Understanding responses to environments for the Prisoner's Dilemma; A machine learning approach

Nikoleta E. Glynatsi

Month 2020

Submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy.

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

School of Mathematics
Ysgol Mathemateg

Executive Summary

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetuer a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

Acknowledgements

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Contents

List of Figures

List of Tables

Summary

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Chapter 1

Introduction

Chapter 2

A systematic literature review of the Prisoner's Dilemma.

The Prisoner's Dilemma is a well known game used since the 1950's as a framework for studying the emergence of cooperation; a topic of continuing interest for mathematical, social, biological and ecological sciences. The iterated version of the game, the Iterated Prisoner's Dilemma, attracted attention in the 1980's after the publication of the "The Evolution of Cooperation" and has been a topic of pioneering research ever since. The aim of this paper is to provide a systematic literature review on Prisoner's Dilemma related research. This is achieved by reviewing selected pieces of work and partition the literature into five different sections with each reviewing a different aspect of research. The questions answered in this manuscript are (1) what are the research trends in the field (2) what are the already existing results within the field.

2.1 Introduction

Based on the Darwinian principle of survival of the fittest cooperative behaviour should not be favoured, however, cooperation is plentiful in nature. A paradigm of understanding the emergence of these behaviours is a particular two player non-cooperative game called the Prisoner's Dilemma (PD), originally described in **Flood1958**.

In the PD each player has two choices, to either be selfless and cooperate or to be selfish and defect. Each decision is made simultaneously and independently. The utility of each player is influenced by its own behaviour, and the behaviour of the opponent. Both players do better if they choose to cooperate than if both choose to defect. However, a player has the temptation to deviate as that player will receive a higher payoff than that of mutual cooperation. Players' payoffs are generally represented by (2.1). Both players receive a reward for mutual cooperation, R, and a payoff P for mutual defection. A player that defects while the other cooperates receives a payoff of T, whereas the cooperator receives S. The dilemma exists due to constraints (2.2) and (2.3).

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} \tag{2.1}$$

$$T > R > P > S \tag{2.2}$$

$$2R > T + S \tag{2.3}$$

Another common representation of the payoff matrix is given by (2.4), where b is the benefit of the altruistic behaviour and c it's its cost (constraints (2.2) and (2.3) still hold).

$$\begin{pmatrix} b-c & c \\ b & 0 \end{pmatrix} \tag{2.4}$$

Constraints (2.2-2.3) guarantee that it never benefits a player to cooperate, indeed mutual defection is a Nash equilibrium. However, when the game is studied in a manner where prior outcome matters, defecting is no longer necessarily the dominant choice.

The repeated form of the game is called the Iterated Prisoner's Dilemma (IPD) and theoretical works have shown that cooperation can emerge once players interact repeatedly. Arguably, the most important of these works is Robert Axelrod's "The Evolution of Cooperation" **Axelrod1984**. In his book Axelrod reports on a series of computer tournaments he organised. In these tournaments academics from several fields were invited to design computer strategies to compete. Axelrod's work showed that greedy strategies did very poorly in the long run whereas altruistic strategies did better. "The Evolution of Cooperation" is considered a milestone in the field but it is not the only one. On the contrary, the PD has attracted attention ever since the game's origins.

This manuscript presents a qualitative description of selected pieces of work. These have been separated into five sections, each reviewing a different aspect of research. The topics reviewed at each section are the following:

- section 2.2, **Origins of the Prisoner's Dilemma**.
- section 2.3, **Axelrod's tournaments and intelligent design of strategies**.

- section 2.4, **Evolutionary dynamics**
- section 2.5, **Structured strategies and training**.
- section 2.6, **Software**.

The aim of this work is to provide a concrete summary of the existing literature on the PD. This is done to provide a review which will allow the research community to understand overall trends in the field, and already existing results.

2.2 Origins of the prisoner's dilemma

The origin of the PD goes back to the 1950s in early experiments conducted at RAND **Flood1958** to test the applicability of games described in **VonNeumann1944**. The game received its name later the same year. According to **Tucker1983**, Albert W. Tucker (the PhD supervisor of John Nash **Nash1951**), in an attempt to deliver the game with a story during a talk described the players as prisoners and the game has been known as the Prisoner's Dilemma ever since.

The early research on the IPD was limited. The only source of experimental results was through human subject research where pairs of participants simulated plays of the game. Human subject research had disadvantages. Humans could behave randomly and in several experiments both the size and the background of the individuals were different, thus comparing results of two or more studies became difficult.

The main aim of these early research experiments was to understand how conditions such as the gender of the participants **Evans1966**; **Lutzker1961**; **Mack1971**, the physical distance between the participants **Sensenig1972**, the effect of their opening moves **Tedeschi1968** and even how the experimenter, by varying the tone of their voice and facial expressions **Gallo1968**, could influence the outcomes and subsequently the emergence of cooperation. An early figure that sought out to understand several of these conditions was the mathematical psychologist Anatol Rapoport. The results of his work are summarised in **rapoport1965**.

Rapoport was also interested in conceptualising strategies that could promote international cooperation. Decades later he would submit the winning strategy (Tit for Tat) of the first computer tournament, run by Robert Axelrod. In the next section these tournaments, and several strategies that were designed by researchers, such as Rapoport, are introduced.

2.3 Axelrod's tournaments and intelligently designed strategies

As discussed in Section 2.2, before 1980 a great deal of research was done in the field, however, as described in **Axelrod2012**, the political scientist Robert Axelrod believed that there was no clear answer to the question of how to avoid conflict, or even how an individual should play the game. Combining his interest in artificial intelligence and political science Axelrod created a framework for exploring these questions using computer tournaments. Axelrod asked researchers to design a strategy with the purpose of wining an IPD tournament. This section covers Axelrod's original tournaments as well as research that introduced new intelligently designed strategies.

Axelrod's tournaments made the study of cooperation of critical interest. As described in **Rapoport2015**, "Axelrod's "new approach" has been extremely successful and immensely influential in casting light on the conflict between an individual and the collective rationality reflected in the choices of a population whose members are unknown and its size unspecified, thereby opening a new avenue of research". In a collaboration with a colleague, Douglas Dion, Axelrod in **Axelrod1988** summarized a number of works that were immediately inspired from the "Evolution of Cooperation", and **Jurisic2012** gives a review of tournaments that have been conducted since the originals.

The first reported computer tournament took place in 1980 **Axelrod1980a**. A total of 13 strategies were submitted, written in the programming languages Fortran or Basic. Each competed in a 200 turn match against all 12 opponents, itself and a player that played randomly (called **Random**). This type of tournament is referred to as a round robin. The tournament was repeated 5 times to get a more stable estimate of the scores for each pair of play. Each participant knew the exact number of turns and had access to the full history of each match. Furthermore, Axelrod performed a preliminary tournament and the results were known to the participants. This preliminary tournament is mentioned in **Axelrod1980a** but no details were given. The payoff values used for equation (2.1) were $R = 3, P = 1, T = 5$ and $S = 0$. These values are commonly used in the literature and unless specified will be the values used in the rest of the works described here.

The winner of the tournament was determined by the total average score and not by the number of matches won. The strategy that was announced the winner was the strategy submitted by Rapoport, **Tit For Tat**. The success of Tit for Tat came as a surprise. It was not only the simplest submitted strategy, it would always cooperates on the first round and then mimic the opponent's previous move, but it had also won the tournament even though it could never beat any player it was interacting with.

In order to further test the results Axelrod performed a second tournament in 1980 **Axelrod1980b**.

The second tournament received much more attention and had a total of 62 entries. The participants knew the results of the previous tournament and the rules were similar with only a few alterations. The tournament was repeated 5 times and the length of each match was not known to the participants. Axelrod intended to use a fixed probability (refereed to as 'shadow of the future' **Axelrod1988**) of the game ending on the next move. However, 5 different number of turns were selected for each match 63, 77, 151, 308 and 401, such that the average length would be around 200 turns.

Nine of the original participants competed again in the second tournament. Two strategies that remained the same were Tit For Tat and **Grudger**. Grudger is a strategy that will cooperate as long as the opponent does not defect, submitted by James W. Friedman. The name Grudger was give to the strategy in **Li2014**, though the strategy goes by many names in the literature such as, Spite **Beaufils1997**, Grim Trigger **Banks1990** and Grim **Van2015**. New entries in the second tournament included **Tit for Two Tats** submitted by John Maynard Smith and **KPavlovC**. KPavlovC, is also known as Simpleton **rapoport1965**, introduced by Rapoport or just Pavlov **Nowak1993**. The strategy is based on the fundamental behavioural mechanism win-stay, lose-shift. Pavlov is heavily studied in the literature and similarly to Tit for Tat it is used in tournaments today and has had many variants trying to build upon it's success, for example **PavlovD** and **Adaptive Pavlov Li2007**.

Despite the larger size of the second tournament none of the new entries managed to outperform the simpler designed strategy. The winner was once again Tit for Tat. Axelrod deduced the following guidelines for a strategy to perform well:

- The strategy would start of by cooperating.
- It would forgive it's opponent after a defection.
- It would always be provoked by a defection no matter the history.
- It was simple.

The success of Tit for Tat, however, was not unquestionable. Several papers showed that stochastic uncertainties severely undercut the effectiveness of reciprocating strategies and such stochastic uncertainties have to be expected in real life situations **Milinski1987**. For example, in **Molander1985** it is proven that in an environment where **noise** (a probability that a player's move will be flipped) is introduced two strategies playing Tit for Tat receive the same average payoff as two Random players. Hammerstein, pointed out that if by mistake, one of two Tit for Tat players makes a wrong move, this locks the two opponents into a hopeless sequence of alternating defections and cooperations **Hammerstein1984**.

The poor performance of the strategy in noisy environments was also demonstrated in tournaments. In **Bendor1991**; **Donninger1986** round robin tournaments with noise were performed, and Tit For Tat did not win. The authors concluded that to overcome the noise more generous strategies than Tit For Tat were needed. They introduced the strategies **Nice and Forgiving** and **OmegaTFT** respectively.

A second type of stochastic uncertainty is misperception, where a player's action is made correctly but it is recorded incorrectly by the opponent. In **Wu1995**, a strategy called **Contrite Tit for Tat** was introduced that was more successful than Tit for Tat in such environments. The difference between the strategies was that Contrite Tit for Tat was not so fast to retaliate against a defection.

Several works extended the reciprocity based approach which has led to new strategies. For example Gradual **Beaufils1997** which was constructed to have the same qualities as those of Tit for Tat except one, **Gradual** had a memory of the game since the beginning of it. Gradual recorded the number of defections by the opponent and punished them with a growing number of defections. It would then enter a calming state in which it would cooperates for two rounds. In a tournament of 12 strategies, including both Tit for Tat and Pavlov, Gradual managed to outperformed them all. A strategy with the same intuition as Gradual is **Adaptive Tit for Tat tzafestas-2000a**. Adaptive Tit for Tat does not keep a permanent count of past defections, it maintains a continually updated estimate of the opponent's behaviour, and uses this estimate to condition its future actions. In the exact same tournament as in **Beaufils1997** with now 13 strategies Adaptive Tit for Tat ranked first.

Another extension of strategies was that of teams of strategies **J.P.Delahaye1993Lp**; **J.P.Delahaye1995LIeP**; **A.Rogers2007Ctpw** that collude to increase one member's score. In 2004 Graham Kendall led the Anniversary Iterated Prisoner's Dilemma Tournament with a total of 223 entries. In this tournament participants were allowed to submit multiple strategies. A team from the University of Southampton submitted a total of 60 strategies **A.Rogers2007Ctpw**. All these were strategies that had been programmed with a recognition mechanism by default. Once the strategies recognised one another, one would act as leader and the other as a follower. The follower plays as a **Cooperator**, cooperates unconditionally and the leader would play as a **Defector** gaining the highest achievable score. The followers would defect unconditionally against other strategies to lower their score and help the leader. The result was that Southampton had the top three performers. Nick Jennings, who was part of the team, said that "We developed ways of looking at the Prisoner's Dilemma in a more realistic environment and we devised a way for computer agents to recognise and collude with one another despite the noise. Our solution beats the standard Tit For Tat strategy" **southampton˙blog**.

2.3.1 Memory one Strategies

A set of strategies that have received a lot of attention in the literature are **memory one** strategies. In **nowak1989**, Nowak and Sigmund proposed a structure for studying simple strategies that remembered only the previous turn, and moreover, only recorded the move of the opponent. These are called **reactive** strategies and they can be represented by using three parameters (y, p_1, p_2), where y is the probability to cooperate in the first move, and p_1 and p_2 the conditional probabilities to cooperate, given that the opponent's last move was a cooperation or a defection. For example Tit For Tat is a reactive strategy and it can be written as $(1, 1, 0)$. Another reactive strategy well known in the literature is **Generous Tit for Tat Nowak1992**.

In **Nowak1990**, Nowak and Sigmund extended their work to include strategies which consider the entire history of the previous turn to make a decision. These are called **memory one** strategies. If only a single turn of the game is taken into account and depending on the simultaneous moves of the two players there are only four possible states that the players could be in. These are:

- Both players cooperated, denoted as CC.
- First player cooperated while the second one defected, denoted as CD.
- First player defected while the second one cooperated, denoted as DC.
- Both players defected, denoted as DD.

Thus a memory one strategy can be denoted by the probabilities of cooperating after each state and the probability of cooperating in the first round, (y, p_1, p_2, p_3, p_4). For example Pavlov's memory one representation is $(1, 1, 0, 0, 1)$.

Memory one strategies made an impact when a specific set of memory one strategies were introduced called **Zero-determinant** (ZD) **Press2012**. The American Mathematical Society's news section **hilbe2015** stated that "the world of game theory is currently on fire" and in **Stewart2012** it was stated that "Press and Dyson have fundamentally changed the viewpoint on the Prisoner's Dilemma". ZD are a set of extortionate strategies that can force a linear relationship between the long-run scores of both themselves and the opponent, therefore ensuring that the opponent will never do better than them.

Press and Dyson's suggested ZD strategies were the dominant family of strategies in the IPD. Moreover, they argued that memory is not beneficial. In **Adami2013**; **Knight2017**; **Hilbe2013**; **Hilbe2013b**; **hilbe2015**; **KnightHGC17**; **Knight2019**; **Lee2015**; **Stewart2012** the effectiveness of ZD strategies is questioned. In **Adami2013**, it was shown that ZD strategies are not evolutionary stable, and in **Stewart2012** a more generous set of ZDs, the **Generous ZD**, were shown to outperform the more extortionate ZDs. Finally, in **Knight2017**; **KnightHGC17**; **Knight2019**; **Lee2015**, the 'memory does not benefit a strategy' statement was questioned. A set of more complex strategies, strategies that take in account the entire history set of the game, were trained and proven to be more stable than ZD strategies.

This section covered the original computer tournaments of Axelrod and the early success of Tit For Tat in these tournaments. Though Tit For Tat was considered to be the most robust basic strategy, reciprocity was found to not be enough in environments with uncertainties. There are at least two properties, that have been discussed in this section, for coping with such uncertainties; generosity and contrition. Generosity is letting a percentage of defections go unpunished, and contrition is lowering a strategy's readiness to defect following an opponent's defection.

In the later part of this section a series of new strategies which were built on the basic reciprocal approaches were presented, followed by the infamous memory one strategies, the zero-determinant strategies. Though the ZDs can be proven to be robust in pairwise interactions they were found to be lacking in evolutionary settings and in computer tournaments. Evolutionary settings and the emergence of cooperation under natural selection are covered in the next section.

2.4 Evolutionary dynamics

As yet, the emergence of cooperation has been discussed in the contexts of the one shot PD game and the IPD round robin tournaments. In the PD it is proven that cooperation will not emerge, furthermore, in a series of influential works Axelrod demonstrated that reciprocal behaviour favours cooperation when individuals interact repeatedly. But does natural selection favours cooperation? Understanding the conditions under which natural selection can favour cooperative behaviour is important in understanding social behaviour amongst intelligent agents **Boyd1987**.

Imagine a mixed population of cooperators and defectors where every time two individuals meet they play a game of PD. In such population the average payoff for defectors is always higher than cooperators. Under natural selection the frequency of defectors will steadily increase until cooperators become extinct. Thus natural selection favours defection in the PD (Figure 2.1). However, there are several mechanisms that allow the emergence of cooperation in an evolutionary context which will be covered in this section.

In the later sections of **Axelrod1980b**, Axelrod discusses an ecological tournament that he performed using the 62 strategies of the second tournament to understand the reproductive success of Tit for Tat. In his ecological tournament the prevalence of each type of strategy in each round was determined by that strategy's success in the previous

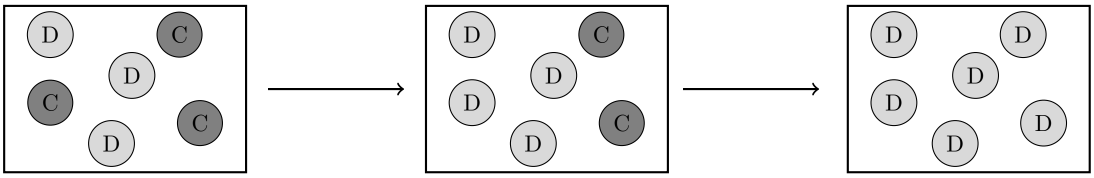


Figure 2.1: Natural selection favours defection in a mixed population of Cooperators and Defectors.

round. The competition in each round would become stronger as weaker performers were reduced and eliminated. The ecological simulation concluded with a handful of nice strategies dominating the population whilst exploitative strategies had died off as weaker strategies were becoming extinct. This new result led Axelrod to study the IPD in an evolutionary context based on several of the approaches established by the biologist John M. Smith **Smith1974**; **Smith1979**; **Smith1973**. John M. Smith was a fundamental figure in evolutionary game theory and a participant of Axelrod's second tournament. Axelrod and the biologist William Donald Hamilton wrote about the biological applications of the evolutionary dynamics of the IPD **Axelrod1984** and won the Newcomb-Cleveland prize of the American Association for the Advancement of Science.

In Axelrod's model **axelrod1981** pairs of individuals from a population played the IPD. The number of interactions between the pairs were not fixed, but there was a probability defined as the importance of the future of the game w, where $0 < w < 1$, that the pair would interact again. In **axelrod1981** it was shown that for a sufficient high w Tit For Tat strategies would become common and remain common because they were "collectively stable". Axelrod argued that collective stability implied evolutionary stability (ESS) and that when a collectively stable strategy is common in a population and individuals are paired randomly, no other rare strategy can invade. However, Boyd and Lorderbaum in **Boyd1987** proved that if w, the importance of the future of the game, is large enough then no pure strategy is ESS because it can always be invaded by any pair of other strategies. This was also independently proven in **Pudaite1987**.

All these conclusions were made in populations where the individuals could all interact with each other. In 1992, Nowak and May, considered a structured population where an individual's interactions were limited to its neighbours. More specifically, in **Nowak1992b** they explored how local interaction alone can facilitate population wide cooperation in a one shot PD game. The two deterministic strategies Defector and Cooperator, were placed onto a two dimensional square array where the individuals could interact only with the immediate neighbours. The number of immediate neighbours could be either, fourth, six or eight, as shown in Figure 2.2, where each node represents a player and the edges denote whether two players will interact. This topology is refereed to as spatial topology. Each cell of the lattice is occupied by a Cooperator or a Defector and at each generation step each cell owner interacts with its immediate neighbours. The score of each player is calculated as the sum of all the scores the player achieved at each generation. At the start of the next generation, each lattice cell is occupied by the player with the highest score among the previous owner and their immediate neighbours.

Local interactions proved that as long as small clusters of cooperators form, where they can benefit from interactions with other cooperators while avoiding interactions with defectors, global cooperation will continue. Thus, local interactions proved that even for the PD cooperation can emerge. Moreover in **Ohtsuki2006**, whilst using the payoff matrix (2.4), it was shown that cooperation will evolve in a structured population as long as the benefit to cost ratio b/c is higher than the number of neighbours. In **Perc2011**, graphs were a probability of rewiring ones connections was considered were studied. The rewire could be with any given node in the graphs and not just with immediate neighbours. Perc et al. concluded that “making new friends” may be an important activity for the successful evolution of cooperation, but also they must be selected carefully and one should keep their number limited.

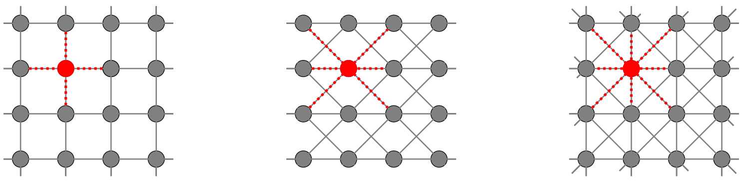


Figure 2.2: Spatial neighbourhoods

Another approach for increasing the likelihood of cooperation by increasing of assortative interactions among cooperative agents, include partner identification methods such as reputation **Janssen2006**; **Nowak1998**; **Suzuki2005**, communication tokens **Miller2002** and tags **Choi2006**; **Hales2000**; **Miller2002**; **Riolo2001**.

In this section evolutionary dynamics and the emergence of cooperation were reviewed. The following section focuses on strategy archetypes, training methods and strategies obtained from training.

2.5 Structured strategies and training

This section covers strategies that are different to that of intelligent design discussed in Section 2.3. These are strategies that have been **trained** using generic strategy archetypes. For example, in **Axelrod1987** Axelrod decided to explore deterministic strategies that took into account the last 3 turns of the game. As discussed in Section 2.3.1, for each turn there are 4 possible outcomes, CC, CD, DC, DD, thus for 3 turns there are a total of $4 \times 4 \times 4 = 64$ possible combinations. Therefore, the strategy can be defined by a series of 64 C's/D's, corresponding to each combination; this type of strategy is called a lookup table. This lookup table was then trained using a genetic algorithm **Koza1997**. During the training process random changes are made to a given lookup table. If the utility of the strategy has increased this change is kept, otherwise not.

In 1996 John Miller considered finite state automata as an archetype **Miller1996**, more specifically, Moore machines **moore1956**. He used a genetic algorithm to train finite state machines in environments with noise. Miller's results showed that even a small difference in noise (from 1% to 3%) significantly changed the characteristics of the evolving strategies. The strategies he introduced were **Punish Twice**, **Punish Once for Two Tats** and **Punish Twice and Wait**. In **Ashlock2006b** finite state automata and genetic algorithms were also used to introduce new strategies. In a series of experiments where the size of the population varied, there were two strategies frequently developed by the training process and more over they were developed only after the evolution had gone on for many generations. These were **Fortess3** and **Fortess4**. Following Miller's work in 1996, the first structured strategies based on neural networks that had be trained using a genetic algorithm was introduced in **Harrald1996** by Harrald and Fogel. Harrald and Fogel considered a single layered neural network which had 6 inputs. These were the last 3 moves of the player and the opponent, similar to **Axelrod1987**. Neural networks have broadly been used to train IPD strategies since then with genetic algorithms **Ashlock2006a**; **Chong2005**; **Marks1999** and particle swarm optimization **Franken2005**.

In **Knight2017**; **KnightHGC17** both genetic algorithm and particle swarm optimization were used to introduce a series of structured strategies based on lookup tables, finite state machines, neural networks, hidden Markov models **eddy1996** and Gambler. Hidden Markov models, are a stochastic variant of a finite state machine and Gamblers are stochastic variants of lookup tables. The structured strategies that arised from the training were put up against a large number of strategies in (1) a Moran process, which is an evolutionary model of invasion and resistance across time during which high performing individuals are more likely to be replicated and (2) a round robin tournament. In a round robin tournament which was simulated using the software **axelrodproject** and the 200 strategies implemented within the software, the top spots were dominated by the trained strategies of all the archetypes. The top three strategies were **Evolved LookUp 2 2 2**, **Evolved HMM 5** and **Evolved FSM 16**.

In **KnightHGC17** it was demonstrated that these trained strategies would overtake the population in a Moran process. The strategies evolved an ability to recognise themselves by using a handshake. This recognition mechanism allowed the strategies to resist invasion by increasing the interactions between themselves, an approach similar to the one described in Section 2.4.

Throughout the different methods of training that have been discussed in this section, a spectrum of structured strategies can be found. Differentiating between strategies is not always straightforward. It is not obvious looking at a finite state diagram how a machine will behave, and many different machines, or neural networks can represent the same strategy. For example Figure 2.3 shows two finite automata and both are a representation of Tit for Tat.

(a) Tit for Tat as a finite state machine with 1 state.

(b) Tit for Tat as a finite state machine with 2 states.

Figure 2.3: Finite state machine representations of Tit for Tat. A machine consists of transition arrows associated with the states. Each arrow is labelled with A/R where A is the opponent's last action and R is the player's response. Finite state machines consist of a set of internal states. In (a) Tit for Tat finite state machine consists of 1 state and in (b) of 2.

To allow for identification of similar strategies a method called fingerprinting was introduced in **Ashlock2005**. The method of fingerprinting is a technique for generating a functional signature for a strategy **Ashlock2008**. This is achieved by computing the score of a strategy against a spectrum of opponents. The basic method is to play the strategy against a probe strategy with varying noise parameters. In **Ashlock2005** Tit for Tat is used as the probe strategy. In Figure 2.4 an example of Pavlov's fingerprint is given. Fingerprinting has been studied in depth in **Ashlock2008**; **Ashlock2009**; **Ashlock2010**; **Ashlock2006a**. Another type of fingerprinting is the transitive fingerprint **axelrodproject**. The method represents the cooperation rate of a strategy against a set of opponents over a number of turns. An example of a transitive fingerprint is given in Figure 2.5.

This section covered structured strategies and training methods. In the following section software that has been developed with main aim simulating the IPD is presented.

2.6 Software

The research of the IPD heavily relies on software. This is to be expected as computer tournaments have become the main means of simulating the interactions in an IPD

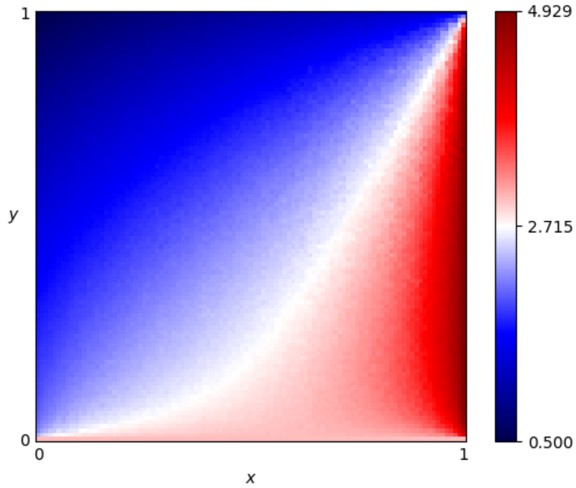


Figure 2.4: Pavlov fingerprinting with Tit for Tat used as the probe strategy. Figure was generated using **axelrodproject**.

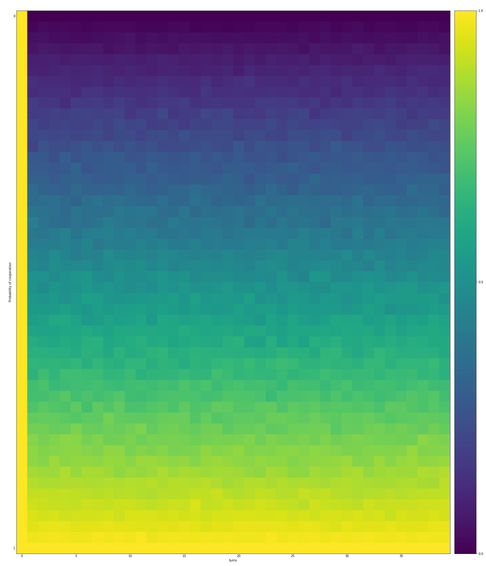


Figure 2.5: Transitive fingerprint of Tit for Tat against a set of 50 random opponents.

game. Many academic fields suffer from lack of source code availability and the IPD is not an exception. Several of the tournaments that have been discussed so far were generated using computer code, though not all of the source code is available. The code for Axelrod's original tournament is known to be lost and moreover for the second tournament the only source code available is the code for the 62 strategies (found on Axelrod's personal website **fortan˙code**).

Several projects, however, are open, available and have been used as research tools or educational platforms over the years. Two research tools **prison**; **axelrodproject** and two educational tools **pd˙trust**; **pd˙game** are briefly mentioned here. Both **prison**; **axelrodproject** are open source projects. The "Game of Trust" **pd˙trust** is an on-line, graphical user interface educational platform for learning the basics of game theory, the IPD and the notion of strategies. It attracted a lot of attention due to being "well-presented with scribble-y hand drawn characters" **trust˙blogb** and "a whole heap of fun" **trust˙bloga**. Finally **pd˙game** is a personal project written in PHP. It is a graphical user interface that offers a big collection of strategies and allows the user to try several matches and tournament configurations.

PRISON **prison** is written in the programming language Java and a preliminary version was launched on 1998. It was used by its authors in several publications, such as **Beaufils1997**, which introduced Gradual, and **Beaufils1988**. The project includes a good number of strategies from the literature but unfortunately the last update of the project dates back to 2004. Axelrod-Python **axelrodproject** is a software used by **Knight2017**; **KnightHGC17**; **Goodman2018**; **Wang2017**. It is written in the programming language Python following best practice approaches **Aberdour2007**; **Benureau2018** and contains the largest collection of strategies, known to the author. The strategy list of the project has been cited by publications **Anastassacos2018**; **Hayes2017**; **Neumann2018**.

2.7 Conclusion

This manuscript presented a literature review on the Iterated Prisoner's Dilemma. The opening sections focused on research trends and published works of the field, followed by a presentation of research and educational software. More specifically, Section 2.2 covered the early years of research. This was when simulating turns of the game was only possible with human subject research. Following the early years, the pioneering tournaments of Axelord were introduced in Section 2.3. Axelrod's work offered the field an agent based game theoretic framework to study the IPD. In his original papers he asked researchers to design strategies to test their performance with the new framework. The winning strategy of both his tournaments was Tit for Tat. The strategy however came with limitations which were explored by other researchers, and new intelligently designed strategies were introduced in order to surpass Tit for Tat with some contributions such as Pavlov and Gradual.

Soon researchers came to realise that strategies should not just do well in a tournament setting but should also be evolutionary robust. Evolutionary dynamic methods were applied to many works in the field, and factors under which cooperation emerges were explored, as described in Section 2.4. This was not done only for unstructured populations, where all strategies in the population can interact with each other, but also in population where interactions were limited to only strategies that were close to each other. In such topologies it was proven that even in the one shot game, cooperation can indeed emerge.

Evolutionary approaches can offer many insights in the study of the PD. In evolutionary settings strategies can learn to adapt and take over population by adjusting their actions; such algorithms can be applied so that evolutionarily robust strategies can emerge. Algorithms and structures used to train strategies in the literature were covered in Section 2.5. From these training methods several strategies are found, and to be able to differentiate between them fingerprinting was introduced. The research of best play and cooperation has been going on since the 1950s, and for simulating the game software has been developed along the way. This software has been briefly discussed in Section 2.6.

