

# Research of Parameters in Strategies of the Iterated Continuous Prisoner's Dilemma

Adrian Hossner

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Theoretical Foundations</b>	<b>4</b>
<b>3</b>	<b>Methods and Implementation</b>	<b>7</b>
<b>4</b>	<b>Results</b>	<b>8</b>
<b>5</b>	<b>Discussion</b>	<b>8</b>
<b>6</b>	<b>Conclusions and Outlook</b>	<b>8</b>
<b>7</b>	<b>Self Reflection</b>	<b>8</b>
<b>8</b>	<b>Annex</b>	<b>9</b>

# Abstract

## 1 Introduction

- Interactions  
implication range is wide:  
from two persons in a relationship to countries in alliance  
from two impala's allogrooming to zebras warning the herd about a lion
- Key hook:  
helping someone with possible cost – cooperation  
being selfish but breaking trust – defection
- model PD:  
single interaction  
already well analysed  
Nash found solution/equilibrium
- Iterated:  
more interesting, insightful  
strategies
- non-suited variants:  
Axelrod's Tournament  
Evolution
- continuous:  
more complex, accurate
- parameters:  
determines behaviour  
create surface  
the part that has not been experimented
- overall and difference:  
overall: wealth of population  
difference: competence between individuals

## 2 Theoretical Foundations

- Game Theory (optional, not needed for understanding)  
mathematical framework to investigate interactions
- Prisoner's Dilemma:

The Prisoner's Dilemma (PD) is a famous intellectual game in Game Theory. It was invented by Merrill Flood and Melvin Dresher in 1950. In order to understand the dilemma, a proper context has to be established. Two persons are arrested and are sued for having stolen something. The police, however, has not enough evidence to imprison them. As a consequence of that, the police interrogates the two criminals. They can either say nothing or confess that the other suspect was involved in the crime. They can either cooperate or defect without having the possibility to talk to each other before the decision. A certain tactic is being used by the police officers. They make it profitable for the criminals to confess. They let the criminal out of prison earlier if he/she confesses. But if the other also confesses, both stay long in prison. If he/she says nothing and the other confesses, he/she will stay very long in prison. If, however, both stay silent, both of them go to prison in a relatively short amount of time. The following matrix visualises the pay-off's in this game in a comprehensive way:

	C	D
C	3, 3	10, 0
D	0, 10	5, 5

C stands for cooperation and D stands for defection. Cooperation is defined by remaining silent whereas defection means confessing. The dilemma consists of the following. Confessing seems attractive since the interrogated criminal can walk out freely without going to prison. However, if the second criminal also confesses, both get five years in prison. This is, nevertheless, the worst outcome for both which could have been avoided as both could have stayed silent.

John Nash, a mathematician which made great contributions in Game Theory, has proved that it is the most logical option for both to confess, always. A Nash-equilibrium is introduced. It defines a stable state in which both would not change their decision even though they would know if the other cooperated or defected.

This game composes a one-time interaction very well in a mathematical and analytical perspective. The pay-off's in the matrix can be changed as long as these rules stay valid.

	C	D
C	R, R	S, T
D	T, S	P, P

$$T > R > P > S$$

$$2 * R > T + S$$

- Iterated:

The Iterated Prisoner's Dilemma (IPD) is a game which extends the PD. As the name suggests, the PD is played a number of times sequentially. After each round the pay-off gets accumulated to the points one strategy has already gained. Since a strategy wants to gain as many points as possible, the pay-off matrix should be altered accordingly.

	C	D
C	3, 3	0, 5
D	5, 0	1, 1

The Nash-equilibrium in the IPD differs from that in the PD. This means that cooperation can earn out long-term. By cooperating, one is building trust with another. One strategy can only get the maximum points, which can be gained by always getting the 5 points each round, by defecting and the opponent's cooperation. Since the opponent will likely not cooperate after one strategy always defects, the maximum is practically impossible to receive. So, the relative high amount and the possibility of getting it for both is by cooperating.

The number of round that will be played has to be unknown to both strategies. There is no sense in cooperating the last round since trust does not matter after the game is ended. And because trust will be given up in the last round, there is no sense for both to keep trust the second last round. This will end up in an inductive defection behaviour for both strategies the whole game.

