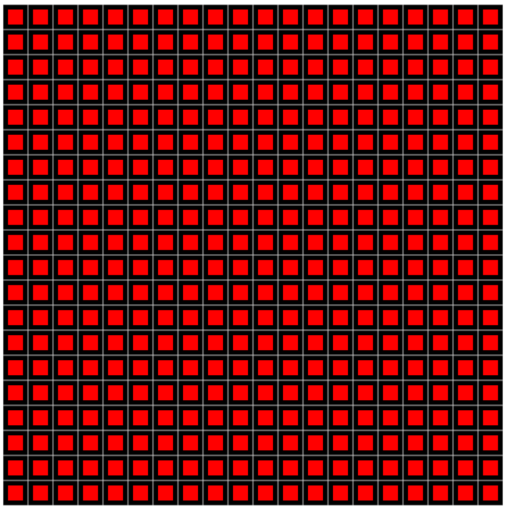


Figure 7 – Final Configuration for Part 2

This experiment was repeated 10 times as well. The following histogram shows the average number of points for each value for k:

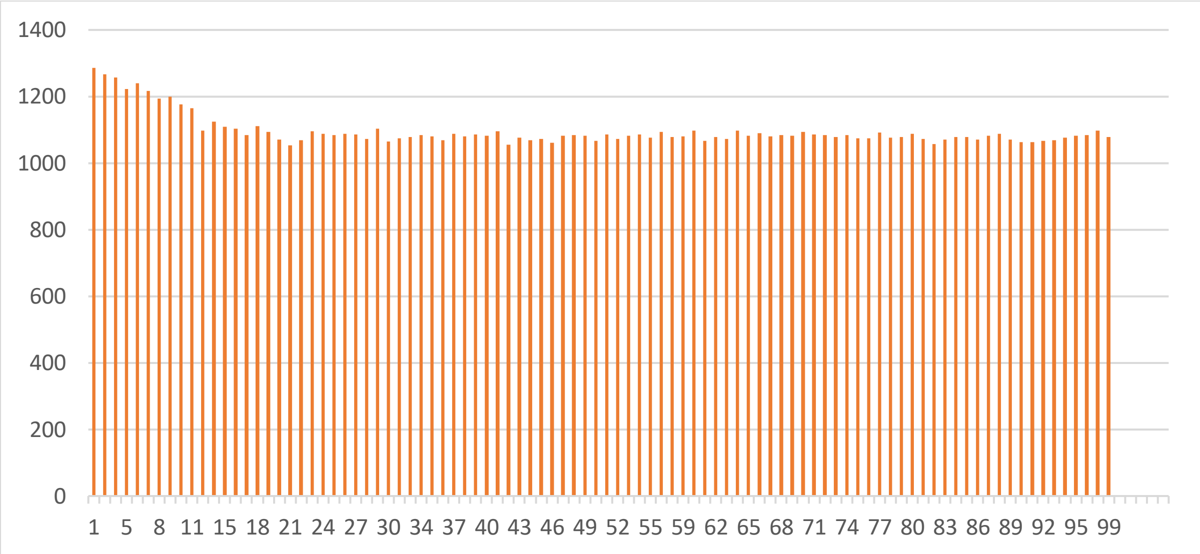


Figure 8 – Part 2 Result Histogram

While most values showed similar results, using 1 as the value for k was clearly the best option. This option allows the player to "steal" 5 points multiple times from the players who are cooperating at the start. Since every player in the board will eventually start defecting, it wouldn't make sense to ever go back to cooperating. Thus, in this board, defecting early is optimal.

VII. Comparison

Since the two parts of this project use different grids, it would be unfair to compare them directly. Thus, I also made a short experiment replacing the Pavlov strategy with the strategy from Part 2.

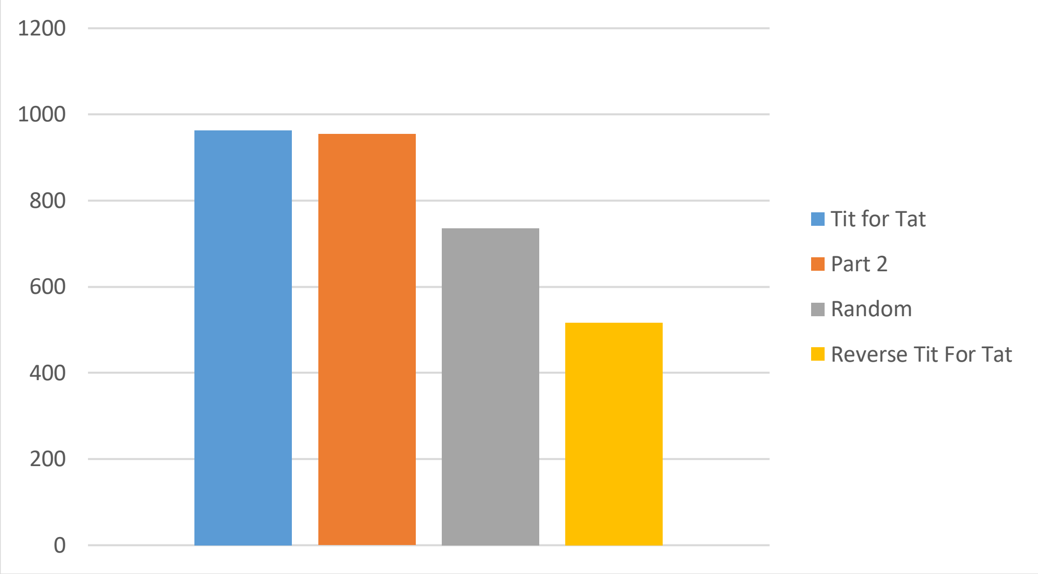


Figure 9 – Part 1 Experiment with k-value

While the average results for this strategy were impressive, they were still slightly behind the Tit For Tat strategy.

VIII. Conclusion

In the future, it would be interesting to complete this project with more strategies. It would be possible to fill the grid with dozens of different strategies and analyse how they interact, and which ones get better results.

With this practical project, we were able to create a useful tool to simulate the Iterated Prisoner's Dilemma. With it, we could obtain a lot of useful information about how different strategies interact, and experiment with a new strategy, able to achieve very competitive results.

IX. Bibliography

[1] http://lizengland.com/masters/iteratedpd.htm

[2] http://www.pnas.org/content/109/26/10409

[3] https://www.psychestudy.com/behavioral/learning-memory/iterated-prisoners-dilemma