The study of the PD is still an ongoing field research where new variants and new structures of strategies are continuously being explored **Ohtsuki2018**. The game now serves as a model in a wide range of applications, for example in medicine and the study of cancer cells **archetti2018**; **Kaznatchee2017**, as well as in social situations and how they can be driven by rewards **Dridi2018**. New research is still ongoing for example in evolutionarily dynamics on graphs **Allen2017**; **hathcock2018**; **Liu2017**.

Chapter 3

A bibliometric study of research topics, collaboration and influence in the field of the Iterated Prisoner's Dilemma

This manuscript explores the research topics and collaborative behaviour of authors in the field of the Prisoner's Dilemma using topic modeling and a graph theoretic analysis of the co-authorship network. The analysis identified five research topics in the Prisoner's Dilemma which have been relevant of the course of time. These are human subject research, biological studies, strategies, evolutionary dynamics on networks and modeling problems as a Prisoner's Dilemma game. Moreover, the results demonstrated the Prisoner's Dilemma is a field of continued interest, and although it is a collaborative field, it is not necessarily more collaborative than other scientific fields. The co-authorship network suggests that authors are focused on their communities and not many connections across the communities are made. The Prisoner Dilemma authors also do not influence or gain much information by their connections, unless they are connected to a "main" group of authors.

3.1 Introduction

The Prisoner's Dilemma (PD) is a well known game used since its introduction in the 1950's **Flood1958** as a framework for studying the emergence of cooperation; a topic of continued interest for mathematical, social **Perc2008**, biological **Turner1999** and ecological **Wu2011** sciences. This manuscript presents a bibliometric analysis of 2,420 published articles on the Prisoner's Dilemma between 1951 and 2018. It presents the dominant topics in the PD publications, which have been identified using Latent

Dirichlet Allocation **Blei2003**, and it explores the changes in the dominant topics over time. The collaborative behaviour of the field is explored using the co-authorship network, and furthermore, the Latent Dirichlet Allocation topic analysis is combined with the co-authorship network analysis to assess the relative influence of authors in these topics. Assessing the collaborative behaviour of the field of collaboration itself is the main aim of this work.

As discussed in **youngblood2018**, bibliometrics (the statistical analysis of published works originally described by **pritchard1969**) has been used to support historical assumptions about the development of fields **raina1998**, identify connections between scientific growth and policy changes **das2016**, develop a quantitative understanding of author order **sekara2018**, and investigate the collaborative structure of an interdisciplinary field **Liu2015**. Most academic research is undertaken in the form of collaborative effort and as **Kyvik2017** points out, it is rational that two or more people have the potential to do better as a group than individually. Indeed this is the very premise of the Prisoner's Dilemma itself. Collaboration in groups has a long tradition in experimental sciences and it has be proven to be productive according to **Etzkowitz1992**. The number of collaborations can be different between research fields and understanding how collaborative a field is not always an easy task. Several studies tend to consider academic citations as a measure for these things. A blog post published by Nature **nature˙blog** argues that depending on citations can often be misleading because the true number of citations can not be known. Citations can be missed due to data entry errors, academics are influenced by many more papers than they actually cite and several of the citations are superficial.

A more recent approach to measuring collaborative behaviour, and to studying the development of a field is to use the co-authorship network, as described in **Liu2015**. The co-authorship network has many advantages as several graph theoretic measures can be used as proxies to explain author relationships. For example the average degree of a node corresponds to the average number of an authors' collaborators, and clustering coefficient corresponds to the extent that two collaborators of an author also collaborate with each other. In **Liu2015**, the approach was applied to analyse the development of the field "evolution of cooperation", and in **youngblood2018** to identify the subdisciplines of the interdisciplinary field of "cultural evolution" and investigate trends in collaboration and productivity between these subdisciplines. Moreover, **Li2019** examined the long-term impact of co-authorship with established, highly-cited scientists on the careers of junior researchers. This paper builds upon the work done by **Liu2015** and **youngblood2018**, and extends their methodology. In **Liu2015**; **youngblood2018**, a data set from a single source, Web of Science, is considered whereas the data set described here, archived at **pd˙data˙2018**, has been collected from five sources.

Latent Dirichlet Allocation (LDA) is a topic modeling technique proposed in **Blei2003** as a generative probabilistic model for discovering underlying topics in collections of data. Applications of the technique include detection in image data **Agarwal2008**; **Coelho2010** and detection in video **Niebles2008**; **Wang2008**. Nevertheless, LDA has been applied by several works on publication data for identifying the topic structure of a subject area. In **Inglis2018**, it was applied to the publications on mathematical education of the journals "Educational Studies in Mathematics" and "Journal for Research in Mathematics Education" to identify the dominant topics that each journal was publishing on. The topics of the North American library and Information Science dissertations were studied chronologically in **Sugimoto2011**, and the main topic of the scientific content presented at EvoLang conferences was identified in **Bergmann2018**. In **Bergmann2018** the LDA approach is combined with clustering and a co-authorship network analysis. A clustering analysis is applied to the LDA topics, and the co-authorship network is analysed as a whole where the clusters are only used to differentiate between the authors' topics. In comparison, this works applies LDA to identify dominant topics in the Prisoner's Dilemma fields and analyses the networks corresponding to these topics individually.

The methodology used in this manuscript (including the data collection) is covered in Section 3.2 and a preliminary analysis of the data set is presented in Section 3.3. This manuscript makes usage of the methodology and data to address the following questions:

1. What are the research topics of the Prisoner's Dilemma?
2. Is one topic currently more in fashion?
3. How have the research topics changed over the years?
4. Is the Prisoner's Dilemma a collaborative field?
5. Are some topics more collaborative than others?
6. Are there authors which benefit more from their position in the network?

Results regarding questions 1-3 are presented in Section 3.4 and questions 4-6 are addressed in Section 3.5. The results are summarised in Section A.5.

3.2 Methodology

Academic articles are accessible through scholarly databases. Several databases and collections today offer access through an open application protocol interface (API). An API allows users to query directly a journal's database and bypass the graphical user interface. Interacting with an API has two phases: requesting and receiving. The request phase includes composing a url with the details of the request.

For example, `http://export.arxiv.org/api/query?search_query=abs:prisoner'sdilemma&max_results=1` represents a request message. The first part of the request is the address of the API. In this example the address corresponds to the API of arXiv. The second part of the request contains the search arguments. In this example it is requested that the word 'prisoners dilemma' exists within the article's title. The format of the request message is different from API to API. The receive phase includes receiving a number of raw metadata of articles that satisfies the request message. The raw metadata are commonly received in extensive markup language (xml) or Javascript object notation (json) formats **nurseitov2009**. Similarly to the request message, the structure of the received data differs from journal to journal.

The data collection is crucial to this study. To ensure that this study can be reproduced all code used to query the different APIs has been packaged as a Python library and is available online **nikoleta˙2017**. The software could be used for any type of projects similar to the one described here, documentation for it is available at: `http://arcas.readthedocs.io/en/latest/`. Project **nikoleta˙2017** allow users to collect articles from a list of APIs by specifying just a single keyword. Articles for which any of the terms "prisoner's dilemma", "prisoners dilemma", "prisoner dilemma", "prisoners evolution", "prisoner game theory" existed within the title, the abstract or the text are included in the analysis. Four prominent journals in the field and a preprint server were used as sources to collect data for this analysis:

- arXiv **mckiernan2000**; a repository of electronic preprints. It consists of scientific papers in the fields of mathematics, physics, astronomy, electrical engineering, computer science, quantitative biology, statistics, and quantitative finance, which all can be accessed online.
- PLOS **plos**; a library of open access journals and other scientific literature under an open content license. It launched its first journal, PLOS Biology, in October 2003 and publishes seven journals, as of October 2015.
- IEEE Xplore Digital Library (IEEE) **ieee**; a research database for discovery and access to journal articles, conference proceedings, technical standards, and related materials on computer science, electrical engineering and electronics, and allied fields. It contains material published mainly by the Institute of Electrical and Electronics Engineers and other partner publishers.
- Nature **nature**; a multidisciplinary scientific journal, first published on 4 November 1869. It was ranked the world's most cited scientific journal by the Science Edition of the 2010 Journal Citation Reports and is ascribed an impact factor of 40.137, making it one of the world's top academic journals.
- Springer **springer**; a leading global

scientific publisher of books and journals. It publishes close to 500 academic and professional society journals.

The data set has been archived and is available at **pd˙data˙2018**. Note that the latest data collection was performed on the 30$^{\text{th}}$ November 2018.

The relationship between the authors within a field will be modeled as a graph $G = (V_G, E_G)$ where V_G is the set of nodes and E_G is the set of edges. The set V_G represents the authors and an edge connects two authors if and only if those authors have written together. This co-authorship network is constructed using the main data set **pd˙data˙2018** and the open source package **networkx**. The PD network is denoted as G where the number of unique authors $|V(G)|$ is 4226 and $|E(G)|$ is 7642 . All authors' names were formatted as their first name and last name (i.e. Martin A. Nowak to Martin Nowak). This was done to avoid errors such as Martin A. Nowak and Martin Nowak being treated as a different person. There are some authors for which only their first initial was found. These entries are left as such.

The collaborativeness of the authors will be analysed using measures such as, isolated nodes, connected components, clustering coefficient, communities, modularity and average degree. These measures show the number of connections authors can have and how strongly connected these people are. The number of isolated nodes is the number of nodes that are not connected to another node, thus the number of authors that have published alone. The average degree denotes the average number of neighbours for each nodes, i.e. the average number of collaborations between the authors. A connected component is a maximal set of nodes such that each pair of nodes is connected by a path **Easley2010**. The number of connected components as well as the size of the largest connected component in the network are reported. The size of the largest connected component represents the scale of the central cluster of the entire network, as will be discussed in the analysis section. Clustering coefficient and modularity are also calculated. The clustering coefficient, defined as 3 times the number of triangles on the graph divided by the number of connected triples of nodes, is a local measure of the degree to which nodes in a graph tend to cluster together in a clique **Easley2010**. It shows to which extent the collaborators of an author also write together. In comparison, modularity is a global measure designed to measure the strength of division of a network into communities. The number of communities will be reported using the Clauset-Newman-Moore method **clauset2004**. Also the modularity index is calculated using the Louvain method described in **Blondel2008**. The value of the modularity index can vary between $[-1, 1]$, a high value of modularity corresponds to a structure where there are dense connections between the nodes within communities but sparse connections between nodes in different communities. That means that there are many sub communities of authors that write together but not across communities. Two further points are aimed to be explored in this work, (1) which people control the flow of information; as in which people influence the field the most and (2) which are the authors that gain the most from the influence of the field. To measure these concepts centrality measures are going to be used. Centrality measures are often used to understand different aspects of social networks **Landherr2010**. Two centrality measures have been chosen for this paper and these are closeness and betweenness centrality.

1. In networks some nodes have a short distance to a lot of nodes and consequently are able to spread information on the network very effectively. A representative of this idea is **closeness centrality**, where a node is seen as centrally involved in the network if it requires only few intermediaries to contact others and thus is structurally relatively independent. Closeness centrality is interpreted as influence. Authors with a high value of closeness centrality, are the authors that spread scientific knowledge easier on the network and they have high influence.

2. Another centrality measure is the **betweenness centrality**, where the determination of an author's centrality is based on the quotient of the number of all shortest paths between nodes in the network that include the node in question and the number of all shortest paths in the network. In betweenness centrality the position of the node matters. Nodes with a higher value of betweenness centrality are located in positions that a lot of information pass through, this is interpreted as the gain from the influence, thus these authors gain the most from their networks.

The articles contained in the data set (**pd˙data˙2018**) will be classified into research topics using LDA an unsupervised machine learning technique designed to summarize large collections of documents by a small number of conceptually connected topics or themes **Blei2003**; **Grimmer2013**. The documents are the articles' abstracts and LDA was carried out using **rehurek˙lrec**. In LDA, each document/abstract is represented by a distribution over topics, and the topics themselves are represented by a distribution over words. More specifically, each topics is described by weights associated with words and each document by the probabilities of belonging to a specific topic. The probability of a document belonging to a topic is referred to as the percentage contribution denoted as c. For example the words and their associated weights for two topics A and B could be:

- Topic A: $0.039\times$"cooperation", $0.028\times$"study" and $0.026\times$"human".
- Topic B: $0.020\times$"cooperation", $0.028\times$"agents" and $0.026\times$"strategies".

The percentage contribution for a document with abstract "The study of cooperation in humans" has a $c_A = 0.039{+}0.028{+}0.026 = 0.093$ and $c_B = .020{+}0.0{+}0.0 = 0.020$. The topic to which a document is assigned to is based on the highest percentage contribution denoted as c^*. For the given example the dominant topic is Topic A $c^* = c_A$. LAD requires that the number of topics is specified in advance before running the algorithm. The number of topics can be chosen using the coherence value **Roder2015** or through subjective minimisation of the overlapping keywords between two topics. Both these approaches will be used in this work.

Several of the approaches described in this section have previously been carried out in **Bergmann2018**; **Liu2015**; **Sugimoto2011**; **youngblood2018**, the novelty here is combining the approaches as well as applying them to a new data set. A preliminary analysis of the data set is presented in the following section.

3.3 Preliminary Analysis

The data set **pd˙data˙2018** consists of 2422 articles with unique titles. In case of duplicates the preprint version of an article (collected from arXiv) was dropped. Similarly to **Liu2015**, 76 articles have not been collected from the aforementioned APIs but have been manually added because they are of interest. Examples of such papers include **Flood1958** the first publication on the PD, **Ohtsuki2006**; **Stewart2012** two well cited articles in the field, and a series of works from Robert Axelrod **Axelrod1980**; **Axelrod1980more**; **Axelrod1987**; **Axelrod1981**; **Riolo2001** a leading author of the field. A more detailed summary of the articles' provenance is given by Table 3.1. Only 3% of the data set consists of articles that were manually added and 27% of the articles were collected from arXiv. The average number of publications is also included in Table 3.1. Overall an average of 43 articles are published per year on the topic. The most significant contribution to this appears to be from arXiv with 11 articles per year, followed by Springer with 9 and PLOS with 8.

	Number of Articles	Percentage %	Year of first publication	Average number of publications per year
IEEE	294	12.14%	1973	5
Manual	76	3.14%	1951	1
Nature	436	18.00%	1959	8
PLOS	477	19.69%	2005	8
Springer	533	22.01%	1966	9
arXiv	654	27.00%	1993	11
Overall	2470	100.00%	1951	43

Table 3.1: Summary of **pd˙data˙2018** per provenance.

The data handled here is in fact a time series from the 1950s, the formulation of the game, until 2018 (Figure 3.1). Two observations can be made from Figure 3.1.

1. There is a steady increase of the number of publications since the 1980s and the introduction of computer tournaments **Axelrod1981** (work by Robert Axelrod).

2. There is a decrease in 2017-2018. This is due to our data set being incomplete. Articles that have been written in 2017-2018 have either not being published or were not retrievable by the APIs at the time of the last data collection.

These observations can be confirmed by studying the time series. Using **scipy**, an exponential distribution is fitted to the data. The fitted model can be used to forecast the behaviour of the field for the next 5 years. Even though the time series has indicated a slight decrease, the model forecasts that the number of publications will keep increasing, thus demonstrating that the field of the PD continues to attract academic attention.

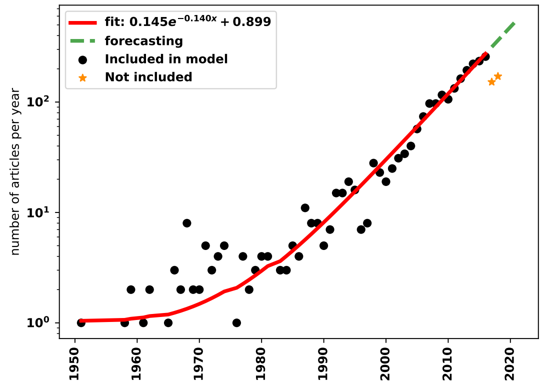


Figure 3.1: Number of articles published on the PD 1951-2018 (on a log scale), with a fitted exponential line, and a forecast for 2017-2022.

There are a total of 4226 authors in the data set (**pd˙data˙2018**) and several of these authors have had multiple publications collected from the data collection process. The highest number of articles collected for an author is 83 publications for Matjaz Perc. The distribution of the number of papers per author is given by Figure 3.2, and it can be seen that Matjaz Perc is an outlier. More specifically, most authors have 1 to 6 publications in the data set.

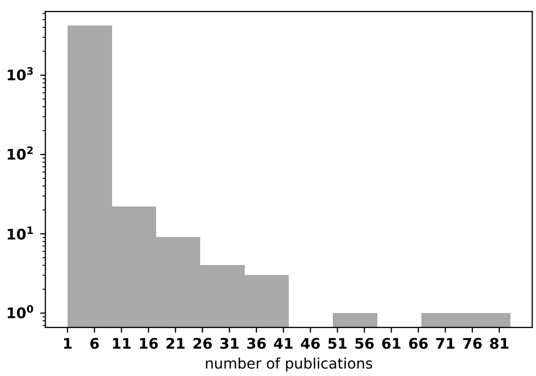


Figure 3.2: Distribution of number of papers per author (on a log scale).

The overall Collaboration Index (CI) or the average number of authors on multi-authored papers is 3.2, thus on average a non single author publication in the PD has 3 authors. This appears to be quite standard compared to other fields such as cultural evolution **youngblood2018**, Astronomy and Astrophysics, Genetics and Heredity, Nuclear and Particle Physics as reported by **nature˙author˙blog**. There are only a total of 545 publications with a single author, which corresponds to the 22% of the papers. It appears that academic publications tend to be undertaken in the form of collaborative effort, which is in line with the claim of **Kyvik2017**. From Figure 3.3 the trend of CI over the years is given. There are some peaks in the early years 1969 and 1980, however, a steady increase appears to happen after 2004. This could be an effect of better communication tools being introduced around that time which enabled more collaborations between researchers.

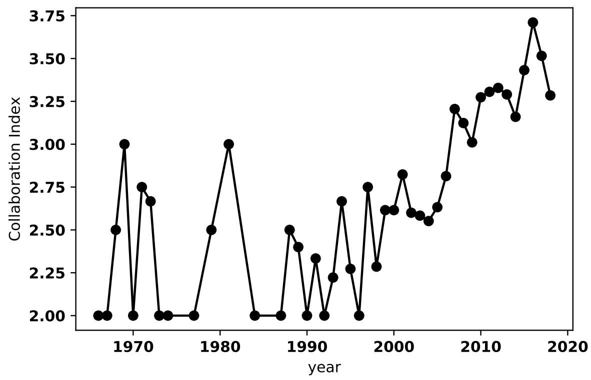


Figure 3.3: Collaboration index over time.

The collaborativeness of the authors is explored in more detail in Section 3.5 using the co-authorship network. The collaborative behaviour and relative influence of authors will also be explored in co-authorship networks which correspond to their publications research topics. These topics are presented in the next section.

3.4 Research topics in the Prisoner's Dilemma research

In order to identify the topics which are being discussed in the field of the PD, the LDA algorithm implemented in **rehurek˙lrec** is applied to the abstracts of the data set. As mentioned before, the number of topics, which will be denoted as n, needs to be specified before running the algorithm. The appropriate number of topics is chosen based on the coherence value **Roder2015**. Figure 3.4 gives the coherence values of 18 models where $n \in \{2, 3, \dots, 19\}$, and it can be seen than the most appropriate number of topics is 6 with a coherence value of 0.418.

An LDA model outputs an $N \times n$ matrix - N rows for N abstracts and n columns for n topics. The cells contain the percentage contributions for each topic for each abstract, c_i^j for $i \in \{1, 2, \dots, n\}$ for $j \in \{1, 2, \dots, N\}$. In essence, LDA maps every paper to a vector space of dimension the number of topics. In the case of 6 topics it is difficult to visualise the clustering of topics. To overcome this a dimensionality reduction approach

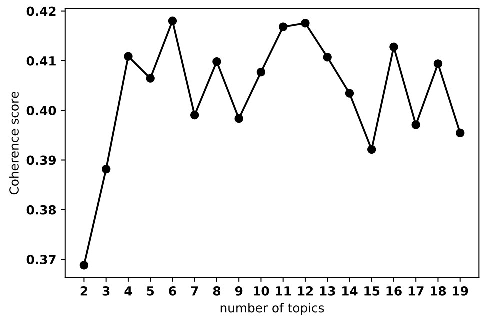


Figure 3.4: Coherence for LDA models over the number of topics.

called t-Distributed Stochastic Neighbor Embedding (t-SNE) **Maaten2008** is applied to the LDA model outputs. More specifically, t-SNE is used to reduce the dimensions of each c^j from n to 2. Figure 3.5, gives the visualisation of LDA for $n = 6$. Each point represents a single document and its color corresponds to the topic with the highest percentage contribution. The documents which are clustered together have a similar percentage contribution distribution over the topics.

Even though the LDA model with $n = 6$ has the highest coherence value, Figure 3.5 shows that documents of the same topic are closer to documents from other topics than each other. For example the documents of topic 2 are divided into two clusters. The one cluster is closer to documents from topic 4 and the other has a few documents closer to topic 1. In the case of $n = 6$ topic 4 appears to be on "evolution of cooperation on networks", and the papers from topic 2 surrounded from topic 4 include the articles "Evolutionary prisoner's dilemma game on hierarchical lattices" **Vukov2005** and "Social evolution in structured populations" **Debarre2014**. Publications that clearly also fit topic 4.

In comparison, 3.6 gives the visualisation of LDA $n = 5$ where the separation of the documents is more clear. Though several models, Figure 3.4, have a higher coherence value than the LDA model with $n = 5$, the separation of topics is not as clear for any model as it is for $n = 5$. Thus, $n = 5$ is chosen to carry out the analysis of this work, and moreover the LDA model for $n = 5$ has a coherence value 0.406 which is close to 0.418.

What are the research topics of the Prisoner's Dilemma?

For $n = 5$ the articles are clustered and assigned to their dominant topic, based on the highest percentage contribution. The keywords associated with a topic, the most representative article of the topic (based on the percentage contribution) and its academic reference are given by Table 3.2. The topics are labelled as A, B, C, D and E, and more specifically:

- Based on the keywords associated with Topic A, and the most representative arti-

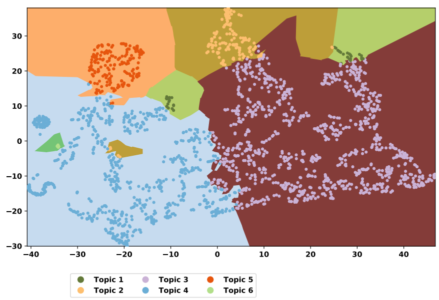


Figure 3.5: Visualisation of LDA with $n = 6$ on 2 dimensions.

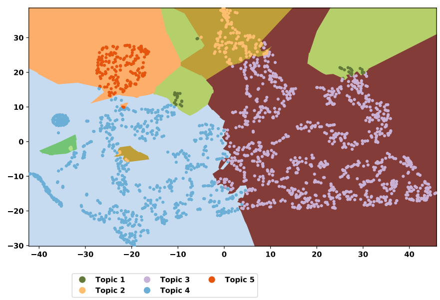


Figure 3.6: Visualisation of LDA with $n = 5$ on 2 dimensions.

cle, Topic A appears to be about **human subject research**. Several publications assigned to the topic study the PD by setting experiments and having human participants simulate the game instead of computer simulations. These articles include **Matsumoto2016** which showed that prosocial behavior increased with the age of the participants, **Li2014** which studied the difference in cooperation between high-functioning autistic and typically developing children, **Molina2013** explored the gender effect in highschool students and **Bell2017** explored the effect of facial expressions of individuals.

- Though it is not immediate from the keywords associated with Topic B, investigating the papers assigned to the topic indicate that it is focused on **biological studies**. Papers assigned to the topic include papers which apply the PD to genetics **Santorelli2008**; **Sistrom2015**, to the study of tumours **archetti2013evolutionary**; **sartakhti2017** and viruses **turner1999prisoner**. Other works include how phenotype affinity can affect the emergence of cooperation **wu2019phenotype** and modeling bacterial communities as a spatial structured social dilemma.

- Based on the keywords and the most representative article Topic C appears to include publications on PD **strategies**. Publications in the topic include the introduction of new strategies **stewart2013extortion**, the search of optimality in strategies **banerjee2007reaching** and the training of strategies **ishibuchi2011evolution** with different representation methods. Moreover, publications that study the evolutionary stability of strategies **adami2013evolutionary** and introduced methods of differentiating between them **ashlock2008fingerprinting** are also assigned to C.

- The keywords associated with Topic D clearly show that the topic is focused on **evolutionary dynamics on networks**. Publications include **ichinose2013robustness** which explored the robustness of cooperation on networks, **wang2012spatial** which studied the effect of a strategy's neighbourhood on the emergence of cooperation and **chen2016fixation** which explored the fixation probabilities of any two strategies is spatial structures.

- The publication assigned to Topic E are on **modeling problems as a PD game**. Though Topic B is also concerned with problems being formulated as a PD, it includes only biological problems. In comparison, the problems in Topic E include decision making in operational research **ormerod2010or**, information sharing among members in a virtual team **feng2008trilateral**, the measurement of influence in articles based on citations **hutchins2016relative** and the price spikes in electric power markets **Guan2002**, and not on biological studies.

Dominant Topic	Topic Keywords	Most Representative Article Title	Reference	# Documents	% Documents
A	social, behavior, human, study, experiment, cooperative, cooperation, suggest, find, behaviour	Facing Aggression: Cues Differ for Female versus Male Faces	**Geniole2012**	496.0	0.2008
B	individual, group, good, show, high, increase, punishment, cost, result, benefit	Genomic and Gene-Expression Comparisons among Phage-Resistant Type-IV Pilus Mutants of Pseudomonas syringae pathovar phaseolicola	**Sistrom2015**	309.0	0.1251
C	game, strategy, player, agent, dilemma, play, payoff, state, prisoner, equilibrium	Fingerprinting: Visualization and Automatic Analysis of Prisoner's Dilemma Strategies	**Sistrom2015**	561.0	0.2271
D	cooperation, network, population, evolutionary, evolution, interaction, dynamic, structure, cooperator, study	Influence of initial distributions on robust cooperation in evolutionary Prisoner's Dilemma	**Chen2007**	556.0	0.2251
E	model, theory, base, system, problem, paper, propose, information, provide, approach	Gaming and price spikes in electric power markets and possible remedies	**Guan2002**	548.0	0.2219

Table 3.2: Keywords for each topic and the document with the most representative article for each topic.

Note that the whilst for the choice of 5 topics the actual clustering is not subjective (the algorithm is determining the output) the interpretation above is.

Five topics in the PD publications identified by the data set of this work are human subject research, biological studies, strategies, evolutionary dynamics on networks and modeling problems as a PD.

These 5 topics nicely summarise the PD research. They highlight the interdisciplinarity of the field; how it brings together applied modeling of real world situations (Topic B and E) and more theoretical notions such as evolutionary dynamics and optimality of strategies.

Is one topic currently more in fashion?

Figure 3.7 gives the number of articles per topic over time. The topics appear to have had a similar trend over the years, with topics B and D having a later start. Following the introduction of a topic the publications in that topic have been increasing. There is no decreasing trend in any of the topics. All the topics have been publishing for years and they still attract the interest of academics. Thus, **there does not seem to be any given topic more or less in fashion**.

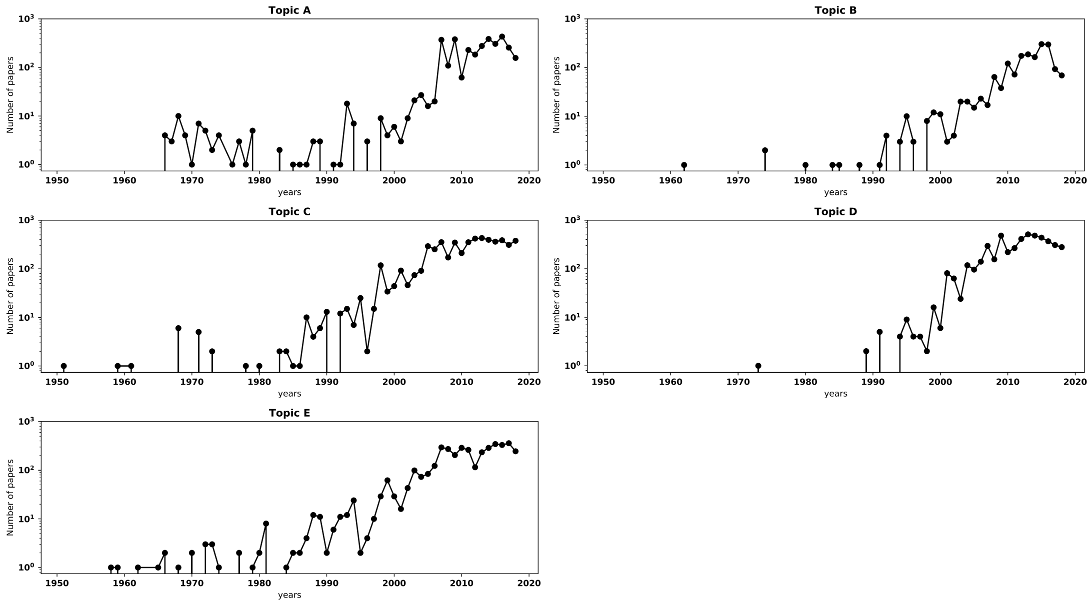


Figure 3.7: Number of articles per topic over the years (on a logged scale).

How do the research topics change over the years?

To gain a better understanding regarding the change in the topics over the years, LDA is applied to the cumulative data set over 8 time periods. These periods are 1951-1965, 1951-1973, 1951-1980, 1951-1988, 1951-1995, 1951-2003, 1951-2010, 1951-2018. The number of topics for each cumulative subset is chosen based on the coherence value and no objective approach is used. As a result, the period 1951-2018 has been assigned $n = 6$ which had the highest coherence value instead of 5. The chosen models for each period including the number of topics, their keywords and number of articles assigned to them are given by Table 3.3.