This extension of the PD allows us to simulate a relationship over time where the two players interact multiple times.

- Strategies:

A strategy is to be considered as an autonomous player which has its own behaviour. A strategy has the whole history of its own contribution and the ones of the opponent at hand to generate the decision for the next round. These decisions are calculated by using conditions and probabilities on the given data.

- Continuous:

The continuous variant is considered more accurate to the real world since the decisions are rather rare only either cooperation or defection. The pay-off system needs to be altered accordingly. The pay-off of [sciencedirect.com] is suitable for this project.

$$p1 = x - c * y$$

$$p2 = y - c * x$$

The amount of cooperation is called investment and can range from zero to one (0 to 1) where full defection is represented by 0 and cooperation corresponds to 1. The variable x is the investment of strategy 1 and y is the investment of strategy 2. The parameter c also ranges from 0 to 1 and describes the cost of cooperation.

- Noise:

Noise in the ICPD is the equivalent to miscommunication or misunderstanding in the real world. In this variant of the PD, noise is essential to trigger interesting outcomes. When a game of the ICPD is started and two strategies who start with a full investment, many strategies only respond with full investments. So noise is necessary to trigger continuous investments. Without noise, in many cases, the ICPD would be equivalent to the IPD.

- Parameter-based:

The main idea of this paper is to use parameter-based strategies. These strategies hold a parameter that defines the behaviour of the strategy. This parameter can determine the investment of one strategy in many ways. [Example]

- Surfaces:

The end result of this paper will be several surface plots. The data will be generated by letting two parameter-based strategies play against each other the ICPD. The game will be structured so that every parameter came against every other parameter in the ICPD. Like this a surface can be plotted by having the x-axis being the parameter of strategy 1 and the y-axis being the parameter of strategy 2. The z-axis will indicate the points one strategy gained. That being said, one surface will be shown for each strategy. Further more, the two z-axis of the two surfaces can be added and form a new surface which shows the overall points gained by

both strategies. The complement to that would be to subtract the two z-axis to plot a surface which describes how much better one strategy was than the other.

- Simultaneous vs Alternating:

Two main differences can be seen when simulating an iterated variant of the PD. On one hand there is the simultaneous form whereby the strategies submit their contribution at the same time. On the other hand, there is the alternating form in which one strategy submits its contribution and the other strategy can respond to this contribution in that round. Since the alternating form gives an advantage to the strategy which is allowed to respond, this only makes things more complicated than necessary. So, this analysis is only dedicated to the simultaneous form.

- Current Findings:

In the simpler variants, e.g. the IPD without noise, it has been proved many times over that the strategy tit-for-tat is the most successful one. [sources] The success of strategies, however, can vary depending on certain conditions such as the influence of noise or different pay-off systems where continuous investments can be submitted. So, the very specific area of the simultaneous ICPD, has not been very well explored.

### 3 Methods and Implementation

- Programming language:

The simulation, data generation and visualisation is completely written in Python (3).

- Data Generation:

two strategies  
every parameter against every parameter  
play several times to smooth out noise (average)

- Implemented Variants:

Prisoner's Dilemma  
Iterated Prisoner's Dilemma  
Iterated Continuous Prisoner's Dilemma  
Parameter-Based Strategies  
only PBS important for MA

- Simulated Strategies and explanation of parameter  
Average  
RandomContinuous  
RandomDiscrete  
Adapt  
...

## 4 Results

- Surfaces:  
plots  
surface of strategy 1  
surface of strategy 2  
surface of both (added)  
surface showing difference (subtracted)
- Analysis:  
hills and valleys

## 5 Discussion

- ?

## 6 Conclusions and Outlook

- Summary and Quintessence  
which strategy 1 better than which strategy 2  
which parameter to be used
- Application in the Real World  
strategies translated to real persons  
giving examples (Politics, economics to simple human beings)  
prospect to the future (which ones will win?)

## 7 Self Reflection

- Challenges  
estimation of time consumption  
transforming ideas during process  
dedication



- Learnings
  - reading professional papers
  - writing English
  - experimental workflow (usage of neovim, tmux and terminal)

## 8 Annex