But how well do the five topics which were presented earlier fit the publications over time? This is answered by comparing the performance of three LDA models over the cumulative periods' publications. The three models are LDA models for the entire data set for n equal to 5, 6 and the optimal number of topics over time. For each model the c^* is estimated for each document in the cumulative data sets. The performance of the models are then compared based on:

$$\bar{c^*} \times n \tag{3.1}$$

where $\bar{c^*}$ is the median highest percentage contribution and n is the number of topics of a given period. A model with more topics will have more difficulty to assign

Period	Topic	Topic Keywords	Num of Documents	Percentage of Documents
1951-1965	1	problem, technology, divert, euler, subsystem, requirement, trace, technique, system, untried	3	0.375
1951-1965	2	interpret, requirement, programme, evolution, article, increase, policy, system, trace, technology	2	0.25
1951-1965	3	equipment, agency, conjecture, development, untried, programme, trend, technology, weapon, technique	1	0.125
1951-1965	4	variation, celebrated, trend, untried, change, involve, month, technique, subsystem, research	1	0.125
1951-1965	5	give, good, modern, trace, technique, ambiguity, problem, trend, technology, system	1	0.125
1951-1973	1	study, shock, cooperative, money, part, vary, investigate, good, receive, equipment	12	0.3243
1951-1973	2	cooperation, level, significantly, sequence, reward, provoke, descriptive, principal, display, argue	4	0.1081
1951-1973	3	player, make, effect, triad, experimental, motivation, dominate, hypothesis, instruction, trend	3	0.0811
1951-1973	4	ss, sex, male, female, dyad, design, suggest, college, factor, tend	3	0.0811
1951-1973	5	result, research, format, change, operational, analysis, relate, understanding, decision, money	2	0.0541
1951-1973	6	condition, give, high, treatment, conflict, cc, real, original, replication, promote	2	0.0541
1951-1973	7	group, competitive, show, interpret, scale, compete, escalation, free, variable, individualistic	2	0.0541
1951-1973	8	outcome, strategy, choice, type, pdg, difference, dummy, conclude, compare, consistent	2	0.0541
1951-1973	9	game, difference, pair, approach, behavior, person, weapon, occur, advantaged, differential	2	0.0541
1951-1973	10	response, present, dilemma, influence, cooperate, bias, point, amount, participate, factor	2	0.0541
1951-1973	11	trial, problem, previous, involve, prisoner, experiment, follow, tit, increase, initial	1	0.027
1951-1973	12	matrix, behavior, rational, black, model, research, broad, distance, complex, trace	1	0.027
1951-1973	13	play, finding, individual, noncooperative, white, nature, race, ratio, represent, prisoner	1	0.027
1951-1980	1	play, trial, group, follow, white, interpret, scale, black, trend, small	14	0.25
1951-1980	2	outcome, level, effect, type, dyad, vary, pdg, participate, understanding, arise	9	0.1607
1951-1980	3	game, strategy, cooperation, significant, difference, sentence, text, occur, differential, hypothesis	4	0.0714
1951-1980	4	male, female, find, result, sex, subject, experimental, situation, treatment, computer	4	0.0714
1951-1980	5	research, problem, influence, matrix, format, model, analysis, year, crime, equipment	4	0.0714
1951-1980	6	condition, dilemma, bias, free, attempt, book, year, dummy, prison, design	4	0.0714
1951-1980	7	variable, result, factor, individual, ability, triad, half, migration, change, investigate	3	0.0536
1951-1980	8	show, present, suggest, rational, compete, approach, characteristic, examine, person, conduct	3	0.0536
1951-1980	9	behavior, high, finding, relate, obtain, assistance, ratio, good, weapon, competition	3	0.0536
1951-1980	10	ss, shock, money, competitive, part, difference, pair, amount, man, information	3	0.0536
1951-1980	11	player, conflict, theory, decision, determine, produce, maker, cooperate, specialist, programming	2	0.0357
1951-1980	12	study, prisoner, make, response, experiment, noncooperative, standard, separate, conclude, initial	2	0.0357
1951-1980	13	give, cooperative, choice, cognitive, real, operational, set, subject, ascribe, concern	1	0.0179
1951-1988	1	trial, difference, find, choice, significant, competitive, effect, triad, interact, occur	24	0.2553
1951-1988	2	ss, shock, money, pair, response, part, high, tit, receive, amount	13	0.1383
1951-1988	3	suggest, paper, case, debate, view, achieve, framework, natural, assumption, finitely	10	0.1064
1951-1988	4	prisoner, dilemma, behavior, model, present, involve, person, increase, trust, experiment	8	0.0851
1951-1988	5	game, player, show, approach, repeat, previous, move, tat, related, include	8	0.0851
1951-1988	6	cooperation, level, mutual, equilibrium, standard, provide, information, human, real, question	6	0.0638
1951-1988	7	play, result, male, subject, female, cooperative, sex, experimental, treatment, computer	5	0.0532
1951-1988	8	research, study, variable, ability, factor, conflict, matrix, year, student, interpret	4	0.0426
1951-1988	9	problem, group, small, scale, social, issue, large, base, bias, party	4	0.0426
1951-1988	10	game, strategy, outcome, type, cooperate, ethical, pdg, explain, dependent, separate	4	0.0426
1951-1988	11	give, condition, individual, major, dyad, behaviour, produce, conflict, assistance, collectively	3	0.0319
1951-1988	12	situation, iterate, statement, rational, card, side, paradox, true, consequence, front	2	0.0213
1951-1988	13	inflation, hypothesis, rate, run, change, demand, nominal, cost, output, growth	2	0.0213
1951-1988	14	theory, make, analysis, decision, system, examine, work, soft, lead, hard	1	0.0106
1951-1995	1	strategy, population, evolution, iterate, tit, opponent, evolve, dynamic, set, tat	31	0.1732
1951-1995	2	game, repeat, assumption, rule, person, equilibrium, general, finitely, indefinitely, analyze	24	0.1341
1951-1995	3	inflation, long, rate, hypothesis, run, policy, cost, nominal, demand, programming	20	0.1117
1951-1995	4	condition, outcome, trial, find, difference, cooperation, experiment, level, significant, response	15	0.0838
1951-1995	5	rational, result, receive, statement, money, paradox, shock, iterate, consequence, common	14	0.0782
1951-1995	6	cooperation, show, competitive, high, probability, conflict, simulation, altruism, yield, natural	14	0.0782
1951-1995	7	prisoner, dilemma, give, point, defect, form, cooperator, increase, relate, ethical	10	0.0559
1951-1995	8	player, give, decision, provide, cooperative, game, previous, pair, determine, interact	9	0.0503
1951-1995	9	play, cooperate, result, male, subject, female, time, relationship, suggest, student	8	0.0447
1951-1995	10	problem, group, theory, good, approach, society, large, scale, issue, level	8	0.0447
1951-1995	11	study, situation, behaviour, computer, argue, change, implication, characteristic, real, associate	8	0.0447
1951-1995	12	model, paper, behavior, examine, present, mutual, expectation, develop, type, variable	7	0.0391
1951-1995	13	make, research, system, analysis, choice, work, base, relation, world, wide	6	0.0335
1951-1995	14	individual, social, behavior, standard, choose, evolutionary, partner, payoff, defection, small	5	0.0279
1951-2003	1	game, player, dilemma, prisoner, theory, give, paper, make, group, problem	151	0.4266
1951-2003	2	cooperation, result, play, show, cooperate, condition, cooperative, high, level, time	106	0.2994
1951-2003	3	strategy, model, agent, study, behavior, individual, population, evolutionary, state, player	97	0.274
1951-2010	1	model, theory, paper, base, make, present, problem, provide, human, decision	325	0.3454
1951-2010	2	game, strategy, player, agent, play, dilemma, system, behavior, show, state	322	0.3422
1951-2010	3	cooperation, network, study, population, individual, evolutionary, social, evolution, interaction, structure	294	0.3124
1951-2018	1	model, theory, system, base, paper, problem, propose, present, approach, provide	556	0.2251
1951-2018	2	behavior, social, human, decision, study, experiment, make, suggest, result, behaviour	482	0.1951
1951-2018	3	individual, group, good, social, punishment, level, cost, mechanism, dilemma, cooperative	428	0.1733
1951-2018	4	game, strategy, player, agent, play, dilemma, state, prisoner, payoff, equilibrium	380	0.1538
1951-2018	5	population, evolutionary, dynamic, model, selection, result, evolution, evolve, show, process	351	0.1421
1951-2018	6	cooperation, network, interaction, structure, study, evolution, find, behavior, cooperative, simulation	273	0.1105

Table 3.3: Topic modeling result for the cumulative data set over the periods

papers. Thus, equation (refeq:ratio) is a measure of confidence in assigning a given paper to its topic weighted by the number of topics. The performances are given by Figure 3.8.

The five topics of the PD presented in this manuscript appear to always be less good at fitting the publications compared to the six topics of LDA $n = 6$. Moreover, there are less good than the topics of the optimal number of topics from 1951 to 1995. The difference in the performance values, equation (3.1), however are small. **The relevances of the five topics has been increasing over time, and though, the topics did not always fit the majority of published work over time, there were still papers being published on those topics**.

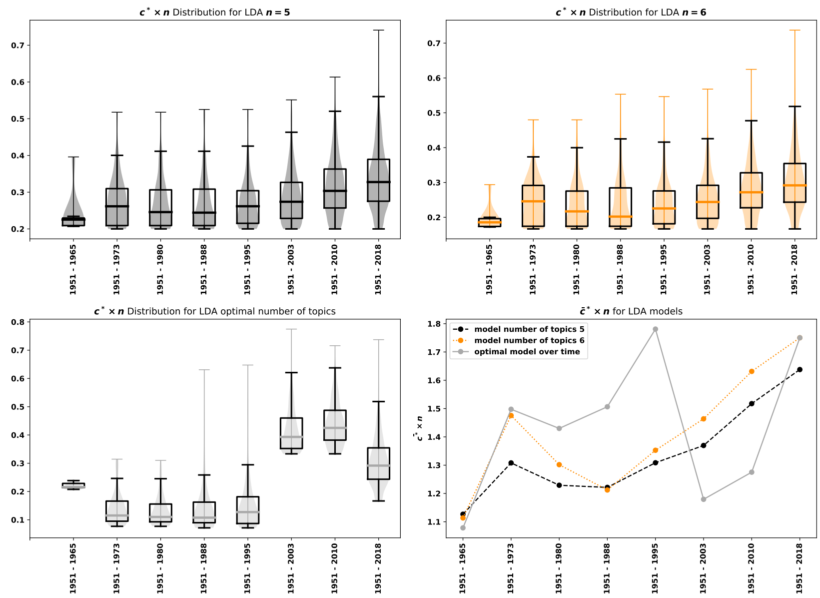
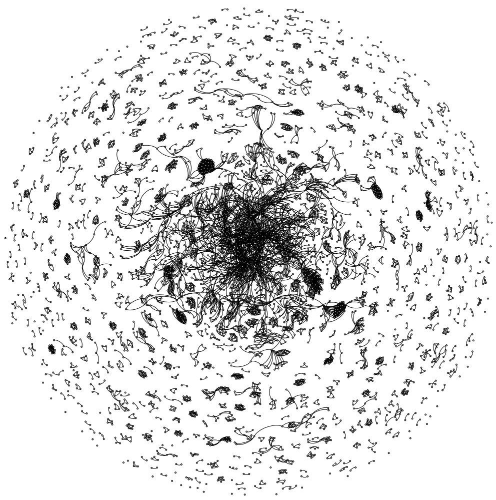


Figure 3.8: Maximum percentage contributions (c^*) over the time periods, for the LDA models for the entire data set for n equal to 5, 6 and the optimal number of topics over time.

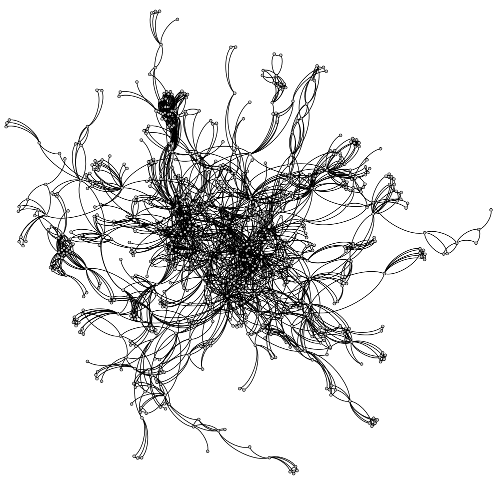
In the following section the collaborative behaviour of authors in the field, and within the field's topics as were presented in this section, are explored using a network theoretic approach.

3.5 Analysis of co-authorship network

The collaborative behaviour of authors in the field of the PD is assessed using the co-authorship network, which as mentioned in Section 3.2 is denoted as G. There are a total of 947 connected components in G and the largest component has a size of 796 nodes. The largest connected component is going to be refereed to as the main cluster of the network and is denoted as $\bar{G}$. A graphical representation of both networks is shown in Figure 3.9 and a metrics summary is given by Table 3.4.



(a) G the co-authorship network for the IPD.



(b) $\bar{G}$ the largest connected component of G.

Figure 3.9: A graphical representation of G and $\bar{G}$

Is the Prisoner's Dilemma a collaborative field?

Based on Table 3.4 an author in G has on average 4 collaborators and a 70% probability of collaborating with a collaborator's co-author. An author of $\bar{G}$ on average is 7% more likely to write with a collaborator's co-author and on average has 2 more collaborators. Moreover, there are only 3.2 % of authors in the PD that has no connection to any other author.

How does this compare to other fields? Two more data sets for the topics "Price of Anarchy" and "Auction Games" have been collected in order to compare the collaborative behaviour of the PD to other game theoretic fields. A total of 3444 publications have been collected for Auction games and 748 for Price of Anarchy. Price of Anarchy is relatively a new field, with the first publication on the topic being **Koutsoupias1999** in 1999. This explains the small number of articles that have been retrieved. Both data sets have been archived and are available in **auction˙data˙2018**; **anarchy˙data˙2018**. The networks for both data sets have been generated in the same way as G. A summary of the networks' metrics are given by Table 3.5.

The average degrees for the Price of Anarchy and for Auction games are lower than the PD's. In Auction games an author is more likely to have no collaborators, and in the Price of Anarchy there are almost no authors that are not connected to someone. This could be an effect of the field being introduced in more modern days. Overall, an author in the PD has on average more collaborators and there are less isolated authors compared to another well established game theoretic field. These results seem to indicate that the PD is a *relatively* collaborative field.

However, both G and $\bar{G}$ have a high modularity (larger than 0.84) and a large number of communities (967 and 25 respectively). A high modularity implies that authors create their own publishing communities but not many publications from authors from different communities occur. Thus, author tends to collaborate with authors in their communities but not many efforts are made to create new connections to other communities and spread the knowledge of the field across academic teams. The fields of both Price of Anarchy and Auction games also have high modularity, and that could indicate that is in fact how academic publications are.

Thus, **the PD is indeed a collaborative field but perhaps it is not more collaborative than other fields**, as there is no effort from the authors to write with people outside their community.

	# Nodes	# Edges	% Isolated nodes	# Connected components	Size of largest component	Av. degree	# Communities	Modularity	Clustering coeff
G	4011	7642	3.2	947	796	3.811	967	0.96491	0.701
$\bar{G}$	796	2214	0.0	1	796	5.563	25	0.84406	0.773

Table 3.4: Network metrics for G and $\bar{G}$ respectively.

	# Nodes	# Edges	# Isolated nodes	% Isolated nodes	# Connected components	Size of largest component	Av. degree	# Communities	Modularity	Clustering coeff
Auction Games	5165	7861	256	5.0	1272	1348	3.044	1294	0.957	0.622
Price of Anarchy	1155	1953	4	0.3	245	222	3.382	253	0.965	0.712

Table 3.5: Network metrics for auction games and price of anarchy networks respectively.

The evolution of the networks was also explored over time by constructing the network cumulatively over 51 periods. Except from the first period 1951-1966 the rest of the periods have a yearly interval (data for the years 1975 and 1982 were not retrieved by the collection data process). The metrics of each sub network are given in the Appendix .1.

The results, similarly to the results of **Liu2015**, confirm that the networks grow over time and that the networks always had a high modularity. Since the first publications authors tend to write with people from their communities, and that is not an effect of a specific time period.

Are some topics more collaborative than other?

The networks corresponding to the topics of Section 3.3 have also been generated similarly to G. Note that authors with publications in more than one topic exist, and these authors are included in all the corresponding networks. A metrics' summary for all five topic networks is given by Table 3.6.

Topic B is the network with the highest average degree followed by Topic A. The topic with the smallest average degree, 2.5, is Topic C. In topics A and B the number of isolated nodes is very small $lessthan(0.2)$ compared to Topic E where the percentage of isolated nodes is approximately 6%. Moreover, in topics C and E an author is 10% more likely to collaborate with a collaborator's co-author. Thus, **topics "human subject research" and "biological studies" tend to be more collaborative than the topic of "strategies", and an authors in these are less likely to have at least one collaborator compared to the topic of "modeling problems as a PD"**.

"Evolutionary dynamics on networks" also appear to be a collaborative topic. In fact the network of the topic is a sub graph of $\bar{G}$, the main cluster of G and it will be demonstrated in the following section that authors in this network are more like to gain from the influence of the network compared to any other topic network.

Are there authors which benefit more from their position in the network?

There are two centrality measures reported in this work, closeness and betweenness centrality. Closeness centrality is a measure of how easy it is for an author to contact others, and consequently affect them; influence them. Thus closeness centrality here

	# Nodes	# Edges	# Isolated nodes	% Isolated nodes	# Connected components	Size of largest component	Av. degree	# Communities	Modularity	Clustering coeff
Topic A	1124	2137	15	1.3	264	56	3.802	265	0.983	0.759
Topic B	695	1382	13	1.9	157	80	3.977	158	0.950	0.773
Topic C	900	1141	41	4.6	281	29	2.536	281	0.981	0.636
Topic D	880	1509	17	1.9	174	312	3.430	183	0.918	0.701
Topic E	1045	1964	59	5.6	354	31	3.759	354	0.926	0.664

Table 3.6: Network metrics for topic networks.

is a measure of influence. Betweenness centrality is a measure of how many paths pass through a specific node, thus the amount of information this person has access to. Betweenness centrality is used here as a measure of how much an author gains from the field. All centrality measure can have values ranging from 0 to 1. The influence and the amount of information an author has access to are used to explore which authors benefit more from their position.

For G and $\bar{G}$ the most central authors based on closeness and betweenness centralities are given by Table 3.7. The most central authors in G and $\bar{G}$ are the same. This implies that the results on centrality heavily rely on the main cluster (as expected). Matjaz Perc is an author with 83 publications in the data set and the most central authors based on both centrality measures. The most central authors are fairly similar between the two measures. The author that appear to be central based on one measure and not the other are Martin Nowak, Franz Weissing, Jianye Hao, Angel Sanchez and Valerio Capraro which have access to information due to their positioning but do not influence the network as much, and the opposite is true for Attila Szolnoki, Luo-Luo Jiang Sandro Meloni, Cheng-Yi Xia and Xiaojie Chen.

It is obvious that in G the centralities values are low which suggests that in the PD authors do not benefit from their positions. This could be an effect of information not flowing from one community to another as authors tend to write with people from their communities. Nevertheless, **there are authors that do benefit from their position, but these are only the authors connected to the main cluster**.

	G				$\bar{G}$			
	Name	Betweenness	Name	Closeness	Name	Betweenness	Name	Closeness
1	Matjaz Perc	0.015	Matjaz Perc	0.066	Matjaz Perc	0.373	Matjaz Perc	0.330
2	Zhen Wang	0.011	Long Wang	0.060	Zhen Wang	0.279	Long Wang	0.301
3	Long Wang	0.007	Yamir Moreno	0.059	Long Wang	0.170	Yamir Moreno	0.299
4	Martin Nowak	0.006	Attila Szolnoki	0.059	Martin Nowak	0.159	Attila Szolnoki	0.297
5	Angel Sanchez	0.004	Zhen Wang	0.059	Angel Sanchez	0.114	Zhen Wang	0.296
6	Yamir Moreno	0.004	Arne Traulsen	0.056	Yamir Moreno	0.110	Arne Traulsen	0.281
7	Arne Traulsen	0.004	Luo-Luo Jiang	0.055	Arne Traulsen	0.107	Luo-Luo Jiang	0.280
8	Franz Weissing	0.004	Sandro Meloni	0.055	Franz Weissing	0.101	Sandro Meloni	0.278
9	Jianye Hao	0.004	Cheng-Yi Xia	0.055	Jianye Hao	0.094	Cheng-Yi Xia	0.276
10	Valerio Capraro	0.004	Xiaojie Chen	0.055	Valerio Capraro	0.093	Xiaojie Chen	0.276

Table 3.7: 10 most central authors based on betweenness and closeness centralities for G and $\bar{G}$.

The centrality measures for the topic networks have also been estimated and are given in Tables 3.8-3.9. If information was flowing between the communities of the research topics then there would be an increase to the values of centralities for the sub networks. However, the only topic where authors gain from their positions are the authors of Topic D (topic on evolutionary dynamics on network). From the list of names it is obvious that these authors are part of $\bar{G}$, and that the network of Topic D is a sub network of $\bar{G}$. This confirms the results. The people benefiting from their position in the co-authorship networks corresponding to research topics of the PD are only the people from the main cluster of G.

The fact that most authors of the main cluster are primarily publishing in evolutionary dynamics on networks indicates that publishing in this specific topic differs from the other topics covered in this manuscript. There appears to be more collaboration and influence in the publications on evolutionary dynamics and authors are more likely to gain from their position, though it is not clear as to why.

	Topic A		Topic B		Topic C		Topic D		Topic E	
	Name	Betweeness	Name	Betweeness	Name	Betweeness	Name	Betweeness	Name	Betweeness
1	David Rand	0.002	Long Wang	0.006	Daniel Ashlock	0.001	Matjaz Perc	0.064	Zengru Di	0.0
2	Valerio Capraro	0.001	Luo-Luo Jiang	0.005	Matjaz Perc	0.000	Luo-Luo Jiang	0.037	Jian Yang	0.0
3	Angel Sanchez	0.001	Martin Nowak	0.004	Karl Tuyls	0.000	Yamir Moreno	0.031	Yevgeniy Vorobeychik	0.0
4	Feng Fu	0.001	Matjaz Perc	0.003	Philip Hingston	0.000	Christoph Hauert	0.027	Otavio Teixeira	0.0
5	Martin Nowak	0.000	Attila Szolnoki	0.003	Eun-Youn Kim	0.000	Long Wang	0.024	Roberto Oliveira	0.0
6	Nicholas Christakis	0.000	Christian Hilbe	0.002	Wendy Ashlock	0.000	Zhen Wang	0.024	M. Nowak	0.0
7	Pablo Branas-Garza	0.000	Yamir Moreno	0.002	Attila Szolnoki	0.000	Han-Xin Yang	0.023	M. Harper	0.0
8	Toshio Yamagishi	0.000	Xiaojie Chen	0.002	Seung Baek	0.000	Martin Nowak	0.020	Xiao Han	0.0
9	James Fowler	0.000	Arne Traulsen	0.002	Martin Nowak	0.000	Angel Sanchez	0.017	Zhesi Shen	0.0
10	Long Wang	0.000	Zhen Wang	0.002	Thore Graepel	0.000	Zhihai Rong	0.016	Wen-Xu Wang	0.0

Table 3.8: 10 most central authors based on betweenness centrality for topics' networks.

	Topic A		Topic B		Topic C		Topic D		Topic E	
	Name	Closeness	Name	Closeness	Name	Closeness	Name	Closeness	Name	Closeness
1	David Rand	0.027	Long Wang	0.043	Karl Tuyls	0.022	Matjaz Perc	0.123	Stefanie Widder	0.029
2	Valerio Capraro	0.023	Matjaz Perc	0.041	Thore Graepel	0.019	Zhen Wang	0.109	Rosalind Allen	0.029
3	Jillian Jordan	0.022	Attila Szolnoki	0.040	Joel Leibo	0.018	Long Wang	0.107	Thomas Pfeiffer	0.029
4	Nicholas Christakis	0.021	Martin Nowak	0.040	Edward Hughes	0.017	Yamir Moreno	0.105	Thomas Curtis	0.029
5	James Fowler	0.020	Olivier Tenaillon	0.038	Matthew Phillips	0.017	Luo-Luo Jiang	0.104	Carsten Wiuf	0.029
6	Martin Nowak	0.020	Xiaojie Chen	0.038	Edgar Duenez-Guzman	0.017	Attila Szolnoki	0.103	William Sloan	0.029
7	Angel Sanchez	0.019	Bin Wu	0.038	Antonio Castaneda	0.017	Gyorgy Szabo	0.102	Otto Cordero	0.029
8	Gordon Kraft-Todd	0.019	Yanling Zhang	0.037	Iain Dunning	0.017	Xiaojie Chen	0.102	Sam Brown	0.029
9	Akihiro Nishi	0.019	Feng Fu	0.037	Tina Zhu	0.017	Guangming Xie	0.101	Babak Momeni	0.029
10	Anthony Evans	0.019	David Rand	0.037	Kevin Mckee	0.017	Lucas Wardil	0.101	Wenying Shou	0.029

Table 3.9: 10 most central authors based on closeness centrality for topics' networks.

The distributions of both centrality measures for all the networks of this work are given in the Appendix .2.2.

3.6 Conclusion

This manuscript has explored the research topics in the publications of the Iterated Prisoner's Dilemma, and moreover, the authors' collaborative behaviour and their influence in the research field. This was achieved by applying network theoretic approaches and a LDA algorithm to a total of 2422 publications. Both the software **nikoleta˙2017** and the data **nikoleta˙2017** have been archived and are available to be used by other researchers. In fact **nikoleta˙2017** has been used by **brane** and **arcas˙blog**.

The data collection and an introduction to the methodology used in this work were covered in Section 3.2. Section 3.3 covered an initial analysis of the data set which demonstrated that the PD is a field that continues to attract academic attention and publications. In Section 3.4 LDA was applied to the data set to identify topics on which researchers have been publishing. The LDA analysis showed that the data could be classified into 5 topics associated with human subject research, biological studies, strategies, evolutionary dynamics on networks and modeling problems as a PD. These topics summarize the field of the PD well, as they demonstrate its interdisciplinarity and applications to a variety of problems. A temporal analysis explored how relevant these topics have been over the course of time, and it revealed that even though there were not the necessarily always the most discussed topics they were still being explored by researchers.

The collaborative behaviour of the field was explored in Section 3.5 by constructing the co authorship network. It was concluded that the field is a collaborative field, where authors are likely to write with a collaborator's co-authors and on average an author has 4 co-authors, however it not necessarily more collaborative than other fields. The authors tend to collaborate with authors from one community, but not many authors are involved in multiple communities. This however might be an effect of academic research, and it might not be true just for the field of the PD. Exploring the influence of authors and their gain from being in the network of the field demonstrated that authors do not gain much, and the authors with influence are only the ones connected to the main cluster, to a "main" group of authors. This 'main" group of authors consists of authors publishing in evolutionary dynamics on networks. Thus, an author would be aiming to publish on this topic if they were interested in gaining from their position in the publications of the PD.

The study of the PD is the study of cooperation and investigating the cooperative behaviours of authors is what this work has aimed to achieve. Interesting areas of future work would include extending this analysis to more game theoretic sub fields, to evaluate whether the results remain the same.

3.7 Cumulative Networks Metrics

3.7.1 Collaborativeness metrics for cumulative graphs, $\tilde{G} \subseteq G$

Period	# Nodes	# Edges	# Isolated nodes	% Isolated nodes	# Connected components	Size of largest component	Av. degree	# Communities	Modularity	Clustering coeff
1951 - 1966	6	3	0	0.0	3	2	1.000	3	0.667	0.000
1951 - 1967	8	4	0	0.0	4	2	1.000	4	0.750	0.000
1951 - 1968	19	15	0	0.0	8	5	1.579	8	0.684	0.228
1951 - 1969	20	17	0	0.0	8	6	1.700	8	0.630	0.250
1951 - 1970	22	18	0	0.0	9	6	1.636	9	0.667	0.227
1951 - 1971	33	28	0	0.0	13	6	1.697	13	0.827	0.424
1951 - 1972	39	34	0	0.0	15	6	1.744	15	0.867	0.513
1951 - 1973	42	35	1	2.4	17	6	1.667	17	0.873	0.476
1951 - 1974	42	35	1	2.4	17	6	1.667	17	0.873	0.476
1951 - 1976	42	35	1	2.4	17	6	1.667	17	0.873	0.476
1951 - 1977	44	36	1	2.3	18	6	1.636	18	0.880	0.455
1951 - 1978	44	36	1	2.3	18	6	1.636	18	0.880	0.455
1951 - 1979	47	40	1	2.1	18	6	1.702	18	0.884	0.454
1951 - 1980	47	40	1	2.1	18	6	1.702	18	0.884	0.454
1951 - 1981	50	46	1	2.0	18	6	1.840	18	0.889	0.497
1951 - 1983	51	46	2	3.9	19	6	1.804	19	0.889	0.487
1951 - 1984	53	47	2	3.8	20	6	1.774	20	0.894	0.469
1951 - 1985	53	47	2	3.8	20	6	1.774	20	0.894	0.469
1951 - 1986	53	47	2	3.8	20	6	1.774	20	0.894	0.469
1951 - 1987	56	48	3	5.4	22	6	1.714	22	0.898	0.443
1951 - 1988	62	52	4	6.5	25	6	1.677	25	0.909	0.449
1951 - 1989	75	62	5	6.7	31	6	1.653	31	0.926	0.424
1951 - 1990	79	64	5	6.3	33	6	1.620	33	0.930	0.403
1951 - 1991	87	69	6	6.9	37	6	1.586	37	0.937	0.400
1951 - 1992	95	72	10	10.5	42	6	1.516	42	0.941	0.367
1951 - 1993	106	81	12	11.3	47	6	1.528	47	0.947	0.366
1951 - 1994	124	95	16	12.9	56	6	1.532	56	0.955	0.394
1951 - 1995	135	102	17	12.6	61	6	1.511	61	0.960	0.384
1951 - 1996	142	105	18	12.7	65	6	1.479	65	0.962	0.365
1951 - 1997	155	115	20	12.9	71	6	1.484	71	0.966	0.392
1951 - 1998	191	140	21	11.0	87	6	1.466	87	0.973	0.367
1951 - 1999	221	169	25	11.3	99	6	1.529	99	0.977	0.397
1951 - 2000	250	195	27	10.8	110	6	1.560	110	0.979	0.418
1951 - 2001	287	235	30	10.5	125	7	1.638	125	0.977	0.419
1951 - 2002	335	278	36	10.7	146	7	1.660	146	0.979	0.428
1951 - 2003	381	310	40	10.5	168	7	1.627	168	0.982	0.413
1951 - 2004	437	370	40	9.2	185	10	1.693	185	0.983	0.424
1951 - 2005	532	476	41	7.7	214	19	1.789	214	0.985	0.458
1951 - 2006	640	603	43	6.7	246	22	1.884	246	0.987	0.486
1951 - 2007	793	877	46	5.8	283	25	2.212	283	0.985	0.532
1951 - 2008	948	1170	50	5.3	318	33	2.468	319	0.985	0.558
1951 - 2009	1108	1442	54	4.9	356	71	2.603	358	0.982	0.573
1951 - 2010	1300	1936	66	5.1	402	133	2.978	405	0.965	0.592
1951 - 2011	1560	2375	79	5.1	472	157	3.045	475	0.970	0.613
1951 - 2012	1837	2865	80	4.4	534	209	3.119	537	0.969	0.634
1951 - 2013	2149	3420	93	4.3	603	322	3.183	609	0.965	0.644
1951 - 2014	2481	3971	103	4.2	683	399	3.201	694	0.962	0.658
1951 - 2015	2938	4877	110	3.7	765	504	3.320	779	0.965	0.675
1951 - 2016	3469	6532	114	3.3	850	613	3.766	863	0.964	0.696
1951 - 2017	3735	7072	119	3.2	895	706	3.787	912	0.964	0.700
1951 - 2018	4011	7642	128	3.2	947	796	3.811	967	0.966	0.701

3.7.2 Collaborativeness metrics for cumulative graphs' main clusters, $\tilde{G} \subseteq \bar{G}$

Periods	# Nodes	# Edges	# Isolated nodes	% Isolated nodes	# Connected components	Size of largest component	Av. degree	# Communities	Modularity	Clustering coeff
1951 - 1966	2	1	0	0.0	1	2	1.000	1	0.000	0.000
1951 - 1967	2	1	0	0.0	1	2	1.000	1	0.000	0.000
1951 - 1968	5	8	0	0.0	1	5	3.200	1	0.000	0.867
1951 - 1969	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1970	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1971	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1972	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1973	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1974	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1976	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1977	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1978	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1979	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1980	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1981	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1983	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1984	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1985	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1986	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1987	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1988	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1989	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1990	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1991	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1992	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1993	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1994	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1995	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1996	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1997	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1998	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 1999	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 2000	6	10	0	0.0	1	6	3.333	2	0.020	0.833
1951 - 2001	7	21	0	0.0	1	7	6.000	1	0.000	1.000
1951 - 2002	7	21	0	0.0	1	7	6.000	1	0.000	1.000
1951 - 2003	7	21	0	0.0	1	7	6.000	1	0.000	1.000
1951 - 2004	10	13	0	0.0	1	10	2.600	2	0.376	0.553
1951 - 2005	19	28	0	0.0	1	19	2.947	3	0.544	0.730
1951 - 2006	22	35	0	0.0	1	22	3.182	4	0.527	0.720
1951 - 2007	25	39	0	0.0	1	25	3.120	5	0.558	0.686
1951 - 2008	33	62	0	0.0	1	33	3.758	4	0.623	0.736
1951 - 2009	71	148	0	0.0	1	71	4.169	6	0.697	0.698
1951 - 2010	133	387	0	0.0	1	133	5.820	7	0.726	0.749
1951 - 2011	157	465	0	0.0	1	157	5.924	8	0.727	0.725
1951 - 2012	209	611	0	0.0	1	209	5.847	11	0.733	0.737
1951 - 2013	322	892	0	0.0	1	322	5.540	12	0.780	0.743
1951 - 2014	399	1109	0	0.0	1	399	5.559	15	0.794	0.742
1951 - 2015	504	1368	0	0.0	1	504	5.429	24	0.811	0.751
1951 - 2016	613	1677	0	0.0	1	613	5.471	21	0.819	0.761
1951 - 2017	706	1935	0	0.0	1	706	5.482	29	0.830	0.772
1951 - 2018	796	2214	0	0.0	1	796	5.563	25	0.845	0.773

3.8 Centrality Measures Distributions

3.8.1 Distrubutions for G and $\bar{G}$

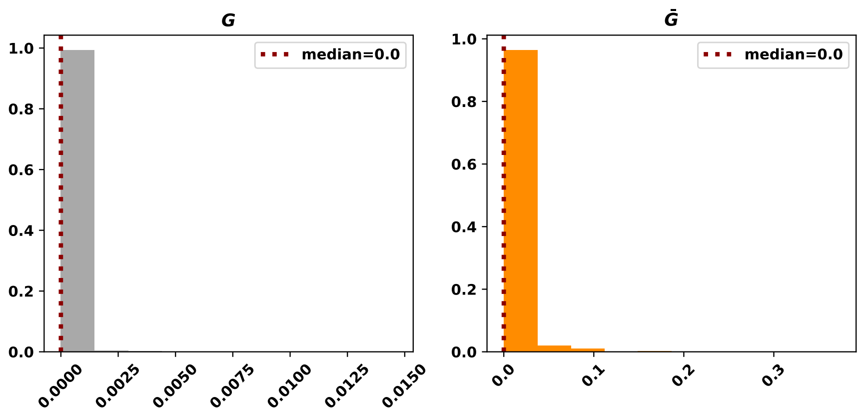


Figure 3.10: Distributions of betweenness centrality in G and $\bar{G}$

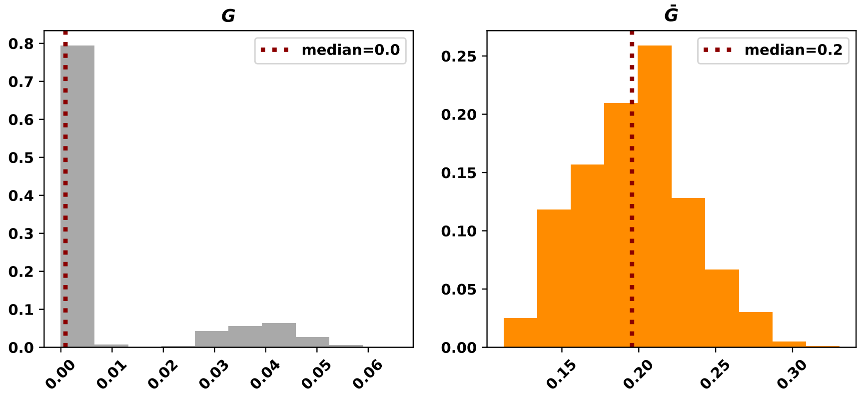


Figure 3.11: Distributions of closeness centrality in G and $\bar{G}$

3.8.2 Distrubutions for Topic Networks

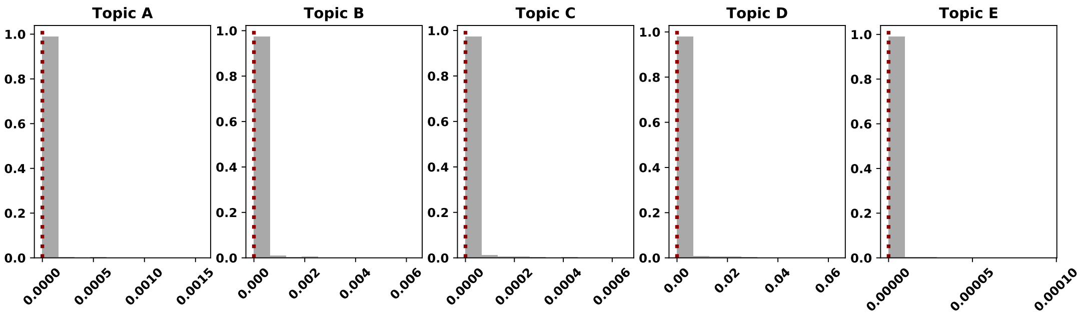


Figure 3.12: Distributions of betweenness centrality in topics' networks.

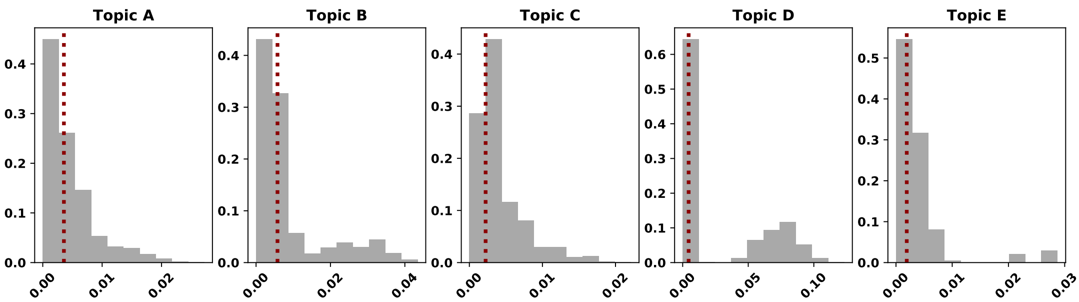


Figure 3.13: Distributions of closeness centrality in topics' networks.

Chapter 4

A meta analysis of tournaments and an evaluation of performance in the Iterated Prisoner's Dilemma.

The Iterated Prisoner's Dilemma has been used for decades as a model of behavioural interactions. From the celebrated performance of Tit for Tat, to the introduction of the zero-determinant strategies, to the use of sophisticated structures such as neural networks, the literature has been exploring the performance of strategies in the game for years. The results of the literature, however, have been relying on the performance of specific strategies in a finite number of tournaments, whereas this manuscript evaluates 195 strategies' effectiveness in more than 40000 tournaments. The top ranked strategies are presented, and moreover, the impact of features on their success are analysed using machine learning techniques. The analysis determines that the cooperating ratio of a strategy in a given tournament compared to the mean and median cooperator is the most important feature. The conclusions are distinct for different types of tournaments. For instance a strategy with a theory of mind would aim to be the mean/median cooperator in standard tournaments, whereas in tournaments with probabilistic ending it would aim to cooperate 10% of the times the median cooperator did.

4.1 Background

The Iterated Prisoner's Dilemma (IPD) is a repeated two player game that models behavioural interactions, and more specifically, interactions where self-interest clashes with collective interest. At each turn of the game both players, simultaneously and independently, decide between cooperation (C) and defection (D) whilst having memory of their prior interactions. The payoffs for each player, at each turn, is influenced by their own choice and the choice of the other player. The payoffs of the game are generally defined by:

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix}$$

where $T > R > P > S$ and $2R > T + S$. The most common values used in the literature **Axelrod1981** are $R = 3, P = 1, T = 5, S = 0$. These values are also used in this work.

Conceptualising strategies and understanding the best way of playing the game has been of interest to the scientific community since the formulation of the game in 1950 **Flood1958**. Following the computer tournaments of Axelrod in the 1980's **Axelrod1980a**; **Axelrod1980b**, a strategy's performance in a round robin computer tournament became a common evaluation technique for newly designed strategies. Today more than 200 strategies exist in the literature and several tournaments, excluding Axelrod's, have been undertaken **Bendor1991**; **Harper2017**; **Kendall2007**; **Stephens2002**; **Stewart2012**.

In the 80's, Axelrod performed two computer tournaments **Axelrod1980a**; **Axelrod1980b**. The contestants were strategies submitted in the form of computer code. They competed against all other entries, a copy of themselves and a random strategy. The winner was decided on the average score a strategy achieved. The winner of both tournaments was the simple strategy Tit For Tat which cooperated on the first turn and then simply copied the previous action of it's opponent. Due to the strategy's strong performance in both tournaments, and moreover in a series of evolutionary experiments **Axelrod1981**, Tit For Tat was thought to be the most robust basic strategy in the IPD.

However, further research proved that the strategy had weakness, and more specifically, it was shown that the strategy suffered in environments with noise **Bendor1991**; **Donninger1986**; **Molander1985**; **Hammerstein1984**. This was mainly due to the strategy's lack of generosity and contrition. The strategy was quick to punish a defection, and in a noisy environment it could lead to a repeated cycle of defections and cooperations. Some new strategies, more robust in tournaments with noise, were soon introduced and became the new protagonists of the game. These include Nice and Forgiving **Bendor1991**, Pavlov **Nowak1993** and Generous Tit For Tat **Nowak1992**.

In 2004, a 20$^{\text{th}}$ Anniversary Iterated Prisoner Dilemma Tournament took place with 233 entries. This time the winning strategy was not designed on a reciprocity based approach but on a mechanism of teams **J.P.Delahaye1993Lp**; **J.P.Delahaye1995LIeP**;

A.Rogers2007Ctpw. A team from Southampton University took advantage of the fact that a participant was allowed to submit multiple strategies. They submitted a total of 60 strategies that could recognised each other and colluded to increase one members score. This resulted with three of the strategies to be ranked in the top spots. The performance of the Southampton University team received mixed attention, though they had won the tournament as stated in **us˙blog** "technically this strategy violates the spirit of the Prisoner's Dilemma, which assumes that the two prisoners cannot communicate with one another".

Another set of IPD strategies that have received a lot of attention are the zero-determinant strategies (ZDs) **Press2012**. By forcing a linear relationship between the payoffs ZDs can ensure that they will never receive less than their opponents. The American Mathematical Society's news section stated that "the world of game theory is currently on fire". ZDs are indeed a set of mathematically unique strategies and robust in pairwise interactions, however, their simplicity and extortionate behaviour have been tested. In **Harper2017** a tournament containing over 200 strategies, including ZDs, was ran and none of them ranked in top spots. Instead, the top ranked strategies were a set of trained strategies based on lookup tables **Axelrod1987**, hidden markov models **Harper2017** and finite state automata **Miller1996**.

Though only a select pieces of work have been discussed, the IPD literature is rich, and new strategies and competitions are being published every year. The question, however, still remains the same: what is the best way to play the game? Compared to other works, whereas a few selected strategies are evaluated on a small number of tournaments, this manuscript evaluates the performance of 195 strategies in 45686 tournaments. These tournaments do not consist by just standard round robin tournaments, but also by tournaments with noise and tournaments with a probabilistic ending. The later part of the paper, evaluates the impact of features on the performance of the strategies using modern machine learning techniques. These features include measures regarding a strategy's behaviour and measures regarding the tournaments. The data set used in this work has been made publicly available **data** and can be used for further analysis and insights.

The different tournament types as well as the data collection, which is made possible due an open source package called Axelrod-Python **axelrodproject**, are covered in Section A.2. Section A.3, focuses on the best performing strategies for each type of tournament and overall. Section A.4, explores the traits which contribute to good performance, and finally the results are summarised in Section A.5. This manuscripts uses several parameters. These are introduced in the following sections, however, the full set of parameters and their definitions are given in Appendix A.6.

4.2 Data collection

For the purposes of this manuscript a data set containing results of IPD tournaments has been generated and is available at **data**. This was done using the open source package Axelrod-Python **axelrodproject**, and more specifically, version 3.0.0. Axelrod-Python allows for different types of IPD computer tournaments to be simulated whilst containing a list of over 180 strategies. Most of these are strategies described in the literature with a few exceptions being strategies that have been contributed specifically to the package. This paper make use of 195 strategies implemented in version 3.0.0. A list of the strategies is given in the Appendix A.7. Though Axelrod-Python features several tournament types, this work considers only standard, noisy, probabilistic ending and noisy probabilistic ending tournaments.

Standard tournaments, are tournaments similar to that of Axelrod's in **Axelrod1980a**. There are N strategies which all play an iterated game of n number of turns against each other. Note that self interactions are not included. Similarly, **noisy tournaments** have N strategies and n number of turns, but at each turn there is a probability p_n that a player's action will be flipped. **Probabilistic ending tournaments**, are of size N and after each turn a match between strategies ends with a given probability p_e. Finally, **noisy probabilistic ending** tournaments have both a noise probability p_n and an ending probability p_e. For smoothing the simulated results a tournament is repeated for k number of times. This was allowed to vary in order to evaluate the effect of smoothing. The winner of each tournament is based on the average score a strategy achieved and not by the number of wins.

The process of collecting tournament results implemented in this manuscript is described by Algorithm 1. For each trial a random size N is selected, and from the 195 strategies a random list of N strategies is chosen. For the given list of strategies a standard, a noisy, a probabilistic ending and a noisy probabilistic ending tournament are performed and repeated k times. The parameters for the tournaments, as well as the number of repetitions, are selected once for each trial. The parameters and their respective minimum and maximum values are given by Table A.1.

parameter	parameter explanation	min value	max value
N	number of strategies	3	195
k	number of repetitions	10	100
n	number of turns	1	200
p_n	probability of flipping action at each turn	0	1
p_e	probability of match ending in the next turn	0	1

Table 4.1: Data collection; parameters' values

The source code for the data collection, as well as the source code for the analysis, which will be discussed in the following sections, have been written following best practices **Aberdour2007**; **Benureau2018**. It has been packaged and is available here.

Algorithm 1: Data collection Algorithm

foreach *seed* $\in [0, 11420]$ **do**
  $N \leftarrow$ randomly select integer $\in [N_{min}, N_{max}]$;
  players $\leftarrow$ randomly select N players;
  $k \leftarrow$ randomly select integer $\in [k_{min}, k_{max}]$;
  $n \leftarrow$ randomly select integer $\in [n_{min}, n_{max}]$;
  $p_n \leftarrow$ randomly select float $\in [p_{n\,min}, p_{n\,max}]$;
  $p_e \leftarrow$ randomly select float $\in [p_{e\,min}, p_{e\,max}]$;

  result standard $\leftarrow$ Axelrod.tournament(players, n, k);
  result noisy $\leftarrow$ Axelrod.tournament(players, n, p_n, k);
  result probabilistic ending $\leftarrow$ Axelrod.tournament(players, p_e, k);
  result noisy probabilistic ending $\leftarrow$ Axelrod.tournament(players, p_n, p_e, k);

return *result standard, result noisy, result probabilistic ending, result noisy probabilistic ending*;

A total of 11420 trials of Algorithm 1 have been run. For each trial the results for 4 different tournaments were collected, thus a total of 45686 (11420×4) tournament results have been retrieved. Each tournament outputs a result summary in the form of Table A.2. Each strategy have participated on average in 5154 tournaments of each type. The strategy with the maximum participation in each tournament type is Inverse Punisher with 5639 entries. The strategy with the minimum entries is EvolvedLookerUp 1 1 1 which was selected in 4693 trials.

The result summary, Table A.2, has N number of rows because each row contains information for each strategy that participated in the tournament. The information includes the strategy's rank, median score, the rate with which the strategy cooperated (C_r), its match win count and the probability that the strategy cooperated in the opening move. Moreover, the probabilities of a strategy being in any of the four states (CC, CD, DC, DD), and the rate of which the strategy cooperated after each state. A measure that has been manually included is the **normalised rank**. The normalised rank, denoted as r, is calculated as a strategy's rank divided by the tournament's size (N). In the next section the performance of these strategies is evaluated based on their normalised rank.

						Rates							
Rank	Name	Median score	Cooperation rating (C_r)	Win	Initial C	CC	CD	DC	DD	CC to C	CD to C	DC to C	DD to C
0	EvolvedLookerUp2 2 2	2.97	0.705	28.0	1.0	0.639	0.066	0.189	0.106	0.836	0.481	0.568	0.8
1	Evolved FSM 16 Noise 05	2.875	0.697	21.0	1.0	0.676	0.020	0.135	0.168	0.985	0.571	0.392	0.07
2	PSO Gambler 1 1 1	2.874	0.684	23.0	1.0	0.651	0.034	0.152	0.164	1.000	0.283	0.000	0.136
3	PSO Gambler Mem1	2.861	0.706	23.0	1.0	0.663	0.042	0.145	0.150	1.000	0.510	0.000	0.122
4	Winner12	2.835	0.682	20.0	1.0	0.651	0.031	0.141	0.177	1.000	0.441	0.000	0.462
...	...	...	...	...	...	...	...	...	...	...	...	...	...

Table 4.2: Output result of a single tournament.

4.3 Top ranked strategies

This section evaluates the performance of 195 IPD strategies. The performance of each strategy is evaluated in four tournament types, which were presented in Section A.2, followed by an evaluation of their performance over all the 45686 simulated tournaments of this work.

Each strategy participated in multiple tournaments of the same type (on average 5154). For example Tit For Tat has participated in a total of 5114 tournaments of each type. The strategy's normalised rank distribution in these is given in Figure A.1. A value of $r = 0$ corresponds to a strategy winning the tournament where a value of $r = 1$ corresponds to the strategy coming last. Because of the strategies' multiple entries their performance is evaluated based on the **median normalised rank** denoted as $\bar{r}$.

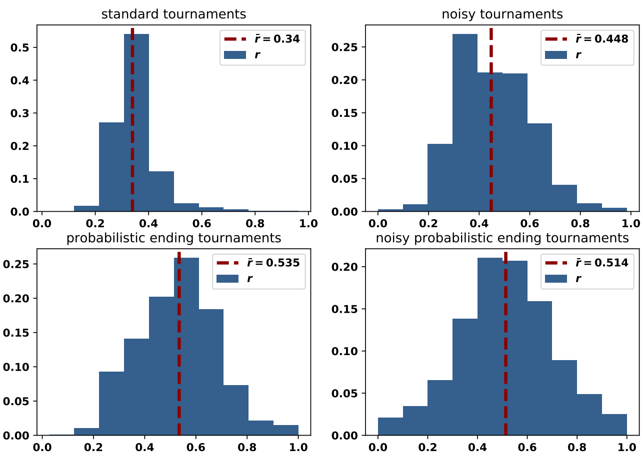


Figure 4.1: Tit For Tat's r distribution in tournaments. The best performance of the strategy has been in standard tournaments where it achieved a $\bar{r}$ of 0.34.

The top 15 strategies for each tournament type based on $\bar{r}$ are given in Table A.3.

In standard tournaments 10 out of the 15 top strategies are introduced in **Harper2017**. These are strategies based on finite state automata (FSM), hidden markov models (HMM), artificial neural networks (ANN), lookup tables (LookerUp) and stochastic lookup tables (Gambler) that have been trained using reinforcement learning algorithms (evolutionary and particle swarm algorithms). They have been trained to perform

	Standard		Noisy		Probabilistic ending		Noisy probabilistic ending	
	Name	$\bar{r}$	Name	$\bar{r}$	Name	$\bar{r}$	Name	$\bar{r}$
0	Evolved HMM 5	0.00667	Grumpy	0.14020	Fortress4	0.01266	Alternator	0.30370
1	Evolved FSM 16	0.00995	e	0.19388	Defector	0.01429	ϕ	0.30978
2	EvolvedLookerUp2 2 2	0.01064	Tit For 2 Tats	0.20617	Better and Better	0.01587	e	0.31250
3	Evolved FSM 16 Noise 05	0.01667	Slow Tit For Two Tats	0.20962	Tricky Defector	0.01875	π	0.31686
4	PSO Gambler 2 2 2	0.02143	Cycle Hunter	0.21538	Fortress3	0.02174	Limited Retaliate	0.35263
5	Evolved ANN	0.02878	Risky QLearner	0.22222	Gradual Killer	0.02532	Anti Tit For Tat	0.35431
6	Evolved ANN 5	0.03390	Retaliate 3	0.22887	Aggravater	0.02778	Retaliate 3	0.35563
7	PSO Gambler 1 1 1	0.03704	Cycler CCCCCD	0.23507	Raider	0.03077	Limited Retaliate 3	0.35563
8	Evolved FSM 4	0.04891	Retaliate 2	0.23913	Cycler DDC	0.04545	Retaliate	0.35714
9	PSO Gambler Mem1	0.05036	Defector Hunter	0.24038	Hard Prober	0.05128	Retaliate 2	0.35767
10	Winner12	0.06011	Retaliate	0.24177	SolutionB1	0.06024	Limited Retaliate 2	0.36134
11	Fool Me Once	0.06140	Hard Tit For 2 Tats	0.25000	Meta Minority	0.06077	Hopeless	0.36842
12	DBS	0.07143	ShortMem	0.25286	Bully	0.06081	Arrogant QLearner	0.40651
13	DoubleCrosser	0.07200	Limited Retaliate 3	0.25316	Fool Me Forever	0.07080	Cautious QLearner	0.40909
14	BackStabber	0.07519	Limited Retaliate	0.25706	EasyGo	0.07101	Fool Me Forever	0.41764

Table 4.3: Top performances for each tournament type based on $\bar{r}$.

well against the strategies in **axelrodproject** in a standard tournament, thus their performance in the specific setting was anticipated. DoubleCrosser, and Fool Me Once, are strategies not from the literature but from **axelrodproject**. DoubleCrosser is a strategy that makes use of the number of turns because is set to defect on the last two rounds. The strategy was expected to not perform as well in tournaments where the number of turns is not specified, but the strategy did not perform well in tournaments with noise either. Finally, Winner 12 **mathieu2017** and DBS **Au2006** are both from the the literature. DBS is strategy specifically designed for noisy environments, however, it ranks highly only in standard ones.

Figure A.2 gives the distributions of r for the top ranked strategies. The distributions are skewed towards zero and the highest median, of the top 15 strategies, is at 0.075. This indicates that the top ranked strategies perform well in any given standard tournament, despite the opponents and the number of turns.

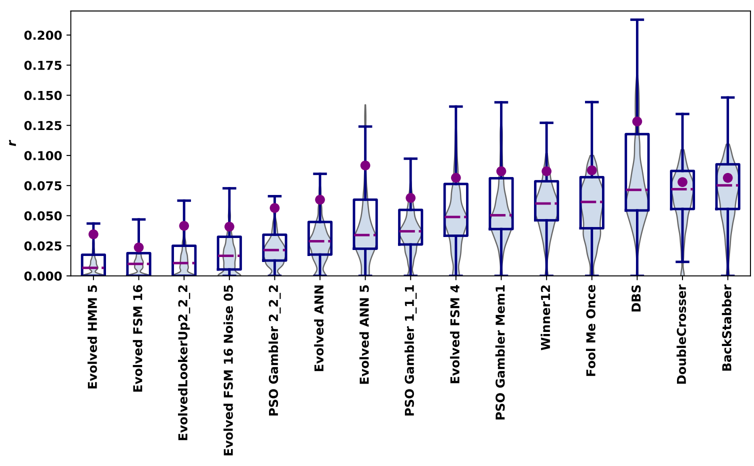


Figure 4.2: r distributions of top 15 strategies in standard tournaments.

The top strategies in noisy tournaments are shown in Figure A.3. These include deterministic strategies, such as Tit For 2 Tats **Axelrod1980b**, Slow Tit For Two Tats **axelrodproject**, Hard Tit For 2 Tats **Stewart2012** and Cycler CCCCCD, and strategies which decide their actions based on the cooperations to defections ratio, such as ShortMem **Carvalho2013**, Grumpy and e **axelrodproject**. Slow Tit For Two Tats is the same strategy as Tit For 2 Tats, and at the time of writing this manuscript the contributors of **axelrodproject** made a new release where the strategy has been removed. However, for the purpose of this work the strategy is kept. The Retaliate and Limited Retaliate strategies are implemented in **axelrodproject** by the same contributor. They are strategies designed to defect if the opponent has tricked them more often than $x\%$ of the times that they have done the same. Finally, in 4^{th} and 9^{th} place are Hunter strategies which trying to extort, equivalently, strategies that play cyclically and defectors.

From Figure A.3, it is evident that the normalised rank distributions in noisy environments are more variant with higher medians compared to standard tournaments. The distributions are bimodal. This indicates that although the top ranked strategies mainly performed well, there are several tournaments that they ranked in the bottom half. To gain a better understanding of the behaviour in noisy tournaments, the r distributions for the top 6 of Figure A.3 strategies over the noise probability p_n, are given in Figure A.4.

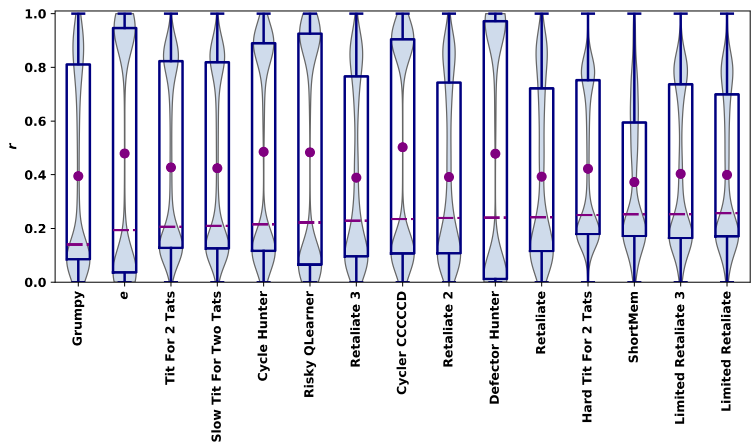


Figure 4.3: r distributions for best performed strategies in noisy tournaments.

Figure A.4 shows that for p_n values lower than 0.5 Grumpy, Tit For 2 Tats and Slow Tit For Two Tat perform moderately, and e, Cycle Hunter and Ricky QLearner perform poorly. At $p_n = 0.5$ all the distributions become bimodal. This is because with a noise probability of 0.5, all strategies correspond to a random player. Interestedly, for a p_n larger than 0.5 all of the 6 strategies become successful. Note that a value $p_n = 1$ corresponds to a strategy playing opposite from what it intended to. Thus, it is demonstrated that the successful strategies is noisy tournaments are sometimes effective when $p_n = 0.5$ but overall they are very successful whn they are playing opposite from their original design. If during the data collection a p_n strictly less 0.5 was considered then the top ranked strategies would be different. There are a total of 5661 trials where $p_n < 0.5$ and the top ranked strategies are given in Table A.4. The median ranks are lower than before and the top spots are mainly overtaken by Meta strategies which include NMWE deterministic and NMWE Long Memory. The Meta strategies **axelrodproject** create a team of strategies for themselves and choose to play as a member of their team based on their scores against a given opponent.

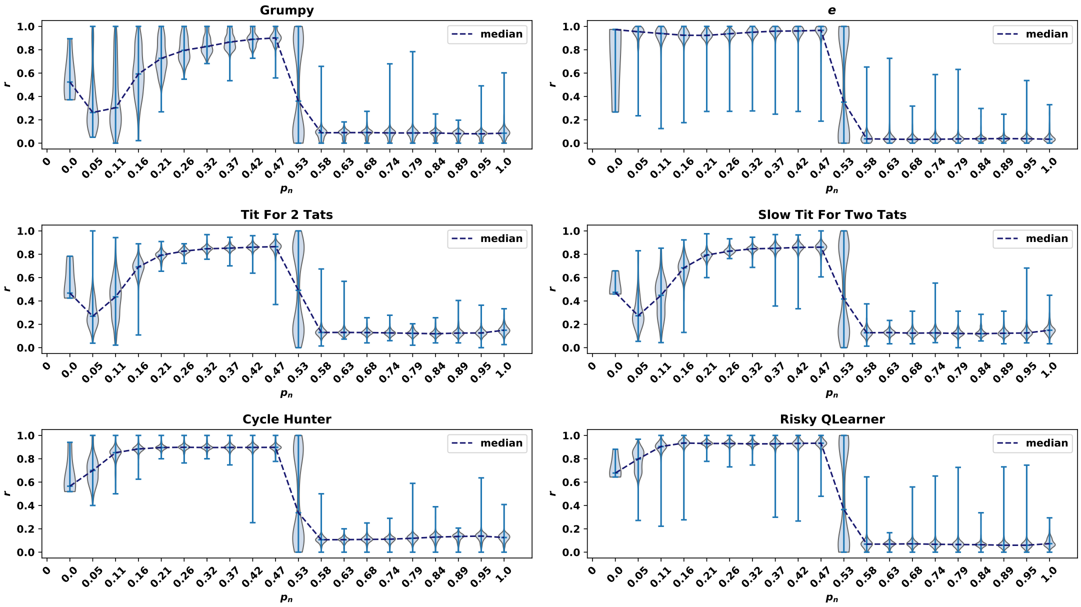


Figure 4.4: r distributions for top 6 strategies in noisy tournaments over the probability of noisy (p_n).

Name	$\bar{r}$
MEM2	0.06135
Spiteful Tit For Tat	0.06344
Nice Meta Winner	0.06620
Grudger	0.06667
Meta Winner Long Memory	0.07339
Forgiver	0.07362
Fool Me Once	0.07362
Meta Winner	0.07487
Meta Winner Memory One	0.07621
Meta Winner Finite Memory	0.07692
Meta Winner Deterministic	0.07792
NMWE Deterministic	0.08696
NMWE Long Memory	0.08696
CollectiveStrategy	0.08696
Defector	0.08889

Table 4.4: Top performances in 5661 noisy tournaments where $p_n < 0.5$.

The 15 top ranked strategies in probabilistic ending tournaments include Fortress 3, Fortress 4 (both introduced in **Ashlock2006**), Raider **Ashlock2014** and Solution B1 **Ashlock2014**, which are strategies based on finite state automata introduced by Daniel and Wendy Ashlock. These strategies have been evolved using reinforcement learning, however, there were trained to maximise their payoffs in tournaments with fixed turns (150 specifically) and not in probabilistic ending ones. In probabilistic ending tournaments it appears that the top ranks are mostly occupied by defecting strategies. These include Better and Better, Gradual Killer, Hard Prober (all from **prison**), Bully (Reverse Tit For Tat) **Nachbar1992** and Defector. Thus, it's surprisingly that EasyGo and Fool Me Forever which are strategies that will defect until their opponent defect, then they will cooperate until the end, ranked 14$^{\text{th}}$ and 15$^{\text{th}}$. Upon inspection, it was found that they are actually the same strategy. This was not known to the authors at the time of data collection. Figure A.5 verifies that their performance is the same. Both strategies have repeatedly ranked highly and there are cases for which they were the winners of the tournament.

The distributions of the normalised rank in probabilistic ending tournaments, shown in Figure A.5, are less variant than those of noisy tournaments. The medians of the top 15 strategies are lower than 0.1 and the distributions are skewed towards 0. Though the large difference between the means and the medians indicates some outliers, the strategies have overall performed well in the probabilistic ending tournaments that they participated.

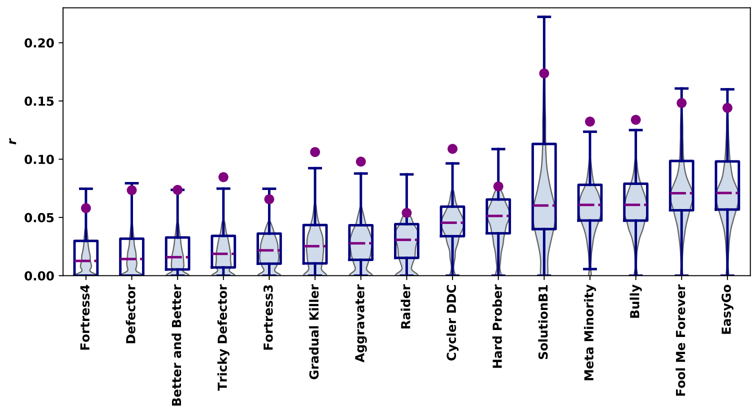


Figure 4.5: r distributions for best performed strategies in probabilistic ending tournaments.

The distributions of r for the top 6 strategies in probabilistic ending tournaments over p_e are given in Figure A.6. Figure A.6 shows that the 6 strategies start of with a high median rank, however, their ranked decreased as the the probability of the game ending increased and at the point of $p_e = 0.1$ they became the dominant strategies in their respective tournaments. In essence, what is demonstrated is that defecting strategies did better when the likelihood of the game ending in the next turn increased, which is inline with the Folk Theorem **Fudenberg2009**. If tournaments where the probability of the game ending was less than 0.1 were considered then the the top ranked spots are not dominated by just defecting strategies anymore, Table A.5. Instead the effective strategies are now the Meta strategies, trained strategies, Grudger **axelrodproject** and Spiteful Tit for Tat **prison**.

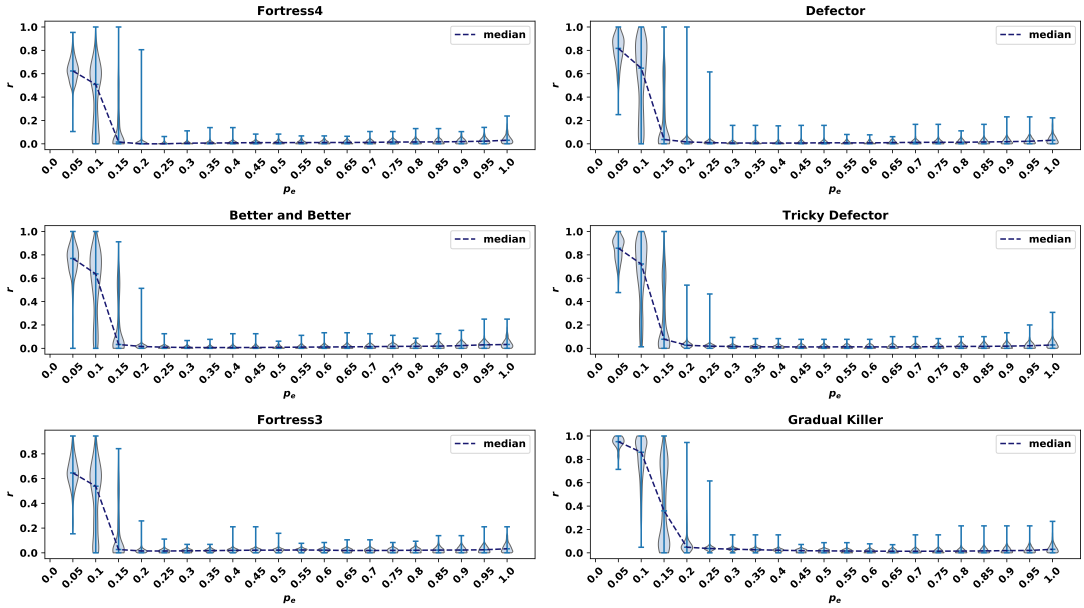


Figure 4.6: r distributions for top 6 strategies in probabilistic ending tournaments over p_e.

Name	$\bar{r}$
Evolved FSM 16	0.00000
Evolved FSM 16 Noise 05	0.01266
MEM2	0.02715
Evolved HMM 5	0.04423
EvolvedLookerUp2 2 2	0.04870
Spiteful Tit For Tat	0.05958
Nice Meta Winner	0.06842
NMWE Finite Memory	0.06923
Grudger	0.06985
NMWE Deterministic	0.07018
NMWE Long Memory	0.07407
Nice Meta Winner Ensemble	0.07595
EvolvedLookerUp1 1 1	0.07692
NMWE Memory One	0.08000
NMWE Stochastic	0.08475

Table 4.5: Top performances in 1139 probabilistic ending tournaments with $p_e < 0.1$

In tournaments with both noise and an unspecified number of turns several of the top ranked strategies are strategies that were highly ranked in noisy tournaments. However, strategies from the top ranks in probabilistic ending tournaments did not rank highly here. Other strategies include π, ϕ which are based on the same approach as e. The distributions of r shown in Figure A.7 have the largest median values compared to the top rank strategies of the other tournament types. A subset of noisy probabilistic ending tournaments has been considered such that $p_e < 0.1$ and $p_n < 0.5$. The top ranked strategies are given in Table A.6 and it is shown that the Meta strategies which performed well in noisy tournaments with $p_n < 0.5$, perform well once again even the number of turns is not specified. Moreover, several strategies that did well in probabilistic ending tournaments such as Fortress 3, Fortress 4, Defector and Better and Better are effective here as well.

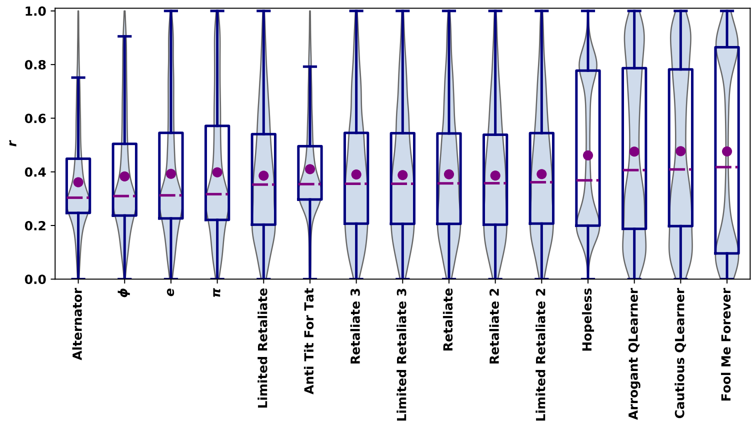


Figure 4.7: r distributions for best performed strategies in noisy probabilistic ending tournaments.

Name	$\bar{r}$
Defector	0.00552
Better and Better	0.01055
Aggravater	0.01399
Fortress4	0.02100
Tricky Defector	0.03857
Meta Winner Long Memory	0.04878
Meta Winner Memory One	0.04955
Meta Winner Finite Memory	0.04972
Meta Winner Stochastic	0.05128
Meta Winner Deterministic	0.05195
Meta Winner	0.05333
Meta Winner Ensemble	0.05882
Fortress3	0.06956
CollectiveStrategy	0.07692
Prober 3	0.08018

Table 4.6: Top performances in 568 probabilistic ending tournaments with $p_e < 0.1$ and $p_n < 0.5$.

Up till now, the performances of the 195 strategies have been evaluated for individual tournament types. The distributions of r for the tournament types indicate that for probabilistic ending and standard tournaments successful strategies do exist. For these settings, the top 15 strategies have frequently ranked in the top spots with only a few exceptions. Contrarily, it appears that noise cause variation in the normalised ranks, and the strategies can always guarantee a spot in the top ranks.

The data set considered in this work, described in Section A.2, contains a total of 45686 tournament results. For this part of the manuscript the strategies are ranked based on the median normalised rank they achieved over the entire data set. The top 15 strategies are given in Table A.7 and their normalised rank distributions are given in Figure A.8.

Name	$\bar{r}$
Limited Retaliate 3	0.28609
Retaliate 3	0.29630
Retaliate 2	0.30250
Limited Retaliate 2	0.30328
Limited Retaliate	0.31000
Retaliate	0.31707
BackStabber	0.32381
DoubleCrosser	0.33136
Nice Meta Winner	0.34921
PSO Gambler 2 2 2 Noise 05	0.35146
Grudger	0.35156
Evolved HMM 5	0.35714
NMWE Memory One	0.35714
Nice Meta Winner Ensemble	0.35870
Forgetful Fool Me Once	0.35884

Table 4.7: Top performances over all the tournaments

The top ranks include strategies that have been previously mentioned. The set of Retaliate strategies occupy the top spots followed by BackStabber and DoubleCrosser. The distributions of the Retaliate strategies have no statistical difference. Thus, in an IPD tournament where the type is not specified, playing as any of the Retaliate strategies will have the result. DoubleCrosser performed well in standard tournaments and the strategy is just an extension of BackStabber. It should be noted that these strategies can be characterised as "cheaters". The source code of the strategies allows them to known the number of turns in a match (if they are specified). PSO Gambler and Evolved HMM 5 are trained strategies introduced in **Harper2017** and Nice Meta Winner and NMWE Memory One are strategies based on teams. Grudger is a strategy from Axelrod's original tournament and Forgetful Fool Me Once is based on the same approach as Grudger. Overall the top 15 strategies are fundamentally different. Some are cheaters, some are complex, others are simple deterministic strategies and strategies based on teams. The results of 45686 tournaments used in this work imply the following: they is not a single type of strategy which can performance well in any IPD interaction.

This section presented the winning strategies in a series of IPD tournaments. In standard tournaments the top spots were dominated by complex strategies that had been

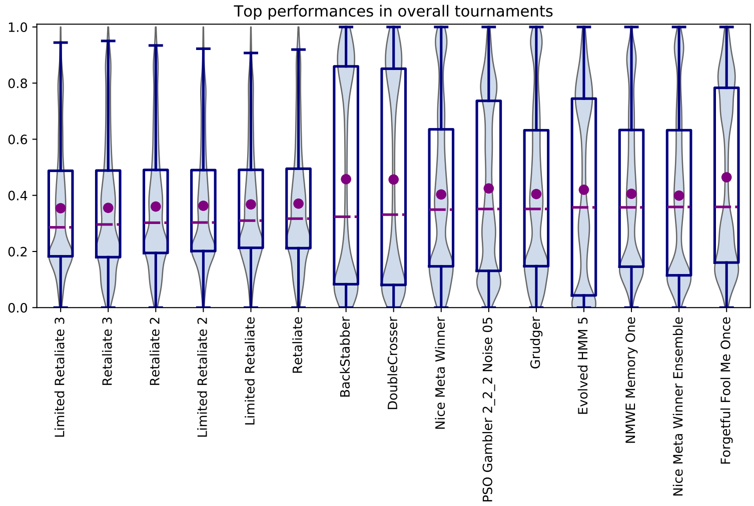


Figure 4.8: r distributions for best performed strategies in the data set **data**.

trained using reinforcement learning techniques. In noisy environments, whether the number of turns was fixed or not, the winning strategies were deterministic strategies designed to defect if the opponent tricked them more than a current amount of the times that they had tricked their opponent. However, if a value of noise strictly less than 0.5 was considered, then the successful strategies were strategies based on teams. In probabilistic ending tournaments most of the winning strategies were defecting strategies and trained finite state automata, designed by the same authors. These strategies only did better when the probability of the game ending after each turn was increased. Finally the performance of all 195 strategies over the 45686 tournaments in this manuscript was assessed on $\bar{r}$. The top ranked strategies were a mixture of behaviours that did well in standard tournaments and tournaments with noise, as well as a few strategies based on teams.

The results of this section imply that successful strategies for specific settings exist for an IPD tournament. The top ranked strategies in both standard tournaments and tournaments with probabilistic ending, managed to rank in the top 10% of the tournament most of the times. Strategies in noisy environments demonstrated that no strategy can be consistently successful, expected if the value of noise is constrained to less than a half. Overall, there has been not a single strategy that has shown to perform well in more than one setting. The aim of the next section is to understand which are the factors that made these strategies successful, in each setting separately but also overall.

4.4 Evaluation of performance

The aim of this section is to explore the factors that contribute to a strategy's successful performance. The factors explored are measures regarding a strategy's behaviour, along with measures regarding the tournaments the strategies competed in. These are given in Table A.8.

measure	measure explanation	source	value type	min value	max value
stochastic	If a strategy is stochastic	strategy classifier from **axelrodproject**	boolean	Na	Na
makes use of game	If a strategy makes used of the game information	strategy classifier from **axelrodproject**	boolean	Na	Na
makes use of length	If a strategy makes used of the number of turns	strategy classifier from **axelrodproject**	boolean	Na	Na
memory usage	The memory size of a strategy divided by the number of turns	memory size from **axelrodproject**	float	0	1
SSE	A measure of how far a strategy is from ZD behaviour	method described in **Knight2019**	float	0	1
max cooperating rate ($C_{\max}$)	The biggest cooperating rate in a given tournament	result summary	float	0	1
min cooperating rate ($C_{\min}$)	The smallest cooperating rate in a given tournament	result summary	float	0	1
median cooperating rate (C_{median})	The median cooperating rate in a given tournament	result summary	float	0	1
mean cooperating rate (C_{mean})	The mean cooperating rate in a given tournament	result summary	float	0	1
C_r / $C_{\max}$	A strategy's cooperating rate divided by the maximum	result summary	float	0	1
C_r / $C_{\min}$	A strategy's cooperating rate divided by the minimum	result summary	float	0	1
C_r / C_{median}	A strategy's cooperating rate divided by the median	result summary	float	0	1
C_r / C_{mean}	A strategy's cooperating rate divided by the mean	result summary	float	0	1
C_r	The cooperating ratio of a strategy	result summary	float	0	1
CC to C rate	The probability a strategy will cooperate after a mutual cooperation	result summary	float	0	1
CD to C rate	The probability a strategy will cooperate after being betrayed by the opponent	result summary	float	0	1
DC to C rate	The probability a strategy will cooperate after betraying the opponent	result summary	float	0	1
DD to C rate	The probability a strategy will cooperate after a mutual defection	result summary	float	0	1
p_n	The probability of a player's action being flip at each interaction	trial summary	float	0	1
n	The number of turns	trial summary	integer	1	200
p_e	The probability of a match ending in the next turn	trial summary	float	0	1
N	The number of strategies in the tournament	trial summary	integer	3	195
k	The number of repetitions of a given tournament	trial summary	integer	10	100

Table 4.8: The measures which are included in the performance evaluation analysis.

Axelrod-Python makes use of classifiers to classify strategies according to various dimensions. These determine whether a strategy is stochastic or deterministic, whether it makes use of the number of turns or the game's payoffs. The memory usage measure is calculated as the memory size of strategy (which is specified in the strategies implementation in **axelrodproject**) divide by the number of turns. For example, Evolved FSM 16 Noise 05 has a memory size of 16 and participated in a tournament where n was 134. In the given tournament Evolved FSM 16 Noise 05 has a memory usage of 0.119. For tournaments with a probabilistic ending the number of turns was not collected, so the memory usage measure is not used for probabilistic ending tournaments. The SSE is a measure introduced in **Knight2019** which shows how close a strategy is to behaving as a ZDs, and subsequently, in an extortionate way. The method identifies the ZDs closest to a given strategy and calculates the algebraic distance between them, defined as SSE. A SSE value of 1 indicates no extortionate behaviour at all whereas a value of 0 indicates that a strategy is behaving a ZDs. The rest of the factors considered are the CC to C, CD to C, DC to C, and DD to C rates as well as cooperating ratio of a strategy. The minimum, maximum, medium and median cooperating ratios of each tournament are also included, and finally the number of turns, the number of strategies, the number of repetitions and the probabilities of noise and the game ending are also included.

Table A.9 shows the correlation coefficients between the measures of Table A.8 the median score and the median normalised rank. Note that the correlation for the classifiers is not included because they are binary variables and they will be evaluated using a different method. The correlation coefficients for all the measures in Table A.8 against themselves have also been calculated and a graphical representation can be found in the Appendix A.8.

	Standard		Noisy		Probabilistic ending		Noisy probabilistic ending		Overall	
	r	median score	r	median score	r	median score	r	median score	r	median score
CC to C rate	-0.501	0.501	0.414	-0.504	0.408	-0.323	0.260	0.022	-0.501	0.501
CD to C rate	0.226	-0.199	0.456	-0.330	0.320	-0.017	0.205	-0.220	0.226	-0.199
C_r	-0.323	0.384	0.711	-0.678	0.714	-0.832	0.579	-0.135	-0.323	0.384
C_r / C_{max}	-0.323	0.381	0.616	-0.551	0.714	-0.833	0.536	-0.116	-0.323	0.381
C_r / C_{mean}	-0.331	0.358	0.731	-0.740	0.721	-0.861	0.649	-0.621	-0.331	0.358
C_r / C_{median}	-0.331	0.353	0.652	-0.669	0.712	-0.852	0.330	-0.466	-0.331	0.353
C_r / C_{min}	0.109	-0.080	-0.358	0.250	-0.134	0.150	-0.368	0.113	0.109	-0.080
C_{max}	-0.000	0.049	0.000	0.023	-0.000	0.046	0.000	-0.004	-0.000	0.049
C_{mean}	-0.000	0.229	-0.000	0.271	0.000	0.200	0.000	0.690	-0.000	0.229
C_{median}	0.000	0.209	-0.000	0.240	-0.000	0.187	-0.000	0.673	0.000	0.209
C_{min}	0.000	0.084	0.000	-0.017	-0.000	0.007	-0.000	0.041	0.000	0.084
DC to C rate	0.127	-0.100	0.509	-0.504	-0.018	0.033	0.341	-0.016	0.127	-0.100
DD to C rate	0.412	-0.396	0.533	-0.436	-0.103	0.176	0.378	-0.263	0.412	-0.396
N	0.000	-0.009	-0.000	0.002	-0.000	0.003	-0.000	0.001	0.000	-0.009
k	0.000	-0.002	-0.000	0.003	-0.000	0.001	-0.000	-0.008	0.000	-0.002
n	0.000	-0.125	-0.000	-0.024	-	-	-	-	0.000	-0.125
p_e	-	-	-	-	0.000	0.165	0.000	-0.058	-0.001	0.001
p_n	-	-	-0.000	0.207	-	-	-0.000	-0.650	0.002	-0.000
Make use of game	-0.003	-0.022	0.025	-0.082	-0.053	-0.108	0.013	-0.016	-0.003	-0.022
Make use of length	-0.158	0.124	0.005	-0.123	-0.025	-0.090	0.014	-0.016	-0.154	0.117
SSE	0.473	-0.452	0.463	-0.337	-0.156	0.223	0.305	-0.259	0.473	-0.452
memory usage	-0.082	0.095	-0.007	-0.017	-	-	-	-	-0.084	0.095
stochastic	0.006	-0.024	0.022	-0.026	0.002	-0.130	0.021	-0.013	0.006	-0.024

Table 4.9: Correlations table between the measures of Table A.8 the normalised rank and the median score.

In standard tournaments the measures CC to C, C_r, C_r/C_{max} and the cooperating ratio compared to C_{median} and C_{mean} have a moderate negative effect on the normalised rank, and a moderate positive on the median score. The SSE error and the DD to C have the opposite effects. Thus, in standard tournaments behaving cooperatively corresponds to a more successful performance. Even though being nice pays off, that's not true against defective strategies. Cooperating after a mutual defection lowers a strategy's success. Figure A.9 confirms that the winners of standard tournaments always cooperate after a mutual cooperation and almost always defects after a mutual defection.

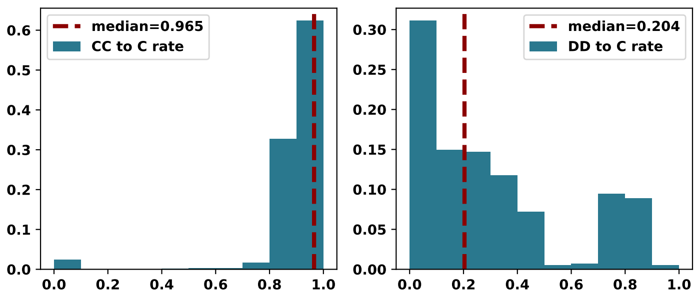


Figure 4.9: Distributions of CC to C and DD to C for the winners in standard tournaments.

Compared to standard tournaments, in both noisy and in probabilistic ending tournaments the higher the rates of cooperation the lower a strategy's success and median score. A strategy would want to cooperate less than both the mean and median cooperator in such settings. In probabilistic ending tournaments the correlations coefficients have a larger values, indicating a stronger effect. Thus a strategy will be punished more by it's cooperative behaviour in probabilistic ending environments, this was seen in Section A.4 as well. The distributions of the C_r of the winners in both tournaments is given by Figure A.10. It confirms that the winners in noisy tournaments cooperated less than 35% of the times and in probabilistic ending tournaments less than 10%. In noisy probabilistic ending tournaments and in over all the tournaments' results, the only measures that had a moderate affect are $C_r/C_{\text{mean}}, C_r/C_{\max}$ and C_r. In such environments cooperative behaviour appears to be punished by not as much as in noisy and probabilistic ending tournaments.

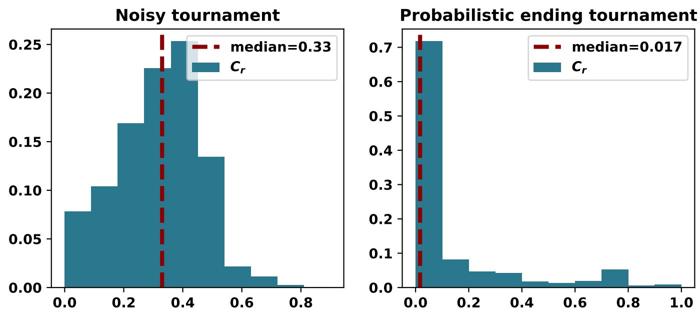


Figure 4.10: C_r distributions of the winners in noisy and in probabilistic ending tournaments.

The performances are clustered based on the normalised rank. More specifically, they are clustered 3 times into 2 different clusters based on on whether their normalised rank was in the top 5%, 25% or 50% respectively. A random forest approach **breiman2001** is then applied to each performance to predict the cluster to which it has been assigned to. The random forest method constructs many individual decision trees and the predictions from all trees are pooled to make the final prediction. The random forest models are trained on a training set of 70% of the tournaments results. The accuracy of each model based on R^2 are given by Table A.10. The out of the bag error **hastie2005** has also been calculated. The models fit well, and a high value of both the accuracy measure on the test data and the OOB error indicate that the model is not over fitting.

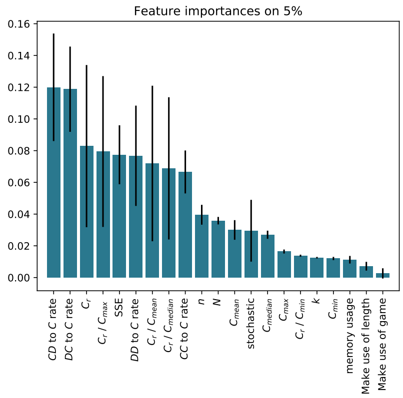
The performances have also been clustered based on their normalised rank and their median score by a $k-$means algorithm **Arthur2007**. The number of clusters is not deterministically chosen but it is based on the silhouette coefficients **Rousseeuw1987**. The chosen cluster for each tournament type, as well as the accuracy for random forest models are also given in Table A.10.

Tournament type	Clustering on	Number of clusters	R^2 training data	R^2 test data	R^2 OOB score
standard	top 5% r	2	0.998831	0.987041	0.983708
	top 25% r	2	0.998643	0.978626	0.969202
	top 50% r	2	0.998417	0.985217	0.976538
	r & normalised score	2	0.998794	0.990677	0.982959
noisy	top 5% r	2	0.996677	0.950572	0.935383
	top 25% r	2	0.996677	0.950572	0.935383
	top 50% r	2	0.996677	0.950572	0.935383
	r & normalised score	3	0.996677	0.950572	0.935383
probabilistic ending	top 5% r	2	0.999592	0.995128	0.992819
	top 25% r	2	0.999592	0.995128	0.992819
	top 50% r	2	0.999592	0.995128	0.992819
	r & normalised score	2	0.999592	0.995128	0.992819
noisy probabilistic ending	top 5% r	2	0.990490	0.813905	0.791418
	top 25% r	2	0.990490	0.813905	0.791418
	top 50% r	2	0.990490	0.813905	0.791418
	r & normalised score	4	0.990490	0.813905	0.791418
over 45686 tournaments	top 5% r	2	0.993396	0.913409	0.898059
	top 25% r	2	0.993396	0.913409	0.898059
	top 50% r	2	0.993396	0.913409	0.898059
	r & normalised score	3	0.993396	0.913409	0.898059

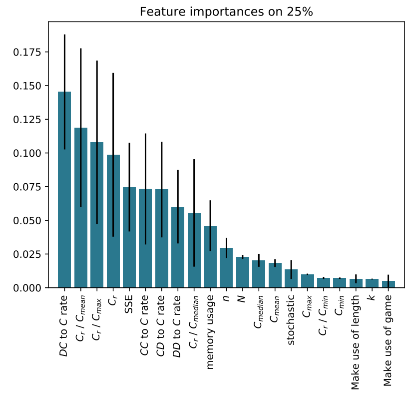
Table 4.10: Accuracy metrics for random forest models.

The importance that the measures of Table A.8 had on each classification task; to which cluster a performance was assigned to based on the normalised rank, and their normalised rank and median score have been calculated and are given by Figures A.11, A.12, A.13, A.14 and A.15. These show that the classifiers stochastic, make use of game and make use of length have no significant effect, and several of the measures that are highted by the importance are inline with the correlation results. Moreover, the smoothing parameter k appears to no have a significant effect either. The most important measures based on the random forest analysis were C_r/C_{median} and C_r/C_{mean}.

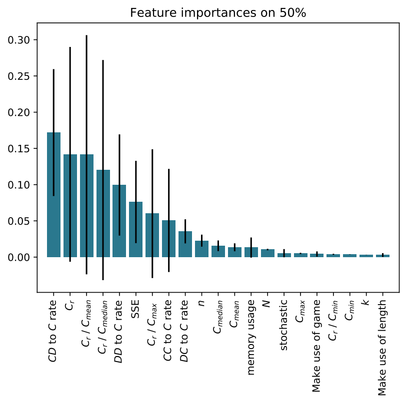
The effect of both these measures can be further explored. In Figure A.16 the distributions of C_r/C_{mean} and C_r/C_{median} are given for the winners in standard tournaments. A value of $C_r/C_{\text{mean}} = 1$ imply that the cooperating ratio of the winner was the same as the mean/median cooperating ratio of the tournament. In standard tournaments, the mean for both ratios is 1. Therefore, an effective strategy in standard tournaments was the mean/median cooperator of its respective tournament. In comparison, Figure A.17 shows the distributions of the measures for the winners in noisy tournaments where the mean is at 0.67. Thereupon the winners cooperated 67% of the times the mean/median cooperator did. This analysis is applied to the rest of the tournaments and the distributions are given by Figures A.18, A.19 and A.20. In a tournament with noisy and a probabilistic ending the winners cooperated 60%, whereas in settings that



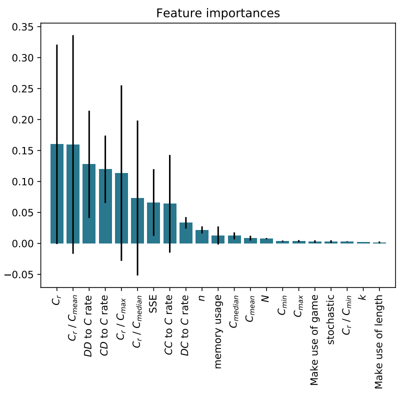
(a) Importance of features for clusters on 5% performance.



(b) Importance of features for clusters on 25% performance.

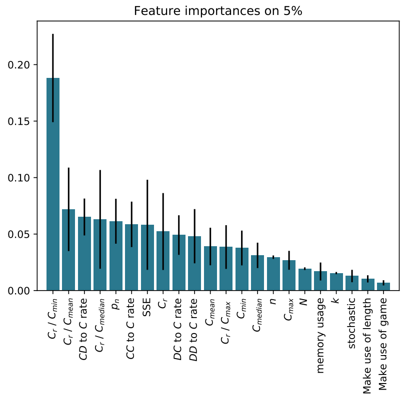


(c) Importance of features for clusters on 50% performance.

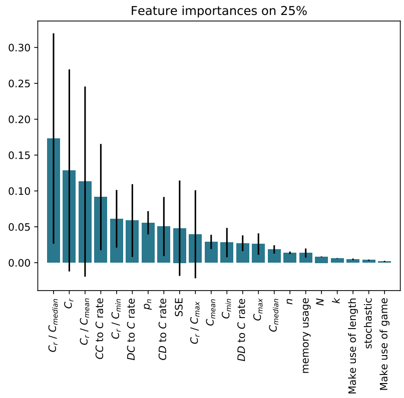


(d) Importance of features for clusters based on $k$means algorithm.

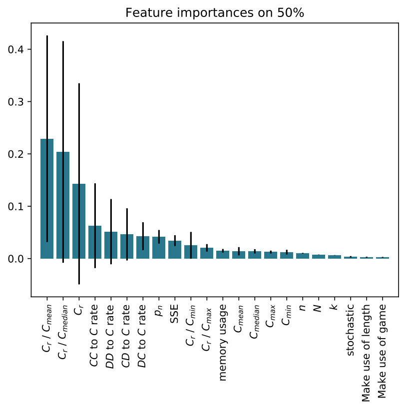
Figure 4.11: Importance of features in standard tournaments for different clustering methods.



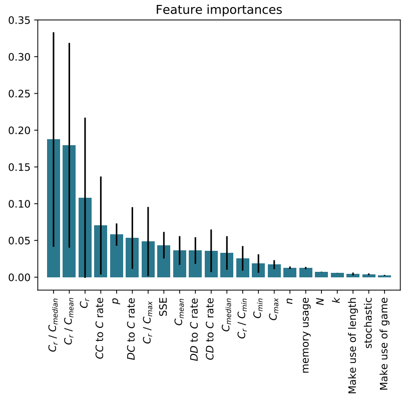
(a) Importance of features for clusters on 5% performance.



(b) Importance of features for clusters on 25% performance.

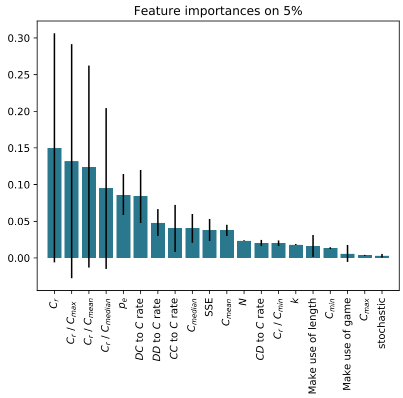


(c) Importance of features for clusters on 50% performance.

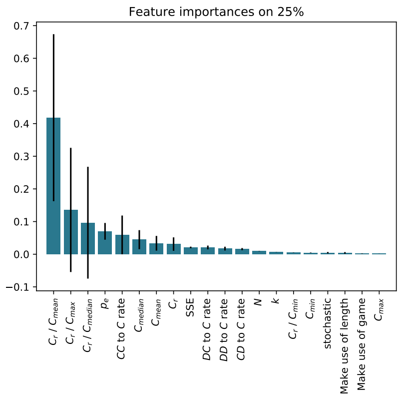


(d) Importance of features for clusters based on $k$means algorithm.

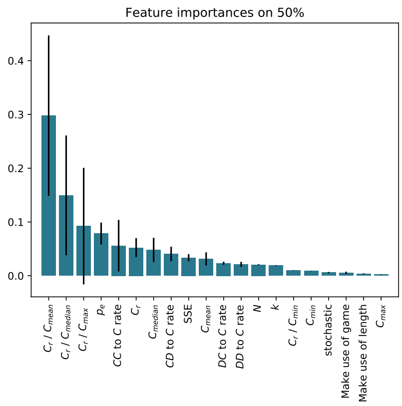
Figure 4.12: Importance of features in noisy tournaments for different clustering methods.



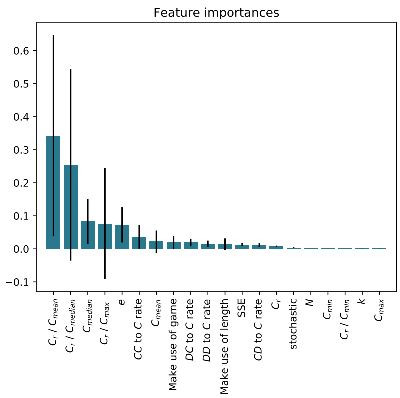
(a) Importance of features for clusters on 5% performance.



(b) Importance of features for clusters on 25% performance.

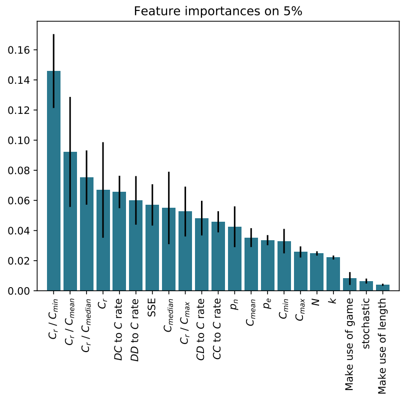


(c) Importance of features for clusters on 50% performance.

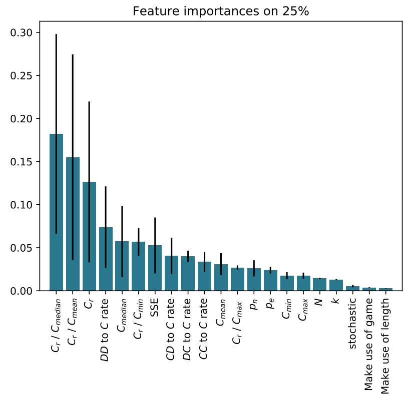


(d) Importance of features for clusters based on $k$means algorithm.

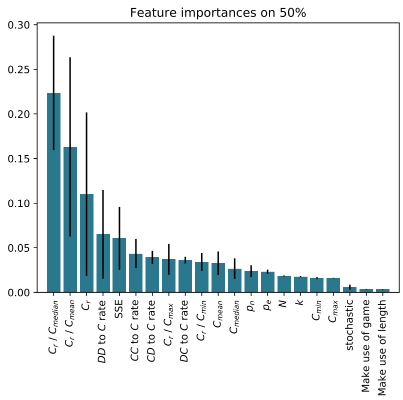
Figure 4.13: Importance of features in probabilistic ending tournaments for different clustering methods.



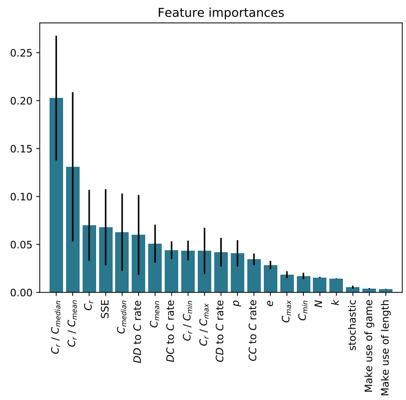
(a) Importance of features for clusters on 5% performance.



(b) Importance of features for clusters on 25% performance.

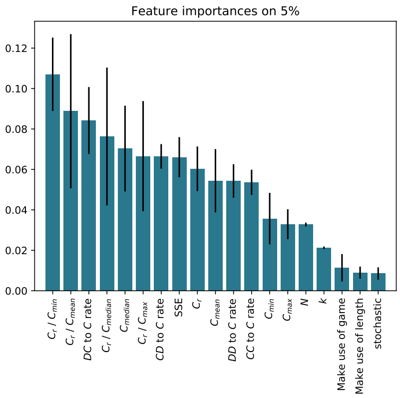


(c) Importance of features for clusters on 50% performance.

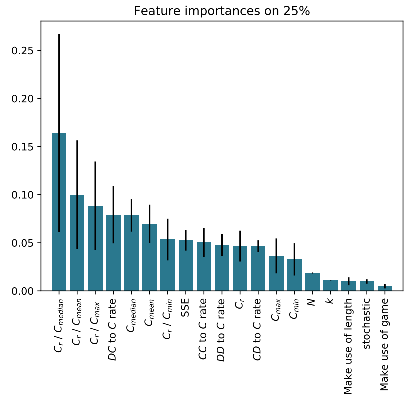


(d) Importance of features for clusters based on $k$means algorithm.

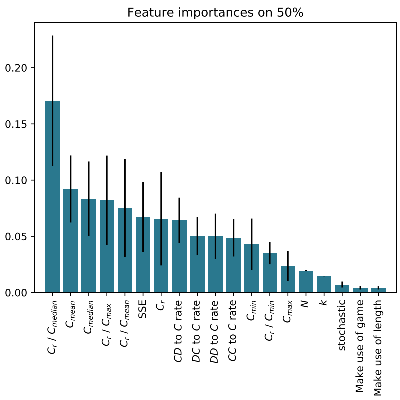
Figure 4.14: Importance of features in noisy probabilistic ending tournaments for different clustering methods.



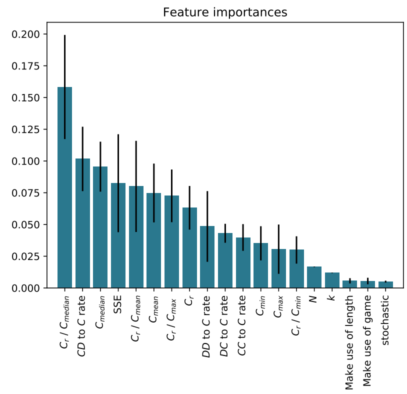
(a) Importance of features for clusters on 5% performance.



(b) Importance of features for clusters on 25% performance.



(c) Importance of features for clusters on 50% performance.



(d) Importance of features for clusters based on $k$means algorithm.

Figure 4.15: Importance of features over all the tournaments for different clustering methods.

the type of the tournament can vary between the types considered in this work the winners cooperated 67% of the times the mean or median cooperator did. Finally, in probabilistic ending tournament it has already been mentioned that defecting strategies prevail and this result is once again confirmed in this section.

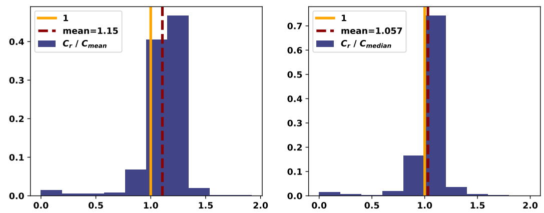


Figure 4.16: Distributions of C_r/C_{median} and C_r/C_{median} for winners of standard tournaments.

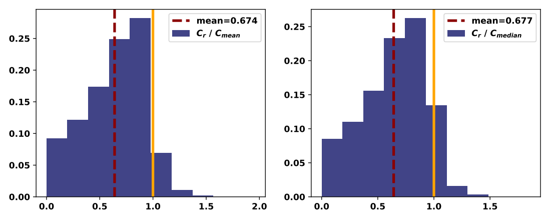


Figure 4.17: Distributions of C_r/C_{median} and C_r/C_{median} for winners of noisy tournaments.

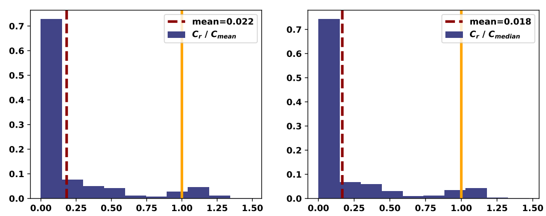


Figure 4.18: Distributions of C_r/C_{median} and C_r/C_{median} for winners of probabilistic ending tournaments.

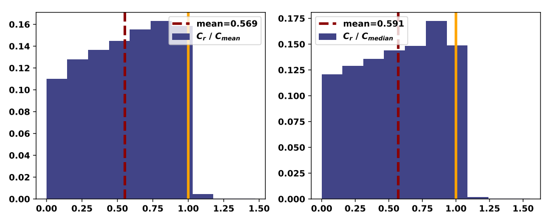


Figure 4.19: Distributions of C_r/C_{median} and C_r/C_{median} for winners of noisy probabilistic ending tournaments.

In this section the effect of several measures, regarding a strategy's behaviour and the tournament in which it participated on its performance were presented. This was done using two approaches. Correlation coefficients and a random forest analysis. The results of these are summarised in the following section.

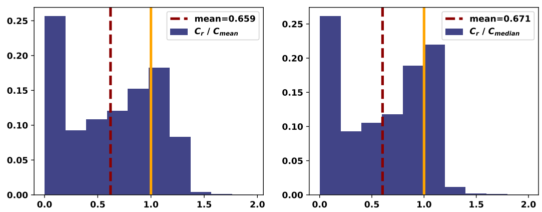


Figure 4.20: Distributions of C_r/C_{median} and C_r/C_{median} for winners of over all the tournaments.

4.5 Conclusion

This manuscript has explored the performance of 195 strategies of the Iterated Prisoner's Dilemma in a large number of computer tournaments. The results of the analysis demonstrated that, although for specific tournament types such as standard and probabilistic ending tournaments, dominant strategies exist there is not a single dominant type of strategies if the environments vary. Moreover, a strategy with a theory of mind should aim to adapt its behaviour based on the mean and median cooperators.

The 195 strategies used in this manuscript have been mainly for the literature, and they have been accessible due to an open source software called Axelrod-Python. The software was used to generate a total of 45686 computer tournaments results with different number of strategies and different participants each time. The data collection was described in Section A.2. In Section A.3, the tournaments results were used to present the top performances. The data set contained results from four different settings, and these were also studied individually. In standard tournaments complex strategies trained using reinforcement learning ranked in the top spots. Some of these strategies ranked again in the top spots in probabilistic ending tournaments when a p_e of less 0.1 was considered. In probabilistic ending tournaments p_e was designed to vary between 0 and 1. It was demonstrated that for values larger than 0.1, as stated in the Folk Theorem, defecting strategies were winning the tournaments because there was a high likelihood of the game ending in the next turn. In tournaments with noise the median ranks of the top 15 strategies had the highest values and the r distributions were bimodal. The top rank strategies were performing both well and bad, and this indicates that in noisy tournaments there are not strategies that can guarantee winning. Overall, the top ranked strategies differed from one tournament type to another and the mechanism behind the winning strategies were all different. Even strategies designed to do good in one setting did better in others. On the whole ... (the ipd interactions are unique and there is no winning strategy)

Section A.4, covered an analysis of performance based on several measures associated with a strategy and with the environments it was competing. The results of this analysis showed that a strategy's characteristics such as whether or not it's stochastic, and the information it used regarding the game had no effect on the strategy's success. The most important factors have been those that compared the strategy's behaviour to it's environment. The cooperating ratio of the strategy compared to the mean and median cooperator was highlighted as the most important feature in the analysis. More specifically, if a strategy were to enter a tournament with a theory of mind of its environment it would choose to be the median cooperator in standard tournaments, to cooperate 10% in probabilistic ending tournaments and 60% in noisy and noisy probabilistic tournaments of the times the median cooperator did. Lastly, if a strategy was aware of the opponents but not of the setting on the tournament, a strategy would be more likely to be successful if it were to identify the median cooperator and cooperated 67% of the times that they did.

The data set described in this work contains the largest number of IPD tournaments, to the authors knowledge, and it available at **data**. Further data mining could be applied and provide new insights in the field.

4.6 A summary of parameters

measure	measure explanation
stochastic	If a strategy is stochastic
makes use of game	If a strategy makes used of the game information
makes use of length	If a strategy makes used of the number of turns
memory usage	The memory size of a strategy divided by the number of turns
SSE	A measure of how far a strategy is from extortionate behaviour
C_{max}	The biggest cooperating rate in the tournament
C_{min}	The smallest cooperating rate in the tournament
C_{median}	The median cooperating rate in the tournament
C_{mean}	The mean cooperating rate in the tournament
C_r / C_{max}	A strategy's cooperating rate divided by the maximum
C_r / C_{min}	A strategy's cooperating rate divided by the minimum
C_r / C_{median}	A strategy's cooperating rate divided by the median
C_r / C_{mean}	A strategy's cooperating rate divided by the mean
C_r	The cooperating ratio of a strategy
CC to C rate	The probability a strategy will cooperate after a mutual cooperation
CD to C rate	The probability a strategy will cooperate after being betrayed by the opponent
DC to C rate	The probability a strategy will cooperate after betraying the opponent
DD to C rate	The probability a strategy will cooperate after a mutual defection
p_n	The probability of a player's action being flip at each interaction
n	The number of turns
p_e	The probability of a match ending in the next turn
N	The number of strategies in the tournament
k	The number that a given tournament is repeated

Table 4.11: The measures which are included in the performance evaluation analysis.

4.7 List of strategies

The strategies used in this study which are from Axelrod version 3.0.0 **axelrodproject**.

1. ϕ **axelrodproject**
2. π **axelrodproject**
3. e **axelrodproject**
4. ALLCorALLD **axelrodproject**
5. Adaptive **Li2011**
6. Adaptive Pavlov 2006 **kendall2007iterated**
7. Adaptive Pavlov 2011 **Li2011**
8. Adaptive Tit For Tat: 0.5 **Tzafestas2000**
9. Aggravater **axelrodproject**
10. Alexei **lesswrong**
11. Alternator **Axelrod1981**; **Mittal2009**
12. Alternator Hunter **axelrodproject**
13. Anti Tit For Tat **Hilbe2013**
14. AntiCycler **axelrodproject**
15. Appeaser **axelrodproject**
16. Arrogant QLearner **axelrodproject**
17. Average Copier **axelrodproject**
18. Backstabber **axelrodproject**
19. Better and Better **prison**
20. Bully **Nachbar1992**
21. Calculator **prison**
22. Cautious QLearner **axelrodproject**
23. Champion **Axelrod1980b**
24. CollectiveStrategy **Li2009**
25. Contrite Tit For Tat **Axelrod1995**
26. Cooperator **Axelrod1981**; **Mittal2009**; **Press2012**
27. Cooperator Hunter **axelrodproject**
28. Cycle Hunter **axelrodproject**
29. Cycler CCC-CCD **axelrodproject**
30. Cycler CCCD **axelrodproject**
31. Cycler CC-CDCD **axelrodproject**
32. Cycler CCD **Mittal2009**
33. Cycler DC **axelrodproject**
34. Cycler DDC **Mittal2009**
35. DBS **Au2006**
36. Davis **Axelrod1980a**
37. Defector **Axelrod1981**; **Mittal2009**; **Press2012**
38. Defector Hunter **axelrodproject**
39. Double Crosser **axelrodproject**
40. Desperate **Van2015**
41. DoubleResurrection **Eckhart2015**
42. Doubler **prison**
43. Dynamic Two Tits For Tat **axelrodproject**
44. EasyGo **Li2011**; **prison**
45. Eatherley **Axelrod1980b**
46. Eventual Cycle Hunter **axelrodproject**
47. Evolved ANN **axelrodproject**
48. Evolved ANN 5 **axelrodproject**
49. Evolved ANN 5 Noise 05 **axelrodproject**
50. Evolved FSM 16 **axelrodproject**
51. Evolved FSM 16 Noise 05 **axelrodproject**
52. Evolved FSM 4 **axelrodproject**
53. Evolved HMM 5 **axelrodproject**
54. EvolvedLookerUp1 1 1 **axelrodproject**
55. EvolvedLookerUp2 2 2 **axelrodproject**
56. Eugine Nier **lesswrong**
57. Feld **Axelrod1980a**
58. Firm But Fair **Frean1994**
59. Fool Me Forever **axelrodproject**
60. Fool Me Once **axelrodproject**
61. Forgetful Fool Me Once **axelrodproject**
62. Forgetful Grudger **axelrodproject**
63. Forgiver **axelrodproject**
64. Forgiving Tit For Tat **axelrodproject**

65. Fortress3 **Ashlock2006**
66. Fortress4 **Ashlock2006**
67. GTFT **Gaudesi2016**; **Nowak1993**
68. General Soft Grudger **axelrodproject**
69. Gradual **Beaufils1997**
70. Gradual Killer **prison**
71. Grofman**Axelrod1980a**
72. Grudger **Axelrod1980a**; **Banks1990**; **Beaufils1997**; **Van2015**; **Li2011**
73. GrudgerAlternator **prison**
74. Grumpy **axelrodproject**
75. Handshake **Robson1990**
76. Hard Go By Majority **Mittal2009**
77. Hard Go By Majority: 10 **axelrodproject**
78. Hard Go By Majority: 20 **axelrodproject**
79. Hard Go By Majority: 40 **axelrodproject**
80. Hard Go By Majority: 5 **axelrodproject**
81. Hard Prober **prison**
82. Hard Tit For 2 Tats **Stewart2012**
83. Hard Tit For Tat **PD2017**
84. Hesitant QLearner**axelrodproject**
85. Hopeless **Van2015**
86. Inverse **axelrodproject**
87. Inverse Punisher **axelrodproject**
88. Joss **Axelrod1980a**; **Stewart2012**
89. Knowledgeable Worse and Worse **axelrodproject**
90. Level Punisher **Eckhart2015**
91. Limited Retaliate 2 **axelrodproject**
92. Limited Retaliate 3 **axelrodproject**
93. Limited Retaliate **axelrodproject**
94. MEM2 **Li2014**
95. Math Constant Hunter **axelrodproject**
96. Meta Hunter Aggressive **axelrodproject**
97. Meta Hunter **axelrodproject**
98. Meta Majority **axelrodproject**
99. Meta Majority Finite Memory **axelrodproject**
100. Meta Majority Long Memory **axelrodproject**
101. Meta Majority Memory One **axelrodproject**
102. Meta Minority **axelrodproject**
103. Meta Mixer **axelrodproject**
104. Meta Winner **axelrodproject**
105. Meta Winner Deterministic **axelrodproject**
106. Meta Winner Ensemble **axelrodproject**
107. Meta Winner Finite Memory **axelrodproject**
108. Meta Winner Long Memory **axelrodproject**
109. Meta Winner Memory One **axelrodproject**
110. Meta Winner Stochastic **axelrodproject**
111. NMWE Deterministic **axelrodproject**
112. NMWE Finite Memory **axelrodproject**
113. NMWE Long Memory **axelrodproject**
114. NMWE Memory One **axelrodproject**
115. NMWE Stochastic **axelrodproject**
116. Naive Prober **Li2011**

117. Negation **PD2017**
118. Nice Average Copier **axelrodproject**
119. Nice Meta Winner **axelrodproject**
120. Nice Meta Winner Ensemble **axelrodproject**
121. Nydegger **Axelrod1980a**
122. Omega TFT **kendall2007iterated**
123. Once Bitten **axelrodproject**
124. Opposite Grudger **axelrodproject**
125. PSO Gambler 1 1 1 **axelrodproject**
126. PSO Gambler 2 2 2 **axelrodproject**
127. PSO Gambler 2 2 2 Noise 05 **axelrodproject**
128. PSO Gambler Mem1 **axelrodproject**
129. Predator **Ashlock2006**
130. Prober **Li2011**
131. Prober 2 **prison**
132. Prober 3 **prison**
133. Prober 4 **prison**
134. Pun1 **Ashlock2006**
135. Punisher **axelrodproject**
136. Raider **Ashlock2014**
137. Random Hunter **axelrodproject**
138. Random: 0.5 **Axelrod1980a; Tzafestas2000**
139. Remorseful Prober **Li2011**
140. Resurrection **Eckhart2015**
141. Retaliate 2 **axelrodproject**
142. Retaliate 3 **axelrodproject**
143. Retaliate **axelrodproject**
144. Revised Downing **Axelrod1980a**
145. Ripoff **Ashlock2008**
146. Risky QLearner **axelrodproject**
147. SelfSteem **Andre2013**
148. ShortMem **Andre2013**
149. Shubik **Axelrod1980a**
150. Slow Tit For Two Tats **axelrodproject**
151. Slow Tit For Two Tats 2 **prison**
152. Sneaky Tit For Tat **axelrodproject**
153. Soft Go By Majority **Axelrod1981; Mittal2009**
154. Soft Go By Majority 10 **axelrodproject**
155. Soft Go By Majority 20 **axelrodproject**
156. Soft Go By Majority 40 **axelrodproject**
157. Soft Go By Majority 5 **axelrodproject**
158. Soft Grudger **Li2011**
159. Soft Joss **prison**
160. SolutionB1 **Ashlock2015**
161. SolutionB5 **Ashlock2015**
162. Spiteful Tit For Tat **prison**
163. Stalker **Carvalho2013**
164. Stein and Rapoport **Axelrod1980a**
165. Stochastic Cooperator **Adami2013**
166. Stochastic WSLS **axelrodproject**
167. Suspicious Tit For Tat **Beaufils1997; Hilbe2013**
168. TF1 **axelrodproject**
169. TF2 **axelrodproject**
170. TF3 **axelrodproject**
171. Tester **Axelrod1980b**
172. ThueMorse **axelrodproject**
173. ThueMorseInverse **axelrodproject**
174. Thumper **Ashlock2008**
175. Tit For 2 Tats (**Tf2T**) **Axelrod1981**
176. Tit For Tat (**TfT**) **Axelrod1980a**
177. Tricky Cooperator **axelrodproject**
178. Tricky Defector **axelrodproject**
179. Tullock **Axelrod1980a**

180. Two Tits For Tat (**2TfT**) **Axelrod1981**
181. VeryBad **Andre2013**
182. Willing **Van2015**
183. Win-Shift Lose-Stay (**WShLSt**) **Li2011**
184. Win-Stay Lose-Shift (**WSLS**) **Kraines1989**; **Nowak1993**; **Stewart2012**
185. Winner12 **mathieu2017**
186. Winner21 **mathieu2017**
187. Worse and Worse**prison**
188. Worse and Worse 2**prison**
189. Worse and Worse 3**prison**
190. ZD-Extort-2 v2 **Kuhn2017**
191. ZD-Extort-2 **Stewart2012**
192. ZD-Extort-4 **axelrodproject**
193. ZD-GEN-2 **Kuhn2017**
194. ZD-GTFT-2 **Stewart2012**
195. ZD-SET-2 **Kuhn2017**

4.8 Correlation coefficients

A graphical representation of the correlation coefficients for the measures in Table A.8.

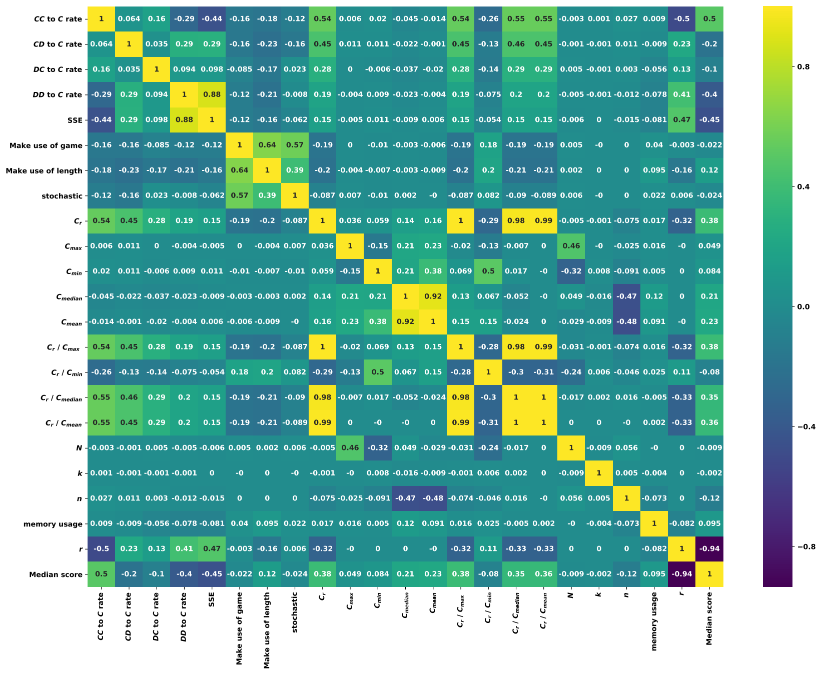


Figure 4.21: Correlation coefficients of measures in Table A.8 for standard tournaments

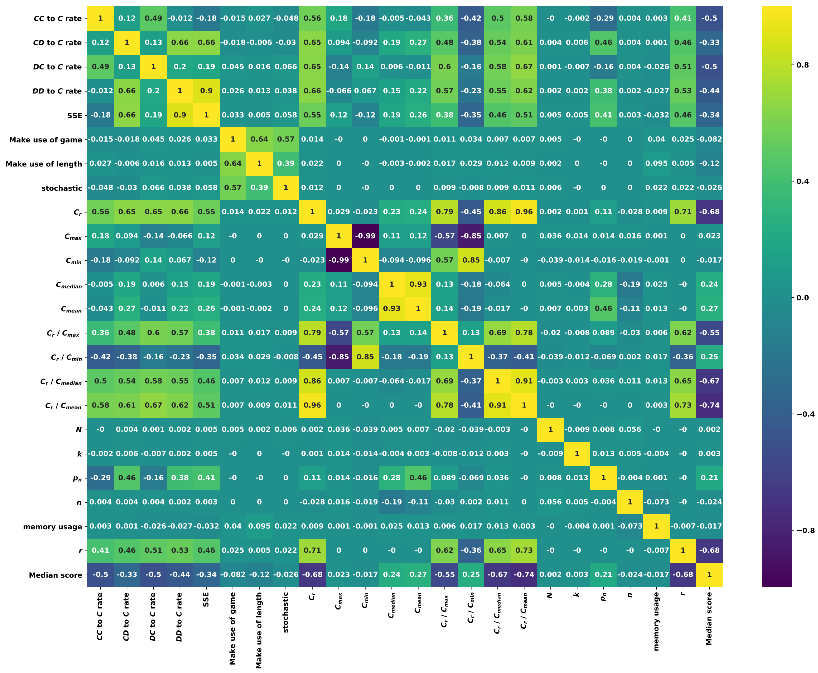


Figure 4.22: Correlation coefficients of measures in Table A.8 for noisy tournaments

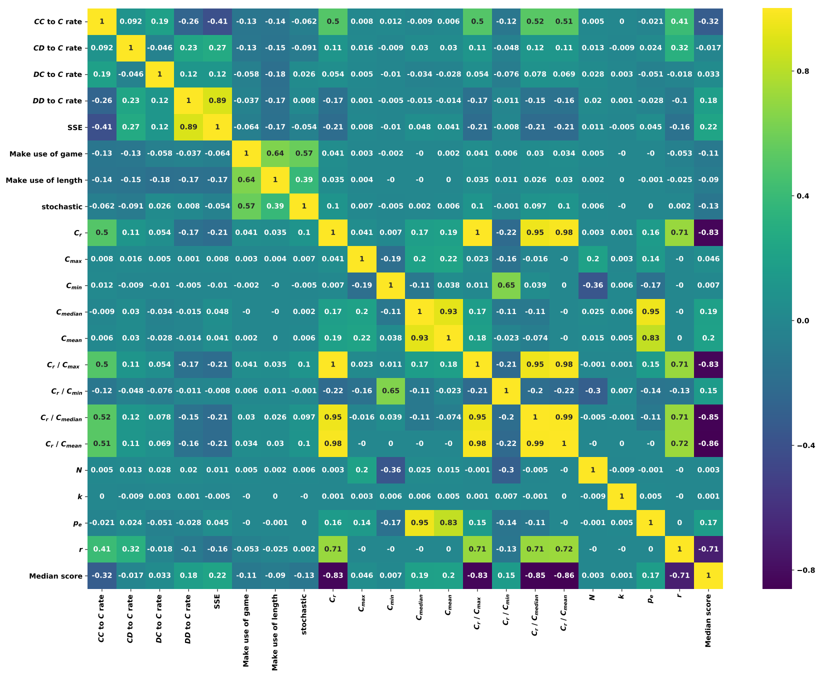


Figure 4.23: Correlation coefficients of measures in Table A.8 for probabilistic ending tournaments

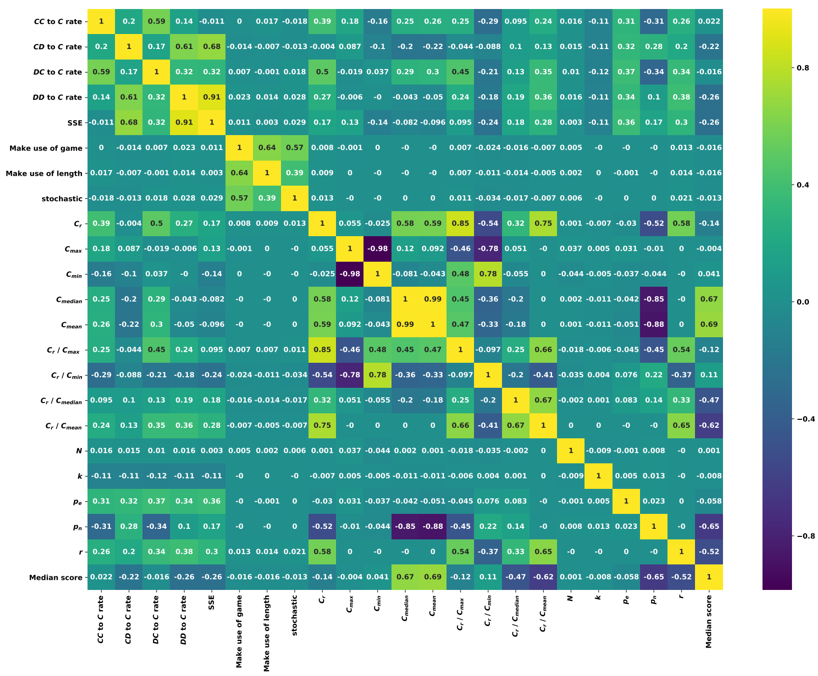


Figure 4.24: Correlation coefficients of measures in Table A.8 for noisy probabilistic ending tournaments

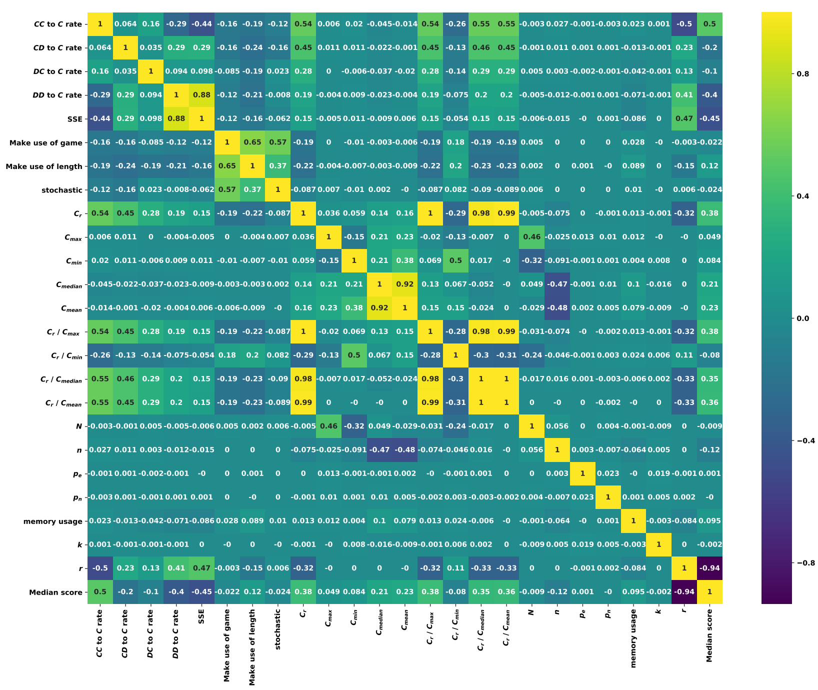


Figure 4.25: Correlation coefficients of measures in Table A.8 for data set

Chapter 5

Stability of defection, optimisation of strategies and the limits of memory in the Prisoner's Dilemma.

Memory-one strategies are a set of Iterated Prisoner's Dilemma strategies that have been praised for their mathematical tractability and performance against single opponents. This manuscript investigates *best response* memory-one strategies as a multi-dimensional optimisation problem. Though extortionate memory-one strategies have gained much attention, we demonstrate that best response memory-one strategies do not behave in an extortionate way, and moreover, for memory one strategies to be evolutionary robust they need to be able to behave in a forgiving way. We also provide evidence that memory-one strategies suffer from their limited memory in multi agent interactions and can be out performed by longer memory strategies.

5.1 Introduction

The Prisoner's Dilemma (PD) is a two player game used in understanding the evolution of cooperative behaviour, formally introduced in **Flood1958**. Each player has two options, to cooperate (C) or to defect (D). The decisions are made simultaneously and independently. The normal form representation of the game is given by:

$$S_p = \begin{pmatrix} R & S \\ T & P \end{pmatrix} \quad S_q = \begin{pmatrix} R & T \\ S & P \end{pmatrix} \tag{5.1}$$

where S_p represents the utilities of the row player and S_q the utilities of the column player. The payoffs, (R, P, S, T), are constrained by equations (B.2) and (B.3). Constraint (B.2) ensures that defection dominates cooperation and constraint (B.3) ensures that there is a dilemma; the sum of the utilities for both players is better when both choose to cooperate. The most common values used in the literature are $(R, P, S, T) = (3, 1, 0, 5)$ **Axelrod1981**.

$$T > R > P > S \tag{5.2}$$

$$2R > T + S \tag{5.3}$$

The PD is a one shot game, however, it is commonly studied in a manner where the history of the interactions matters. The repeated form of the game is called the Iterated Prisoner's Dilemma (IPD) and in the 1980s, following the work of **Axelrod1980a**; **Axelrod1980b** it attracted the attention of the scientific community. In **Axelrod1980a** and **Axelrod1980b**, the first well known computer tournaments of the IPD were performed. A total of 13 and 62 strategies were submitted respectively in the form of computer code. The contestants competed against each other, a copy of themselves and a random strategy, and the winner was then decided on the average score achieved (not the total number of wins). The contestants were given access to the entire history of a match, however, how many turns of history a strategy would incorporate, referred to as the *memory size* of a strategy, was a result of the particular strategic decisions made by the author. The winning strategy of both tournaments was a strategy called Tit for Tat and its success in both tournaments came as a surprise. Tit for Tat was a simple, forgiving strategy that opened each interaction by cooperation, and had won the tournament even though it never scored higher than that its direct opponent. Tit for Tat provided evidence that being nice can be advantageous and became the major paradigm for reciprocal altruism.

Another trait of Tit for Tat is that it considers only the previous move of the opponent. These type of strategies are called *reactive* **Nowak1989** and are a subset of so called *memory-one* strategies, which incorporate both players' latests moves. Memory-one strategies have been studied thoroughly in the literature **Nowak1990**; **Nowak1993**, however, they have gained most of their attention when a certain subset of memory-one strategies was introduced in **Press2012**, the zero-determinants. In **Stewart2012** it was stated that "Press and Dyson have fundamentally changed the viewpoint on the Prisoner's Dilemma". Zero-determinants are a special case of memory-one and extortionate strategies. They choose their actions so that a linear relationship is forced between the players' score ensuring that they will always receive at least as much as their opponents. Zero-determinants are indeed mathematically unique and are proven to be robust in pairwise interactions, however, their true effectiveness in tournaments and evolutionary dynamics has been questioned **adami2013**; **Hilbe2013b**; **Hilbe2013**; **Hilbe2015**; **Knight2018**; **Harper2015**.

In a similar fashion to **Press2012** the purpose of this work is to consider a given memory-one strategy; however, whilst **Press2012** found a way for a player to manipulate a given opponent, this work will consider a multidimensional optimisation approach to identify the best response to a given group of opponents. In particular, this work presents a compact method of identifying the best response memory-one strategy against a given set of opponents, and evaluates whether it behaves extortionately, similar to zero-determinants. Further theoretical and empirical results of this work include:

1. The conditions that ensure a best response memory-one strategy evolutionary robust.
2. A well designed framework that allows the comparison of an optimal memory one strategy and a more complex strategy which has a larger memory and was obtained through reinforcement learning techniques **Harper2017**.
3. An identification of conditions for which defection is known to be stable; thus identifying environments where cooperation will not occur.

5.2 The utility

One specific advantage of memory-one strategies is their mathematical tractability. They can be represented completely as an element of $\mathbb{R}^4_{[0,1]}$. This originates from **Nowak1989** where it is stated that if a strategy is concerned with only the outcome of a single turn then there are four possible 'states' the strategy could be in;

- both players cooperated, denoted as CC
- first players cooperated whilst the second player defected, denoted as CD
- first players defected whilst the second player cooperated, denoted as DC
- both players defected, denoted as DD

Therefore, a memory-one strategy can be denoted by the probability vector of cooperating after each of these states; $p = (p_1, p_2, p_3, p_4) \in \mathbb{R}^4_{[0,1]}$.

In **Nowak1989** it was shown that it is not necessary to simulate the play of a strategy p against a memory-one opponent q. Rather this exact behaviour can be modeled as a stochastic process, and more specifically as a Markov chain (Figure B.1) whose

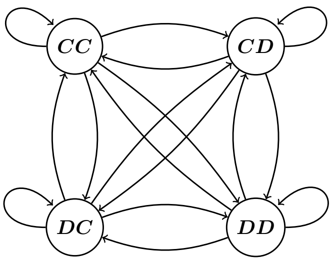


Figure 5.1: Markov Chain

corresponding transition matrix M is given by (B.4). The long run steady state probability vector v, which is the solution to $vM = v$, can be combined with the payoff matrices of (B.1) to give the expected payoffs for each player. More specifically, the utility for a memory-one strategy p against an opponent q, denoted as $u_q(p)$, is given by (B.5).

$$M = \begin{bmatrix} p_1q_1 & p_1(-q_1+1) & q_1(-p_1+1) & (-p_1+1)(-q_1+1) \\ p_2q_3 & p_2(-q_3+1) & q_3(-p_2+1) & (-p_2+1)(-q_3+1) \\ p_3q_2 & p_3(-q_2+1) & q_2(-p_3+1) & (-p_3+1)(-q_2+1) \\ p_4q_4 & p_4(-q_4+1) & q_4(-p_4+1) & (-p_4+1)(-q_4+1) \end{bmatrix} \tag{5.4}$$

$$u_q(p) = v \cdot (R, S, T, P). \tag{5.5}$$

This manuscript has explored the form of $u_q(p)$, to the authors knowledge no previous work has done this, and it proves that $u_q(p)$ is given by a ratio of two quadratic forms **kepner2011**, Theorem 1.

Theorem 1. *The expected utility of a memory-one strategy $p \in \mathbb{R}^4_{[0,1]}$ against a memory-one opponent $q \in \mathbb{R}^4_{[0,1]}$, denoted as $u_q(p)$, can be written as a ratio of two quadratic forms:*

$$u_q(p) = \frac{\frac{1}{2}pQp^T + cp + a}{\frac{1}{2}p\bar{Q}p^T + \bar{c}p + \bar{a}}, \tag{5.6}$$

where $Q, \bar{Q} \in \mathbb{R}^{4\times4}$ are square matrices defined by the transition probabilities of the opponent q_1, q_2, q_3, q_4 as follows:

$$Q = \begin{bmatrix} 0 & -(q_1 - q_3)(q_2 - 5q_4 - 1) & q_3(q_1 - q_2) & -5q_3(q_1 - q_4) \\ -(q_1 - q_3)(q_2 - 5q_4 - 1) & 0 & (q_2 - q_3)(q_1 - 3q_4 - 1) & (q_3 - q_4)(5q_1 - 3q_2 - 2) \\ q_3(q_1 - q_2) & (q_2 - q_3)(q_1 - 3q_4 - 1) & 0 & 3q_3(q_2 - q_4) \\ -5q_3(q_1 - q_4) & (q_3 - q_4)(5q_1 - 3q_2 - 2) & 3q_3(q_2 - q_4) & 0 \end{bmatrix}, \quad (5.7)$$

$$\bar{Q} = \begin{bmatrix} 0 & -(q_1 - q_3)(q_2 - q_4 - 1) & (q_1 - q_2)(q_3 - q_4) & (q_1 - q_4)(q_2 - q_3 - 1) \\ -(q_1 - q_3)(q_2 - q_4 - 1) & 0 & (q_2 - q_3)(q_1 - q_4 - 1) & (q_1 - q_2)(q_3 - q_4) \\ (q_1 - q_2)(q_3 - q_4) & (q_2 - q_3)(q_1 - q_4 - 1) & 0 & -(q_2 - q_4)(q_1 - q_3 - 1) \\ (q_1 - q_4)(q_2 - q_3 - 1) & (q_1 - q_2)(q_3 - q_4) & -(q_2 - q_4)(q_1 - q_3 - 1) & 0 \end{bmatrix}. \quad (5.8)$$

c and $\bar{c} \in \mathbb{R}^{4\times 1}$ are similarly defined by:

$$c = \begin{bmatrix} q_1(q_2 - 5q_4 - 1) \\ -(q_3 - 1)(q_2 - 5q_4 - 1) \\ -q_1q_2 + q_2q_3 + 3q_2q_4 + q_2 - q_3 \\ 5q_1q_4 - 3q_2q_4 - 5q_3q_4 + 5q_3 - 2q_4 \end{bmatrix}, \quad (5.9)$$

$$\bar{c} = \begin{bmatrix} q_1(q_2 - q_4 - 1) \\ -(q_3 - 1)(q_2 - q_4 - 1) \\ -q_1q_2 + q_2q_3 + q_2 - q_3 + q_4 \\ q_1q_4 - q_2 - q_3q_4 + q_3 - q_4 + 1 \end{bmatrix}, \quad (5.10)$$

and the constant terms $a, \bar{a}$ are defined as $a = -q_2 + 5q_4 + 1$ and $\bar{a} = -q_2 + q_4 + 1$.

The proof of Theorem 1 is given in Appendix B.6.1. Furthermore, numerical simulations have been carried out to validate the result. The simulated utility, which is denoted as $U_q(p)$, has been calculated using **axelrodproject** an open source research framework for the study of the IPD (**axelrodproject** is described in **Knight2016**). For smoothing the simulated results the utility has been estimated in a tournament of 500 turns and 200 repetitions. Figure B.2 shows two examples demonstrating that the formulation of Theorem 1 successfully captures the simulated behaviour.

The source code used in this manuscript has been written in a sustainable manner. It is open source (`https://github.com/Nikoleta-v3/Memory-size-in-the-prisoners-dilemma`) and tested which ensures the validity of the results. It has also been archived and can be found at.

Theorem 1 can be extended to consider multiple opponents. The IPD is commonly studied in tournaments and/or Moran Processes where a strategy interacts with a number of opponents. The payoff of a player in such interactions is given by the average payoff the player received against each opponent. More specifically the expected utility of a memory-one strategy against a N number of opponents is given by Theo-

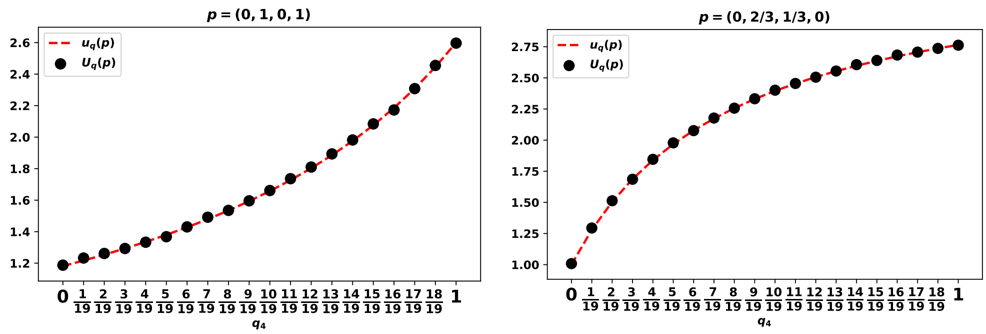


Figure 5.2: Simulated and empirical utilities for $p = (0, 1, 0, 1)$ and $p = (0, \frac{2}{3}, \frac{1}{3}, 0)$ against $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, q_4)$ for $q_4 \in \{0, \frac{1}{19}, \frac{2}{19}, \dots, \frac{18}{19}, 1\}$. $u_q(p)$ is the theoretic value given in Theorem 1, and $U_q(p)$ is simulated numerically.

rem 2.

Theorem 2. *The expected utility of a memory-one strategy* $p \in \mathbb{R}^4_{[0,1]}$ *against a group of opponents* $q^{(1)}, q^{(2)}, \dots, q^{(N)}$, *denoted as* $\frac{1}{N}\sum\limits_{i=1}^{N} {u_q}^{(i)}(p)$, *is given by:*

$$\frac{1}{N}\sum_{i=1}^{N} {u_q}^{(i)}(p) = \frac{1}{N} \frac{\sum\limits_{i=1}^{N} (\frac{1}{2}pQ^{(i)}p^T + c^{(i)}p + a^{(i)}) \prod\limits_{\substack{j=1 \\ j \neq i}}^{N} (\frac{1}{2}p\bar{Q}^{(j)}p^T + \bar{c}^{(j)}p + \bar{a}^{(j)})}{\prod\limits_{i=1}^{N} (\frac{1}{2}p\bar{Q}^{(i)}p^T + \bar{c}^{(i)}p + \bar{a}^{(i)})}. \tag{5.11}$$

The proof of Theorem 2 is a straightforward algebraic manipulation.

Similar to the previous result, the formulation of Theorem 2 is validated using numerical simulations where the 10 memory-one strategies described in **Stewart2012** have been used as the opponents. Figure B.3 shows that the simulated behaviour has been captured successfully.

The list of strategies from **Stewart2012** was also used to check whether the utility against a group of strategies could be captured by the utility against the mean opponent. Thus whether condition (B.12) holds. However condition (B.12) fails, as shown in Figure B.4.

$$\frac{1}{N}\sum_{i=1}^{N} {u_q}^{(i)}(p) = u_{\frac{1}{N}\sum\limits_{i=1}^{N} q^{(i)}}(p), \tag{5.12}$$

Theorem 2 which allows for the utility of a memory-one strategy against any number of opponents to be estimated without simulating the interactions is the main result used in this manuscript. In Section B.3 it is used to define best response memory-one strategies

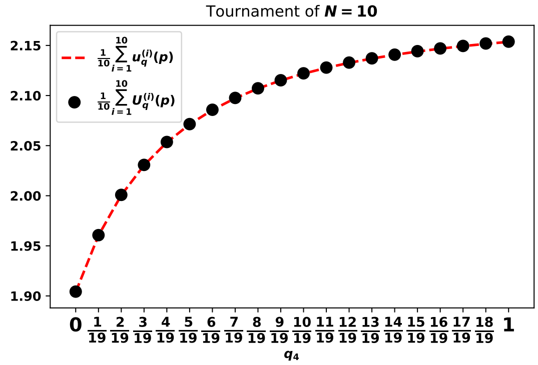


Figure 5.3: The utilities of memory-one strategies $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, p_4)$ for $p_4 \in \{0, \frac{1}{19}, \frac{2}{19}, \dots, \frac{18}{19}, 1\}$ against the 10 memory-one strategies described in **Stewart2012**. $\frac{1}{10}\sum_{i=1}^{10} u_q^{(i)}(p)$ is the theoretic value given in Theorem 1, and $\frac{1}{10}\sum_{i=1}^{10} U_q^{(i)}(p)$ is simulated numerically.

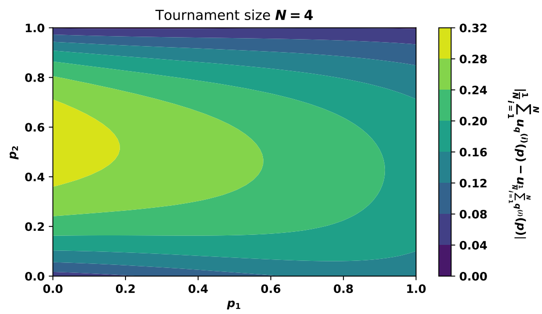


Figure 5.4: The difference between the average utility against the opponents from **Stewart2012** and the utility against the average player of the strategies in **Stewart2012** of a player $p = (p_1, p_2, p_1, p_2)$. A positive difference indicates that condition (B.12) does not hold.

and explore the conditions under which defection dominates cooperation.

5.3 Best responses to memory-one players

This section focuses on best responses and more specifically *memory-one best response* strategies. A *best response* is a strategy which corresponds to the most favorable outcome **Tadelis2013**, thus a memory-one best response to a set of opponents $q^{(1)}, q^{(2)}, \dots, q^{(N)}$ corresponds to a strategy p^* for which (B.11) is maximised. This is considered as a multi dimensional optimisation problem given by:

$$\begin{aligned} \max_p : & \sum_{i=1}^{N} {u_q}^{(i)}(p) \\ \text{such that} : & \ p \in \mathbb{R}_{[0,1]} \end{aligned} \tag{5.13}$$

Optimising this particular ratio of quadratic forms is not trivial. It can be verified empirically for the case of a single opponent that there exists at least one point for which the definition of concavity does not hold, see Appendix B.7.1 for an example. Some results are known for non concave ratios of quadratic forms **Beck2009**; **Hongyan2014**, however, in these works it is assumed that either both the numerator and the denominator of the fractional problem are concave or that the denominator is greater than zero which in this case are not true (as seen in Theorem 3).

Theorem 3. *The utility of a player p against an opponent q, $u_q(p)$, given by (B.6), is not concave. Furthermore neither the numerator or the denominator of (B.6), are concave or strictly greater than zero.*

Proof is given in Appendix B.6.2.

The non concavity of $u(p)$ indicates multiple local optimal points. The approach taken here is to introduce a compact way of constructing the candidate set of all local optimal points, and evaluating which corresponds to the best response strategy (maximises (B.11)).

The problem considered is bounded because $p \in \mathbb{R}^4_{[0,1]}$. Therefore, the candidate solutions will exist either at the boundaries of the feasible solution space, or within that space (the methods of Lagrange Multipliers **bertsekas2014** and Karush-Kuhn-Tucker conditions **Giorgi2016** are based on this). This approach allow us to define the best response memory-one strategy to a group of opponents in the following Lemma:

Lemma 4. *The optimal behaviour of a memory-one strategy player $p^* \in \mathbb{R}^4_{[0,1]}$ against a set of N opponents $\{q^{(1)}, q^{(2)}, \dots, q^{(N)}\}$ for $q^{(i)} \in \mathbb{R}^4_{[0,1]}$ is given by:*

$$p^* = \text{argmax} \sum_{i=1}^{N} u_q(p), \ p \in S_q.$$

The set S_q is defined as all the possible combinations of:

$$S_q = \left\{ p \in \mathbb{R}^4 \left| \begin{array}{l} \bullet \quad p_j \in \{0,1\} \quad \text{and} \quad \dfrac{d}{dp_k} \displaystyle\sum_{i=1}^{N} u_q^{(i)}(p) = 0 \quad \text{for all} \quad j \in J \quad \& \quad k \in K \quad \text{for all} \quad J, K \\ \qquad\qquad\qquad\quad \text{where} \quad J \cap K = \emptyset \quad \text{and} \quad J \cup K = \{1,2,3,4\}. \\ \bullet \quad p \in \{0,1\}^4 \end{array} \right. \right\} \tag{5.14}$$

The proof is given in Appendix B.6.3.

Note that there is no immediate way to find the zeros of $\frac{d}{dp} \sum\limits_{i=1}^{N} u_q(p)$;

$$\frac{d}{dp} \sum_{i=1}^{N} {u_q}^{(i)}(p) =$$
$$= \sum_{i=1}^{N} \frac{\left(pQ^{(i)} + c^{(i)}\right)\left(\frac{1}{2}p\bar{Q}^{(i)}p^T + \bar{c}^{(i)}p + \bar{a}^{(i)}\right) - \left(p\bar{Q}^{(i)} + \bar{c}^{(i)}\right)\left(\frac{1}{2}pQ^{(i)}p^T + c^{(i)}p + a^{(i)}\right)}{\left(\frac{1}{2}p\bar{Q}^{(i)}p^T + \bar{c}^{(i)}p + \bar{a}^{(i)}\right)^2} \tag{5.15}$$

For $\frac{d}{dp} \sum\limits_{i=1}^{N} u_q(p)$ to equal zero then:

$$\sum_{i=1}^{N} \left(\left(pQ^{(i)} + c^{(i)}\right)\left(\frac{1}{2}p\bar{Q}^{(i)}p^T + \bar{c}^{(i)}p + \bar{a}^{(i)}\right) - \left(p\bar{Q}^{(i)} + \bar{c}^{(i)}\right)\left(\frac{1}{2}pQ^{(i)}p^T + c^{(i)}p + a^{(i)}\right) \right) = 0, \quad \textit{while} \tag{5.16}$$

$$\sum_{i=1}^{N} \frac{1}{2}p\bar{Q}^{(i)}p^T + \bar{c}^{(i)}p + \bar{a}^{(i)} \neq 0. \tag{5.17}$$

Finding best response memory-one strategies, more specifically constructing the subset S_q, can be done analytically. The points for any or all of $p_i \in \{0,1\}$ for $i \in \{1,2,3,4\}$ are trivial, and finding the roots of the partial derivatives which are a set of polynomials of equations (B.16) is feasible using resultant theory **Jonsson2005**; however, for large systems building the resultant quickly becomes intractable. As a result, a numerical method taking advantage of the structure will be used for finding best response memory-one strategies. This will be described in Section B.4. The rest of the section focuses on an immediate theoretical result from Lemma 4.

5.3.1 Stability of defection

An immediate result from Lemma 4 can be obtained by evaluating the sign of the derivative (B.15) at $p = (0, 0, 0, 0)$. If at that point the derivative is negative, then the utility of a player only decreases if they were to change their behaviour, and thus defection at that point is stable.

Lemma 5. *In a tournament of N players* $\{q^{(1)}, q^{(2)}, \dots, q^{(N)}\}$ *for* $q^{(i)} \in \mathbb{R}^4_{[0,1]}$ *defection is stable if the transition probabilities of the opponents satisfy conditions (B.18) and (B.19).*

$$\sum_{i=1}^{N} (c^{(i)T}\bar{a}^{(i)} - \bar{c}^{(i)T}a^{(i)}) \leq 0 \tag{5.18}$$

while,

$$\sum_{i=1}^{N} \bar{a}^{(i)} \neq 0 \tag{5.19}$$

Proof. For defection to be stable the derivative of the utility at the point $p = (0, 0, 0, 0)$ must be negative. This would indicate that the utility function is only declining from that point onwards.

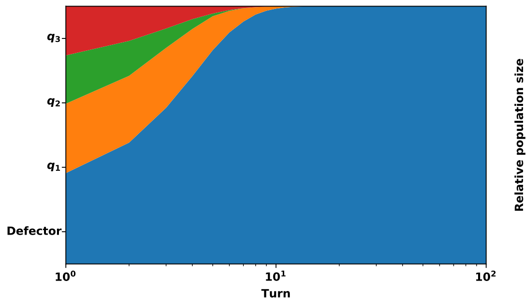
Substituting $p = (0, 0, 0, 0)$ in equation (B.15) gives:

$$\sum_{i=1}^{N} \frac{(c^{(i)T}\bar{a}^{(i)} - \bar{c}^{(i)T}a^{(i)})}{(\bar{a}^{(i)})^2} \tag{5.20}$$

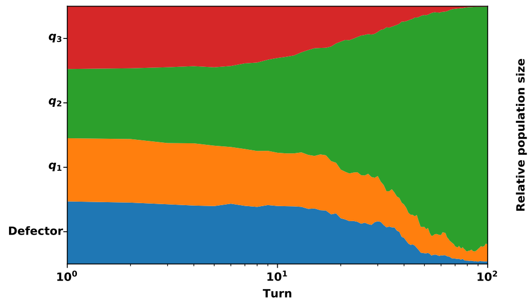
The sign of the numerator $\sum_{i=1}^{N} (c^{(i)T}\bar{a}^{(i)} - \bar{c}^{(i)T}a^{(i)})$ can vary based on the transition probabilities of the opponents. The denominator can not be negative, and otherwise is always positive. Thus the sign of the derivative is negative if and only if $\sum_{i=1}^{N} (c^{(i)T}\bar{a}^{(i)} - \bar{c}^{(i)T}a^{(i)}) \leq 0$. □

Consider a population for which defection is known to be stable. In that population all the members will over time adopt the same behaviour; thus in such population cooperation will never take over. This is demonstrated in Figures B.5a and B.5b.

Lemma 5 gives a condition under which cooperation cannot occur and is the last theoretical result presented in this manuscript. The following section focuses on numerical experiments.



(a) For opponents $q_1 = (\frac{371}{1250}, \frac{4693}{25000}, \frac{4037}{50000}, \frac{18461}{25000})$, $q_2 = (\frac{48841}{100000}, \frac{30587}{50000}, \frac{76591}{100000}, \frac{25921}{50000})$ and $q_3 = (\frac{22199}{100000}, \frac{87073}{100000}, \frac{646}{3125}, \frac{91861}{100000})$ conditions (B.18) and (B.19) hold and Defector takes over the population.



(b) For opponents $q_1 = (\frac{69773}{100000}, \frac{21609}{100000}, \frac{97627}{100000}, \frac{623}{100000})$, $q_2 = (\frac{12649}{50000}, \frac{43479}{100000}, \frac{38969}{50000}, \frac{19769}{100000})$ and $q_3 = (\frac{96703}{100000}, \frac{54723}{100000}, \frac{24317}{25000}, \frac{35741}{50000})$ (B.18) fails and (B.19) holds and Defector does not take over the population.

5.4 Numerical experiments

The results of this section rely on estimating best response memory-one strategies, but as stated in Section B.3, estimating best responses analytically can quickly become an intractable problem. As a result, best responses will be estimated heuristically using Bayesian optimisation **Mokus1978**. Bayesian optimisation is a global optimisation algorithm that has proven to outperform many other popular algorithms **Jones2001**. The algorithm builds a bayesian understanding of the objective function which is well suited to the potential multiple local optimas in the described search space of this work. Differential evolution **Storn1997** was also considered, however, it was not selected due to Bayesian optimisation being computationally more efficient.

As an example of the algorithm's usage let us consider the optimisation problem of (B.13). Figure B.6 illustrates the change of the utility function over iterations of the algorithm. The algorithm is set to run for 60 iterations. After 60 iterations if the utility has changed in the last 10% iterations then algorithm runs for a further 20 iterations. This is repeated until there is no change to the utility in the last 10% of iterations.

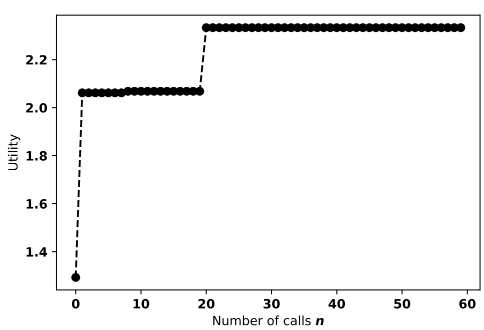


Figure 5.6: Utility over time of calls using Bayesian optimisation. The opponents are $q^{(1)} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $q^{(2)} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. The best response obtained is $p^* = (0, \frac{11}{50}, 0, 0)$

The rest of the section is structured as follows. In Section B.4.1, Bayesian optimisation is used to generate a data set containing memory-one best responses against a number of random opponents. The extortionate behaviour of these best responses is then evaluated using a method introduced in **Knight2019**. In Section B.4.2, a similar data set and approach is discussed but this time the best responses are memory-one best responses in an evolutionary setting where they also incorporate self interactions. This has immediate applications to Moran processes. Finally, Section B.4.3 compares the performances of memory-one and longer-memory best responses against a number of opponents.

5.4.1 Best response memory-one strategies for $N = 2$

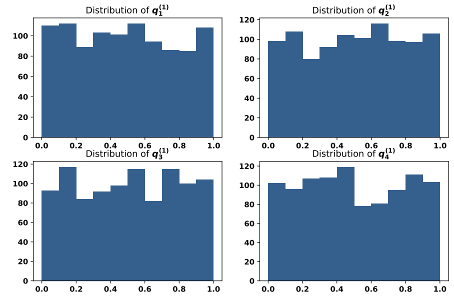
As briefly discussed in Section B.1, zero-determinants have been praised for their robustness against a single opponent. Zero-determinants are evidence that extortion works in pairwise interactions, their behaviour ensures that the strategies will never lose a game. However, this paper argues that in multi opponent interactions, where the payoffs matter, strategies trying to exploit their opponents will suffer.

Compared to zero-determinants, best response memory-one strategies which have a theory of mind of their opponents, utilise their behaviour in order to gain the most from their interactions. The question that arises then is whether best response strategies are optimal because they behave in an extortionate way. To estimate a strategy's extortionate behaviour the SSE method as described in **Knight2019** is used. SSE is defined as how far a strategy is from behaving extortionate, thus a high SSE implies a non extortionate behaviour.

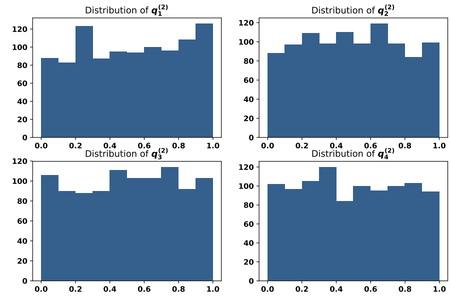
A data set of best response memory-one strategies with $N = 2$ opponents has been generated which is available at **glynatsi2019**. The data set contains a total of 1000 trials corresponding to 1000 different instances of a best response strategy. For each trial a set of 2 opponents is randomly generated and the memory-one best response against them is found. The probabilities q_i of the opponents are randomly generated and Figures B.7a and B.7b, show that they are uniformly distributed over the trials. Thus, the full space of possible opponents has been covered.

The SSE method has been applied to the data set. The distribution of SSE for the best response is given in Figure B.10 and a statistics summary in Table B.2. The distribution of SSE is skewed to the left, indicating that the best response does exhibit extortionate behaviour, however, the best response is not uniformly extortionate. A positive measure of skewness and kurtosis indicates a heavy tail to the right. Therefore, in several cases the strategy is not trying to extort its the opponents.

So although the best response strategy can exhibit extortionate behaviour, its performance is maximised by behaving in a more adaptable way than zero-determinant



(a) Distributions of first opponents' probabilities.



(b) Distributions of second opponents' probabilities.

strategies. This is confirms similar results such as **Knight2019**. This analysis will now be extended to an evolutionary setting.

5.4.2 Memory-one best responses in evolutionary dynamics

As mentioned in Section B.2, the IPD is commonly studied in Moran processes, and generally, in evolutionary processes. In these settings self interactions are key. This section extends the formulation of best responses in evolutionary dynamics, more specifically, the optimisation problem of (B.13) is extended to include self interactions.

Self interactions can be incorporated in the formulation that has been used so far. The utility is given by,

$$\frac{1}{N}\sum_{i=1}^{N} {u_q}^{(i)}(p) + u_p(p) \tag{5.21}$$

and the optimisation problem of (B.13) is modified to give:

$$\begin{aligned} \max_p : \ & \frac{1}{N}\sum_{i=1}^{N} {u_q}^{(i)}(p) + u_p(p) \\ \text{such that} : \ & p \in \mathbb{R}_{[0,1]} \end{aligned} \tag{5.22}$$

For determining the memory-one best response in an evolutionary setting, an algorithmic approach is considered, called *best response dynamics*. Best response dynamics are commonly used in evolutionary game theory. They represent a class of strategy updating rules, where players in the next round are determined by their best responses to some subset of the population. The best response dynamics approach used in this manuscript is given by Algorithm 2.

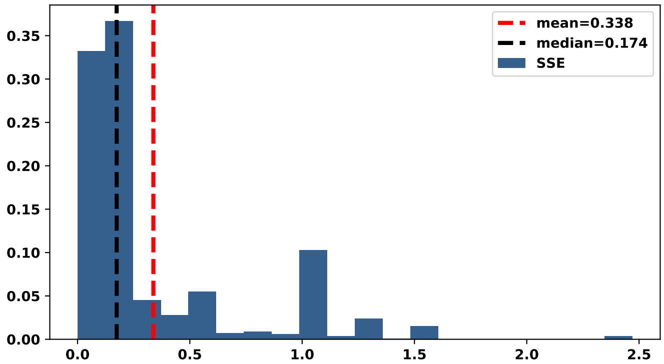


Figure 5.8: Distribution of SSE for memory-one best responses, when $N = 2$.

	SSE
count	1000.00000
mean	0.33762
std	0.39667
min	0.00000
5%	0.02078
25%	0.07597
50%	0.17407
95%	1.05943
max	2.47059
median	0.17407
skew	1.87231
kurt	3.60029

Table 5.1: Summary statistics SSE of best response memory one strategies included tournaments of $N = 2$.

Algorithm 2: Best response dynamics Algorithm

$p^{(t)} \leftarrow (1, 1, 1, 1)$;

while $p^{(t)} \neq p^{(t-1)}$ **do**

$p^{(t+1)} =$

$\text{argmax} \frac{1}{N} \sum_{i=1}^{N} u_q{}^{(i)}(p^{(t+1)}) + u_p^{(t)}(p^{(t+1)})$;

The best response dynamics algorithm starts by setting an initial solution $p^{(1)} = (1, 1, 1, 1)$, and repeatedly finds a strategy that maximises (B.22) using Bayesian optimisation. The algorithm stops once a cycle (a sequence of iterated evaluated points) is detected. A numerical example of the algorithm is given in Figure B.9.

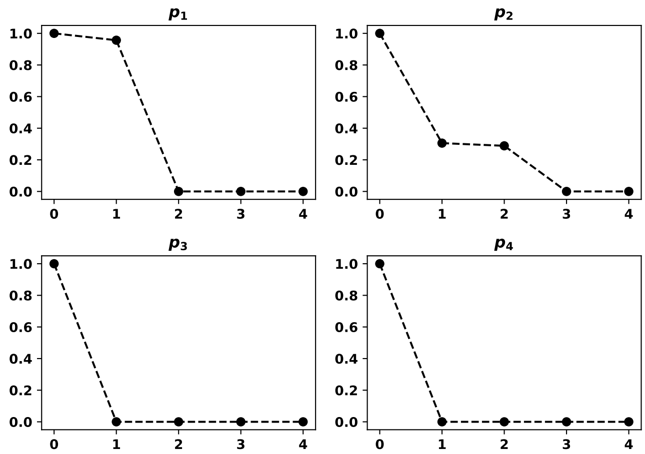


Figure 5.9: Best response dynamics with $N = 2$. More specifically, for $q^{(1)} = (\frac{59}{250}, \frac{1031}{10000}, \frac{99}{250}, \frac{1549}{10000})$ and $q^{(2)} = (\frac{133}{2000}, \frac{803}{2000}, \frac{9179}{10000}, \frac{2001}{2500})$.

The algorithm has been used to estimate the best response in an evolutionary setting for each of the 1000 pairs of opponents described in Section B.4.1. These are also included in the data set **glynatsi2019**, and moreover, the SSE method has also been applied. The distribution of SSE is given by Figure B.10 and a statistical summary by Table B.2.

Similarly to the results of Section B.4.1, the evolutionary best response strategy does not behave uniformly extortionately. A larger value of both the kurtosis and the skewness of the SSE distribution indicates that in evolutionary settings a memory-one best response is even more adaptable.

The difference between best responses in tournaments and in evolutionary settings are further explored by Figure B.11. Though, Table B.3 details that no statistically significant differences has been found, from Figure B.11, it seems that evolutionary best response has a higher p_2 median. Thus, they more likely to forgive after being tricked.

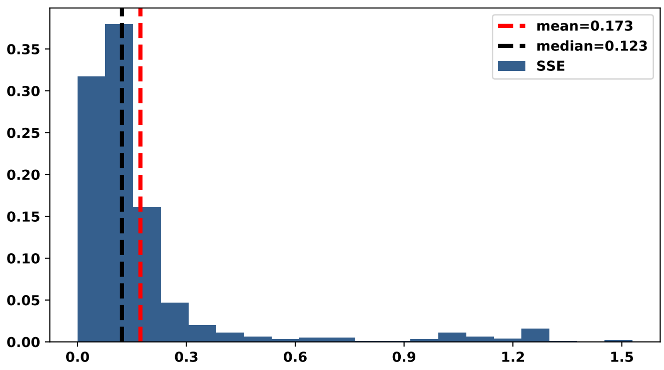


Figure 5.10: Distribution of SSE of best response memory-one strategies in evolutionary settings, when $N = 2$.

	SSE
count	1000.00000
mean	0.17326
std	0.23489
min	0.00001
5%	0.01497
25%	0.05882
50%	0.12253
95%	0.67429
max	1.52941
median	0.12253
skew	3.41839
kurt	11.92339

Table 5.2: Summary statistics SSE of best response memory-one strategies in evolutionary settings, when when $N = 2$.

Best Response Median in:	Tournament	Evolutionary Settings	p-values
Distribution p_1	0.0	0.00000	0.0
Distribution p_2	0.0	0.19847	0.0
Distribution p_3	0.0	0.00000	0.0
Distribution p_4	0.0	0.00000	0.0

Table 5.3: A non parametric test, Wilcoxon Rank Sum, has been performed to tests the difference in the median values of the cooperation probabilities in tournaments versus evolutionary settings. A non parametric test is used because is evident that the data are skewed.

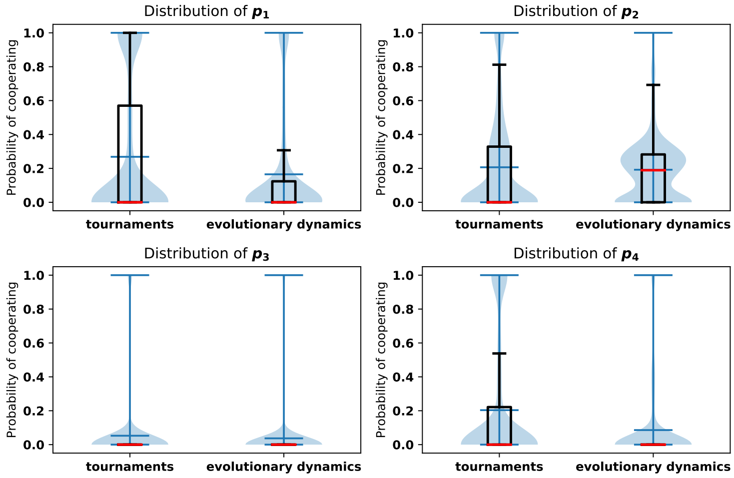


Figure 5.11: Distributions of p^* for both best response and evo memory-one strategies.

5.4.3 Longer memory best response

This section focuses on the memory size of strategies. The effectiveness of memory in the IPD has been previously explored in the literature, as discussed in Section B.1, however, none of the previous works has compared the performance of longer-memory strategies to memory-one best responses.

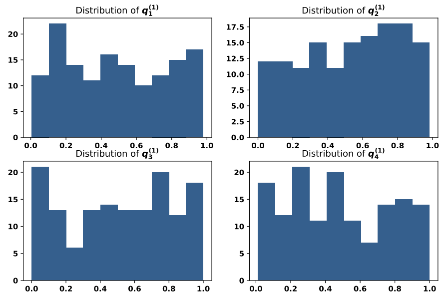
In **Harper2017**, a strategy called *Gambler* which makes probabilistic decisions based on the opponent's n_1 first moves, the opponent's m_1 last moves and the player's m_2 last moves was introduced. In this manuscript Gambler with parameters: $n_1 = 2, m_1 = 1$ and $m_2 = 1$ is used as a longer-memory strategy.

By considering the opponent's first two moves, the opponents last move and the player's last move, there are only 16 $(4 \times 2 \times 2)$ possible outcomes that can occur, furthermore, Gambler also makes a probabilistic decision of cooperating in the opening move. Thus, Gambler is a function $f : \{\text{C, D}\} \to [0, 1]_{\mathbb{R}}$. This can be hard coded as an element of $[0, 1]_{\mathbb{R}}^{16+1}$, one probability for each outcome plus the opening move. Hence, compared to (B.13), finding an optimal Gambler is a 17 dimensional problem given by:

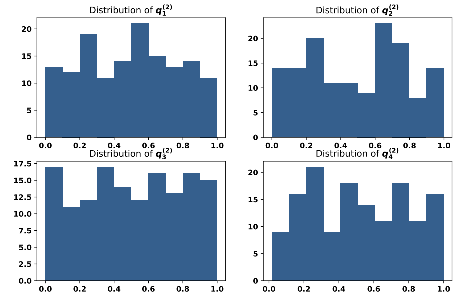
$$\begin{aligned} \max_p : & \sum_{i=1}^{N} {U_q}^{(i)}(f) \\ \text{such that} : & \ f \in \mathbb{R}_{[0,1]}^{17} \end{aligned} \tag{5.23}$$

Note that (B.11) can not be used here for the utility of Gambler, and actual simulated players are used. This is done using **axelrodproject** with 500 turns and 200 repetitions, moreover, (B.23) is solved numerically using Bayesian optimisation.

Similarly to previous sections, a large data set has been generated with instances of an optimal Gambler and a memory-one best response, available at **glynatsi2019**. Estimating a best response Gambler (17 dimensions) is computational more expensive compared to a best response memory-one (4 dimensions). As a result, the analysis of this section is based on a total of 130 trials. For each trial two random opponents have been selected. The 130 pair of opponents are a sub set of the opponents used in Section B.4.1- B.4.2. The distributions of their transition probabilities are given in Figures B.12a and B.12a.



(a) Distributions of first opponents' probabilities for longer memory experiment.



(b) Distributions of second opponents' probabilities for longer memory experiment.

The utilities of both strategies are plotted against each other in Figure B.13. Although Gambler has an infinite memory (in order to remember the opening moves of the opponent) the information the strategy considers is not significantly larger than memory-one strategies. Even so, it is evident from Figure B.13 that Gambler always performs as well as the best response memory-one or better. This seems to be at odd with the result of **Press2012** that against a memory-one opponent having a longer memory will not give a strategy any advantage. However, against two memory-one opponents Gambler's performance is better than the optimal memory-one strategy. This is evidence that in the case of two opponents having a shorter memory is limiting.

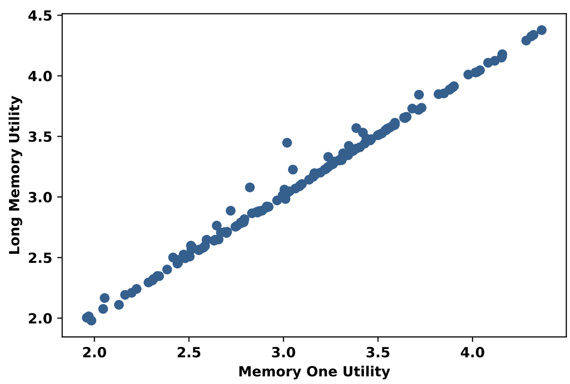


Figure 5.13: Utilities of Gambler and best response memory-one strategies for 130 different pair of opponents.

5.5 Conclusion

This manuscript has considered *best response* strategies in the IPD game, and more specifically, *memory-one best responses.* It has proven that there is a compact way of identifying a memory-one best response to a group of opponents, and moreover, that there exists a condition for which in an environment of memory-one opponents defection is the stable choice. The later parts of this paper focused on a series of empirical results, where it was shown that the performance and the evolutionary stability of memory-one strategies rely not on extortion but on adaptability. Finally, it was shown that memory-one strategies' performance is limited by their memory in cases where they interact with multiple opponents.

Following the work described in **Nowak1989**, where it was shown that the utility between two memory-one strategies can be estimated by a Markov stationary state, we proved that the utilities can be written as a ration of two quadratic forms in R^4, Theorem 1. This was extended to include multiple opponents, as the IPD is commonly studied in such situations, Theorem 2. The formulation of Theorem 2 allowed us to introduce an approach for identifying memory-one best responses to any number of opponents; Lemma 4. This does not only have game theoretic novelty, but also a mathematical novelty of solving quadratic ratio optimisation problem where the quadratics are non concave. The results of Lemma 4 were also used to define a condition for which defection is known to be stable.

This manuscript presented several experimental results. These results were mainly to investigate the behaviour of memory-one strategies and their limitations. In Sections B.4.1 and B.4.2, a large data set which contained best responses in tournaments and in evolutionary settings for $N = 2$ was generated. This allowed us to investigate their respective behaviours, and whether it was extortionate acts that made them the most favorable strategies. However, it was shown that it was not extortion but adaptability that allowed the strategies to gain the most from their interactions. In evolutionary settings it was specifically shown that being adaptable and being able to forgive after being tricked were key factors. In Section B.4.3, the performance of memory-one strategies was put against the performance of a longer memory strategy called Gambler. There were several cases where Gambler would outperform the memory-one strategy, however, a memory-one strategy never managed to outperform a Gambler. This result occurred whilst considering a Gambler with a sufficiently larger memory but not a sufficiently larger amount of information regarding the game.

All the empirical results presented in this manuscript have been for the case of $N = 2$. In future work we would consider larger values of N, however, we believe that for larger values of N the results that have been presented here would only be more evident.

5.6 Proofs of the Theorems

5.6.1 Proof of Theorem 1

The utility of a memory one player p against an opponent q, $u_q(p)$, can be written as a ratio of two quadratic forms on R^4.

Proof. In Section B.2, it was discussed that $u_q(p)$ it's the product of the steady states v and the PD payoffs,

$$u_q(p) = v \cdot (R, S, T, P).$$

More specifically, with $(R, P, S, T) = (3, 1, 0, 5)$

$$u_q(p) = \left(\frac{\begin{array}{c} p_1p_2(q_1q_2 - 5q_1q_4 - q_1 - q_2q_3 + 5q_3q_4 + q_3) + p_1p_3(-q_1q_3 + q_2q_3) + p_1p_4(5q_1q_3 - 5q_3q_4) + p_3p_4(-3q_2q_3 + 3q_3q_4) + \\ p_2p_3(-q_1q_2 + q_1q_3 + 3q_2q_4 + q_2 - 3q_3q_4 - q_3) + p_2p_4(-5q_1q_3 + 5q_1q_4 + 3q_2q_3 - 3q_2q_4 + 2q_3 - 2q_4) + \\ p_1(-q_1q_2 + 5q_1q_4 + q_1) + p_2(q_2q_3 - q_2 - 5q_3q_4 - q_3 + 5q_4 + 1) + p_3(q_1q_2 - q_2q_3 - 3q_2q_4 - q_2 + q_3) + \\ p_4(-5q_1q_4 + 3q_2q_4 + 5q_3q_4 - 5q_3 + 2q_4) + q_2 - 5q_4 - 1 \end{array}}{\begin{array}{l} p_1p_2(q_1q_2 - q_1q_4 - q_1 - q_2q_3 + q_3q_4 + q_3) + p_1p_3(-q_1q_3 + q_1q_4 + q_2q_3 - q_2q_4) + p_1p_4(-q_1q_2 + q_1q_3 + q_1 + q_2q_4 - \\ q_3q_4 - q_4) + \\ p_2p_3(-q_1q_2 + q_1q_3 + q_2q_4 + q_2 - q_3q_4 - q_3) + p_2p_4(-q_1q_3 + q_1q_4 + q_2q_3 - q_2q_4) + p_3p_4(q_1q_2 - q_1q_4 - q_2q_3 - \\ q_2 + q_3q_4 + q_4) + \\ p_1(-q_1q_2 + q_1q_4 + q_1) + p_2(q_2q_3 - q_2 - q_3q_4 - q_3 + q_4 + 1) + p_3(q_1q_2 - q_2q_3 - q_2 + q_3 - q_4) + p_4(-q_1q_4 + q_2 + \\ q_3q_4 - q_3 + q_4 - 1) + \\ \qquad q_2 - q_4 - 1 \end{array}} \right) \tag{5.24}$$

Let us consider the numerator of the $u_q(p)$. The cross product terms p_ip_j are given by,

$$\begin{aligned} & p_1p_2(q_1q_2 - 5q_1q_4 - q_1 - q_2q_3 + 5q_3q_4 + q_3) + p_1p_3(-q_1q_3 + q_2q_3) + p_1p_4(5q_1q_3 - 5q_3q_4) + p_3p_4(-3q_2q_3 + 3q_3q_4) + \\ & \qquad p_2p_3(-q_1q_2 + q_1q_3 + 3q_2q_4 + q_2 - 3q_3q_4 - q_3) + p_2p_4(-5q_1q_3 + 5q_1q_4 + 3q_2q_3 - 3q_2q_4 + 2q_3 - 2q_4). \end{aligned}$$

This can be re written in a matrix format given by (B.25).

$$(p_1, p_2, p_3, p_4)\frac{1}{2}\begin{bmatrix} 0 & -(q_1 - q_3)(q_2 - 5q_4 - 1) & q_3(q_1 - q_2) & -5q_3(q_1 - q_4) \\ -(q_1 - q_3)(q_2 - 5q_4 - 1) & 0 & (q_2 - q_3)(q_1 - 3q_4 - 1) & (q_3 - q_4)(5q_1 - 3q_2 - 2) \\ q_3(q_1 - q_2) & (q_2 - q_3)(q_1 - 3q_4 - 1) & 0 & 3q_3(q_2 - q_4) \\ -5q_3(q_1 - q_4) & (q_3 - q_4)(5q_1 - 3q_2 - 2) & 3q_3(q_2 - q_4) & 0 \end{bmatrix}\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} \tag{5.25}$$

Similarly, the linear terms are given by,

$$p_1(-q_1q_2+5q_1q_4+q_1)+p_2(q_2q_3-q_2-5q_3q_4-q_3+5q_4+1)+p_3(q_1q_2-q_2q_3-3q_2q_4-q_2+q_3)+$$
$$p_4(-5q_1q_4+3q_2q_4+5q_3q_4-5q_3+2q_4).$$

and the expression can be written using a matrix format as (B.26).

$$(p_1,p_2,p_3,p_4)\begin{bmatrix} q_1(q_2-5q_4-1) \\ -(q_3-1)(q_2-5q_4-1) \\ -q_1q_2+q_2q_3+3q_2q_4+q_2-q_3 \\ 5q_1q_4-3q_2q_4-5q_3q_4+5q_3-2q_4 \end{bmatrix} \tag{5.26}$$

Finally, the constant term of the numerator, which is obtained by substituting $p = (0,0,0,0)$, is given by (B.27).

$$q_2-5q_4-1 \tag{5.27}$$

Combining equations (B.25), (B.26) and (B.27) gives that the numerator of $u_q(p)$ can be written as,

$$\frac{1}{2}p\begin{bmatrix} 0 & -(q_1-q_3)(q_2-5q_4-1) & q_3(q_1-q_2) & -5q_3(q_1-q_4) \\ -(q_1-q_3)(q_2-5q_4-1) & 0 & (q_2-q_3)(q_1-3q_4-1) & (q_3-q_4)(5q_1-3q_2-2) \\ q_3(q_1-q_2) & (q_2-q_3)(q_1-3q_4-1) & 0 & 3q_3(q_2-q_4) \\ -5q_3(q_1-q_4) & (q_3-q_4)(5q_1-3q_2-2) & 3q_3(q_2-q_4) & 0 \end{bmatrix}p^T+$$
$$\begin{bmatrix} 0 & -(q_1-q_3)(q_2-5q_4-1) & q_3(q_1-q_2) & -5q_3(q_1-q_4) \\ -(q_1-q_3)(q_2-5q_4-1) & 0 & (q_2-q_3)(q_1-3q_4-1) & (q_3-q_4)(5q_1-3q_2-2) \\ q_3(q_1-q_2) & (q_2-q_3)(q_1-3q_4-1) & 0 & 3q_3(q_2-q_4) \\ -5q_3(q_1-q_4) & (q_3-q_4)(5q_1-3q_2-2) & 3q_3(q_2-q_4) & 0 \end{bmatrix}p+q_2-5q_4-1$$

and equivalently as,

$$\frac{1}{2}pQp^T+cp+a$$

where $Q \in \mathbb{R}^{4\times 4}$ is a square matrix defined by the transition probabilities of the opponent q_1,q_2,q_3,q_4 as follows:

$$Q=\begin{bmatrix} 0 & -(q_1-q_3)(q_2-5q_4-1) & q_3(q_1-q_2) & -5q_3(q_1-q_4) \\ -(q_1-q_3)(q_2-5q_4-1) & 0 & (q_2-q_3)(q_1-3q_4-1) & (q_3-q_4)(5q_1-3q_2-2) \\ q_3(q_1-q_2) & (q_2-q_3)(q_1-3q_4-1) & 0 & 3q_3(q_2-q_4) \\ -5q_3(q_1-q_4) & (q_3-q_4)(5q_1-3q_2-2) & 3q_3(q_2-q_4) & 0 \end{bmatrix},$$

$c \in \mathbb{R}^{4\times1}$ is similarly defined by:

$$c = \begin{bmatrix} q_1 (q_2 - 5q_4 - 1) \\ -(q_3 - 1)(q_2 - 5q_4 - 1) \\ -q_1q_2 + q_2q_3 + 3q_2q_4 + q_2 - q_3 \\ 5q_1q_4 - 3q_2q_4 - 5q_3q_4 + 5q_3 - 2q_4 \end{bmatrix},$$

and $a = -q_2 + 5q_4 + 1$.

The same process is done for the denominator. □

5.6.2 Proof of Theorem 3

The utility $u_q(p)$ is non concave and neither are it's numerator or denominator. Furthermore, the denominator is not always strictly positive.

Proof. The utility $u_q(p)$ is non concave because the concavity condition fails for at least one pair of points see Appendix B.7.1.

Furthermore, regarding the numerator and denominator of $u_q(p)$ in **Anton2014** it is stated that a quadratic form will be concave if and only if it's symmetric matrix is semi-negative definite. A matrix A is semi-negative definite if:

$$|A|_i \leq 0 \text{ for } i \text{ is odd and } |A|_i \geq 0 \text{ for } i \text{ is even,} \tag{5.28}$$

where $|A|_i$ is the eigenvalues of the submatrix A_i.

For (B.6), neither $\frac{1}{2}pQp^T + cp + a$ or $\frac{1}{2}p\bar{Q}p^T + \bar{c}p + \bar{a}$ are concave because for an even $i = 2$:

$$|Q|_2 = -(q_1 - q_3)^2 (q_2 - 5q_4 - 1)^2 \text{ and}$$
$$|\bar{Q}|_2 = -(q_1 - q_3)^2 (q_2 - q_4 - 1)^2$$

are negative.

Moreover, for a quadratic to be strictly positive it has to be positive definite. A quadratic form is positive definite iff every eigenvalue of is positive, however, $\frac{1}{2}p\bar{Q}p^T + \bar{c}p + \bar{a}$ is not positive definite because:

$$|\bar{Q}|_2 = -(q_1 - q_3)^2 (q_2 - q_4 - 1)^2$$

is negative. □

5.6.3 Proof of Lemma 4

Proof. The optimal behaviour of a memory-one strategy player $p^* \in \mathbb{R}^4_{[0,1]}$ against a set of N opponents $\{q^{(1)}, q^{(2)}, \dots, q^{(N)}\}$ for $q^{(i)} \in \mathbb{R}^4_{[0,1]}$ is established by:

$$p^* = \text{argmax}\left(\sum_{i=1}^{N} u_q(p)\right), \ p \in S_q,$$

where S_q is given by (B.14).

The optimisation problem of (B.13) can be written as:

$$\begin{aligned} \max_p : & \sum_{i=1}^{N} {u_q}^{(i)}(p) \\ \text{such that} : p_i & \leq 1 \text{ for } \in \{1,2,3,4\} \\ -p_i & \leq 0 \text{ for } \in \{1,2,3,4\} \end{aligned} \tag{5.29}$$

The optimisation problem has two inequality constraints and regarding the optimality this means that:

- either the optimum is away from the boundary of the optimization domain, and so the constraints plays no role;
- or the optimum is on the constraint boundary.

Thus, the following three cases must be considered:

Case 1: The solution is on the boundary and any of the possible combinations for $p_i \in \{0,1\}$ for $i \in \{1,2,3,4\}$ are candidate optimal solutions.

Case 2: The optimum is away from the boundary of the optimization domain and the interior solution p^* necessarily satisfies the condition $\frac{d}{dp}\sum\limits_{i=1}^{N} u_q(p^*) = 0$.

Case 3: The optimum is away from the boundary of the optimization domain but some constraints are equalities. The candidate solutions in this case are any combinations of $p_j \in \{0,1\}$ and $\frac{d}{dp_k}\sum\limits_{i=1}^{N} u_q^{(i)}(p) = 0$ forall $j \in J$ & $k \in K$ forall J, K where $J \cap K = \emptyset$ and $J \cup K = \{1,2,3,4\}$.

Combining cases 1-3 a set of candidate solution is constructed as:

$$S_q = \left\{ p \in \mathbb{R}^4 \;\middle|\; \begin{array}{l} \bullet \quad p_j \in \{0,1\} \quad \text{and} \quad \dfrac{d}{dp_k}\displaystyle\sum_{i=1}^{N} u_q^{(i)}(p) = 0 \quad \text{for all} \quad j \in J \quad \& \quad k \in K \quad \text{for all} \quad J, K \\ \qquad\qquad\qquad \text{where} \quad J \cap K = \emptyset \quad \text{and} \quad J \cup K = \{1,2,3,4\}. \\ \bullet \quad p \in \{0,1\}^4 \end{array} \right\}$$

This set is denoted as S_q and the optimal solution to (B.13) is the point from S_q for which the utility is maximised.

□

5.7 Further Examples

5.7.1 Example of non concavity for $u(p)$

A function $f(x)$ is concave on an interval $[a,b]$ if, for any two points $x_1, x_2 \in [a,b]$ and any $\lambda \in [0,1]$,

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda) f(x_2). \tag{5.30}$$

Let f be $u_{(\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{3})}$. For $x_1 = (\frac{1}{4},\frac{1}{2},\frac{1}{5},\frac{1}{2}), x_2 = (\frac{8}{10},\frac{1}{2},\frac{9}{10},\frac{7}{10})$ and $\lambda = 0.1$, direct substitution in (B.30) gives:

$$u_{(\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{3})}\left(0.1\left(\frac{1}{4},\frac{1}{2},\frac{1}{5},\frac{1}{2}\right) + 0.9\left(\frac{8}{10},\frac{1}{2},\frac{9}{10},\frac{7}{10}\right)\right) \geq 0.1 \times u_{(\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{3})}\left(\left(\frac{1}{4},\frac{1}{2},\frac{1}{5},\frac{1}{2}\right)\right) + 0.9 \times u_{(\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{3})}\Big(\Big(\frac{8}{10},\frac{1}{2},\frac{9}{10},$$

$$1.485 \geq 0.1 \times 1.790 + 0.9 \times 1.457 \Rightarrow$$

$$1.485 \geq 1.490$$

which can not hold. Thus $u_{(\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{3})}$ is not concave.